

## Self-Saboteurs and Ethical Relationships<sup>1</sup>

**Abstract:** Common sense morality tells us we should help our friends, family, and loved ones when they suffer. However, some people are hell-bent on being unhappy, and will do everything in their power to ensure their own suffering. This creates a dilemma about what we ought to do for someone who puts themselves in pain and refuses to budge. In this paper, I discuss mechanisms of and motivations for self-sabotaging behavior. I then turn to the ethical complications of these cases, focusing on the risk of becoming complicit in another's self-sabotage, the acceptable limits of caring for a self-saboteur, and the permissibility of paternalistic interference. I argue that while there is some degree of leeway we can permissibly give to meet another's needs—including submitting to their low-stakes manipulation—doing so poses the risk of damaging the relationship. While interference in another's self-sabotage for his own good may seem justified, I argue that this approach is also a morally problematic denial of the self-saboteur's agency. Instead, I offer an alternative route between complicity and interference: carers ought to try to maintain a genuine relationship build on the honest recognition of each other's reasons. This will mean the carer should talk to the self-saboteur about her own concerns and the strain the self-sabotage puts on the relationship while also appreciating that the self-saboteur may have legitimate reasons to suffer in an effort to find a resolution that both parties can accept.

**Keywords:** Ethics; Care Ethics; Moral Psychology; Self-Sabotage; Paternalism

### 1. Introduction

Many, if not most, of our relationships are non-ideal. Friends can force us to make impossible choices; family members can make unfair demands of us. And some people are just difficult. It also seems true that at least some relationships are worth preserving despite these challenges. After all, if perfection were a requirement of close relationships, we probably wouldn't have many. This means that insofar as our relationships give rise to ethical demands, these demands may be messy, imperfect, and tricky to navigate—yet important to get right.

In this paper, I consider some of the ethical demands that arise in one type of non-ideal close relationship: caring for a self-saboteur. Common sense morality tells us that we ought to help the people we love when they are suffering if we can. But some people are hell-bent on being unhappy, and do everything in their power to ensure their own suffering. This creates a

---

<sup>1</sup> This paper was published in *Social Theory and Practice*. LaGuardia-LoBianco, A. W. 2019. "Self-Saboteurs and Ethical Relationships." *Social Theory and Practice*, 45(2): 249-285. DOI: 10.5840/soctheorpract20197259.

[https://www.pdcnet.org/soctheorpract/content/soctheorpract\\_2019\\_0999\\_7\\_2\\_59](https://www.pdcnet.org/soctheorpract/content/soctheorpract_2019_0999_7_2_59)

© Copyright by Social Theory and Practice

dilemma about what we ought to do for someone who puts themselves in pain and refuses to budge: how do we care for a friend or family member who *wants* to suffer?

The question becomes difficult to answer when we realize that in these cases there is a tension between what self-saboteurs want for themselves and what another ought to provide for them. Specifically, because the self-saboteur explicitly wants to suffer, and their friend or family member has some duty to alleviate that suffering, discharging this duty will be at odds with the desires of the self-saboteur. This presents a question of whether a carer should interfere with a self-saboteur for the sake of the latter's well-being and against her wishes—that is, if paternalistic interference is permissible. Additionally, the strains of caring for a self-saboteur—in which the self-saboteur may be determined to ruin any help they receive, yet continue to need that help—raise questions of the limits of care in these cases. That a self-saboteur may leverage their suffering against a carer complicates these limits even further.

In order to unpack these ethical demands of caring for a self-saboteur, it will be crucial to understand the psychology behind self-sabotage. To that end, I first discuss different mechanisms of self-sabotaging behavior. I then turn to the reasons one might have for behaving this way, outlining a number of incentives for self-sabotage. Finally, I discuss the ethical complications of these cases, focusing on the limitations of caring for a self-saboteur when this care is manipulatively solicited and exhaustively demanding and the question of whether it is permissible to paternalistically interfere with another's self-sabotage. I argue that while there is some degree of leeway we can permissibly give to meet another's needs—including submitting to their low-stakes manipulation—doing so poses the risk of damaging the relationship. While interference in another's self-sabotage for her own good may seem justified, I argue that this approach is also a morally problematic denial of the self-saboteur's agency. Instead, I offer an

alternative route between complicity and interference: carers ought to try to maintain a genuine relationship build on the honest recognition of each other's reasons. This means the carer should talk to the self-saboteur about her own concerns and the strain the self-sabotage puts on the relationship while also appreciating that the self-saboteur may have legitimate, and important, reasons to suffer, in an effort to find a resolution that both parties can accept.

## 2. Self-Sabotage

We do not always do what will make us better off. Sometimes we inadvertently act in ways that result in our own suffering. But some deliberately undertake a campaign of self-sabotage in order to make their lives difficult. Consider the following example of a patient from psychiatrist

Arnold Cooper:

A man tells his wife that he is not sure what time he is coming home that evening. He arrives earlier than usual and feels disappointed that his wife is not home. When she arrives a half-hour later he berates her for not having been there. She feels unjustly accused and is angry in turn, another evening is ruined, and he mopes through the evening obsessing about how he could have made such a bad marriage.<sup>2</sup>

Cooper's patient demonstrates self-defeating or self-sabotaging behavior: acts of destroying or undermining the circumstances in one's life to one's own detriment.<sup>3</sup> Paradigm examples include procrastination, inciting conflict with others, or binge eating,<sup>4</sup> but the cases of concern here include any deliberate, routine behavior that detrimentally interferes with one's life

---

<sup>2</sup> Cooper, A. M. (1993). Psychotherapeutic approaches to masochism. *Journal of Psychotherapy Practice and Research*, 2(1): 51-63. p. 56-57.

<sup>3</sup> Klein, D. N. & Vocisano, C. (1999). Depressive and self-defeating (masochistic) personality disorder. In T. Millon, P. H. Blaney, & R. D. Davis (Eds.), *Oxford textbook of psychopathology* (653—673). New York: Oxford University Press.

<sup>4</sup> Sherry, S. B., Stoeber, J., & Ramasubbu, C. (2016). Perfectionism explains variance in self-defeating behaviors beyond self-criticism: evidence from a cross-national sample. *Personality and Individual Differences*, 95, 196—199.

to a greater or lesser degree.<sup>5</sup> We can understand these behaviors with the criteria used to describe self-defeating personality disorder. Although never officially incorporated into the main body of the DSM, self-defeating personality disorder, formerly masochistic personality disorder, illustrates this tendency toward self-sabotage.<sup>6</sup> ‘Masochistic’ here does not refer to the link between physical pain and sexual satisfaction,<sup>7</sup> but rather indicates “psychological phenomena in which individuals engage in activities that lead to their otherwise avoidable suffering.”<sup>8</sup> Self-defeating personality disorder is the habituation of this behavior, part of “a lifelong pattern” of self-defeating behaviors of choosing harmful paths of action while avoiding beneficial ones.<sup>9</sup>

I appeal to self-defeating personality disorder only as a useful apparatus to understand ethically relevant self-sabotage, but do not import the implications that self-sabotage is necessarily pathological. Rather, self-sabotaging behavior is better understood along a spectrum than as a fixed category. Psychologist Roy Baumeister, for instance, notes that this behavior is common in the general population, not only in those with mental illness,<sup>10</sup> and Cooper writes, “All of us engage in this technique [of self-sabotage] at some time or other, to some degree, but the masochist makes it his life’s work.”<sup>11</sup> At one extreme of the self-sabotage spectrum are

---

<sup>5</sup> I am concerned with behaviors that detrimentally impact both the objective conditions of one’s life, like one’s career or relationships (what some prefer to call ‘well-being’) and, as a result, one’s psychological or emotional state (one’s ‘happiness’ or misery). These agents sabotage both their objective situations and their (related) emotional health with the same self-destructive act.

<sup>6</sup> Self-defeating personality disorder was included in the DSM-III-R under the appendix marked for further study, but was excluded from the DSM-IV. Whether or not SDPD is properly included as a psychopathology is immaterial for my argument; my question is how to ethically respond to self-saboteurs, and self-defeating personality disorder just gives a helpful description of such behavior. As I’ll argue, there is a spectrum of self-defeating behavior that ranges from typical to pathological. While I don’t hold that all self-defeating behavior is pathological, I don’t exclude cases that are from my discussion. Op. cite note 2.

<sup>7</sup> First discussed by von Krafft-Ebing, based on the novel by Leopold Sacher-Masoch. von Krafft-Ebing, R. (1886). *Psychopathia sexualis, with especial reference to antipathic sexual instinct: A medico-legal study*. Philadelphia, PA: F. A. Davis; Sacher-Masoch, L. (1870). *Venus in furs*. Waiheke Island: Floating Press.

<sup>8</sup> Bekes, V., Perry, J.C. & Robertson, B.M. (2016). Psychological masochism: a systematic review of the literature on conflicts, defenses, and motives. *Psychotherapy Research*, doi: 10.1080/10503307.2016.1189618. p. 2.

<sup>9</sup> Op. cite note 2, p. 664.

<sup>10</sup> Baumeister, Roy F. (1991). *Escaping the Self: Alcoholism, Spirituality, Masochism, and Other Flights from the Burden of Selfhood*. New York: Basic Books. p. xi

<sup>11</sup> Op. cite note 1, p. 57.

behaviors that overlap with or are caused by pathological conditions, such as depression, addiction, borderline personality disorder, or SDPD (if such a category is ever substantiated); at the other extreme are common, non-pathological (yet still deliberate and habitual) behaviors.<sup>12</sup> This means that population of self-saboteurs as I understand it is actually quite wide, as it can include both pathological and non-pathological self-sabotage (so long as they meet the conditions discussed below).<sup>13</sup> Rather than confining my discussion to one or the other end of the spectrum, I take the whole range of cases to count as ‘self-sabotage.’ This is partly because reasons for self-sabotage are often plural and overdetermined, so that it is difficult to pin any instance of self-sabotage to a definite point on the spectrum. But more importantly, this wide scope is justified since (as I’ll argue further below) both pathological and non-pathological self-sabotage complicate ethical relationships, if in slightly different ways. And this means that the whole gamut of cases of self-sabotage is relevant for ethical analysis.

Using the framework of masochistic personality disorder as a start, we can identify various ways in which one can self-sabotage (understood as deliberately behaving in self-destructive ways).<sup>14</sup> Self-sabotage can be accomplished through acts of commission (i.e. creating or putting oneself in a bad situation, say by choosing an unhealthy coping mechanism or ruining a happy relationship), acts of omission (e.g. neglecting to secure resources when one knows they are needed and available, avoiding one’s support system), or some combination therein. As I’ll argue in the next section, motivations for self-sabotage are plural, and can consist of either or both an intrinsic desire for self-destruction and desires for some instrumental gain that comes

---

<sup>12</sup> Assuming that ‘pathological’ and ‘non-pathological’ populations are reliably distinct, which I’m dubious of.

<sup>13</sup> I sincerely thank an anonymous referee for pushing me to rethink an earlier exclusion of pathological cases of self-sabotage.

<sup>14</sup> One way to understand ‘self-destructive’ is as the antipode of ‘self-preservation,’ as Baumeister does: “Psychologists have long been fascinated with the human penchant for self-destructiveness. From procrastination to suicide, people seem to do things that will bring them distress and loss...What could be more basic than self-preservation? Yet people often behave in ways incompatible with self-preservation: harming themselves, compromising their values and goals, or defeating their projects and ambitions.” *Op. cite* note 9, p. 39.

about through self-sabotage.<sup>15</sup> For now, the crucial point is that there are a variety of behaviors that count as self-sabotage, and that a self-saboteur may engage in any number and combination of them.

Consider some standard examples of self-sabotage. Self-sabotaging acts of commission include being “drawn to situations or relationships that will result in suffering.”<sup>16</sup> Frederic Kass, who helped design the criteria for self-defeating personality disorder, describes the self-saboteur as, for example, “a person who repeatedly chooses to enter relationships with others who turn out to be alcoholics or who are in some way emotionally unavailable. They may continually enter employment situations where despite their excellent skills, their work is unrecognized and they are underpaid.”<sup>17</sup> These sorts of behaviors bind an individual to the suffering-inducing situation, both guaranteeing these conditions maintain and justifying them. Or, one may “ruin or spoil pleasurable situations,” as in Cooper’s example of the man who picks a fight with his wife.<sup>18</sup> Others can maintain their misery by interfering with its relief. One example of this is the tendency to sabotage therapy even when it proves effective, as with the case of Jack. Though he entered therapy at his wife’s behest, “...the idea that he needed therapy was not completely alien to him. He sensed that things were wrong with him, that he was on a bad road...He resisted every step of the way, showing up late to appointments, skipping the next session with no notice just when it seemed as if we were getting somewhere, playing all sorts of games to maintain his

---

<sup>15</sup> As will be explained in more detail below, this means I take both of the following types of cases to count as self-sabotage (or equivalently, ‘deliberate self-sabotage’): behavior that has self-destruction as its primary aim, and behavior for which self-destruction is a secondary aim, but is instrumental to securing the primary aim (of, for instance, attention from others). Any given case of self-sabotage won’t be so neatly diagnosed, I suspect. Motivations for any behavior are often overdetermined, plural, and subject to change. By applying my definition of ‘self-sabotage’ to both intrinsic and instrumental motivations for self-destruction (and their intersections), I mean to better capture the messy reality of self-sabotage. I thank an anonymous referee for suggesting this clarification.

<sup>16</sup> Op. cite note 8.

<sup>17</sup> Kass, F. (1987). In response: new controversial diagnoses: self-defeating and sadistic personality disorders. *Jefferson Journal of Psychiatry*. 5(1), 92-96. p. 93.

<sup>18</sup> Huprich, S. K. & Nelson, S. M. (2014). Malignant self-regard: accounting for commonalities in vulnerably narcissistic, depressive, self-defeating, and masochistic personality disorders. *Comprehensive Psychiatry*, 55, 989—998. p. 990. Huprich and Nelson call these sorts of actions “the hallmark of self-defeating behavior,” (p. 990).

sense of superiority to the therapeutic process.”<sup>19</sup> Certain omissive acts can also count as self-sabotage. Avoidance behaviors like failing to seek beneficial therapy, refusing to follow up with health-care providers,<sup>20</sup> ignoring bill collectors, or repeated truancy are ways of securing further difficulty by failing to act in one’s best interest.

Having given example of self-sabotaging behavior, it may help to describe some contrast cases in order to get at a clearer definition. First, one-off, single incidents of self-sabotage should be set aside. A single self-undermining act does not necessarily make one a self-saboteur, as anyone may occasionally act on an errant self-loathing impulse or fail to do what is in her best interest. Rather, deliberate self-sabotage involves the habituated pattern of self-destructive tendencies (whether this is a result of pathology or not), not isolated acts of self-undermining.<sup>21</sup>

Self-sabotage also involves choosing a painful option where better options are available, so the second case-type of contrast is one where agents have no other option but to self-sabotage. Certain environments force bad options on agents by design. Systemic oppression and moral dilemmas, for instance, characteristically beget double binds, in which the best option an agent can choose is still a terrible one. In these cases, an agent’s situation restricts her choices such that suffering is effectively her only choice. Though she may have chosen an option that ultimately

---

<sup>19</sup> Rosner, S. & Hermes, P. (2006). The self-sabotage cycle: why we repeat behaviors that create hardships and ruin relationships. Westport: Praeger. p. 123-124.

<sup>20</sup> Op. cite note 9.

<sup>21</sup> The proposed diagnostic criteria for self-defeating personality disorder “should...not be made when the patient has a clinical depression or during the recovery period from depression, a state in which it is virtually impossible to discern the masochistic trait” (Mackinnon, R. A., Michaels, R., Buckley, P.J. (2006). The masochistic patient. In *The Psychiatric Interview in Clinical Practice*, 2<sup>nd</sup> ed. Washington: American Psychiatric Publishing. 205-228. p. 210). This suggests that, at least as a potential personality disorder (PD) is concerned, self-sabotage is distinct from the typical depressive behavior. However, since I’m not concerned with self-sabotage as a PD, but rather only using the proposed criteria to partially inform my understanding of self-sabotage, both cases of self-sabotage that are motivated primarily by self-destructive aims (which could, therefore, be used to diagnose SDPD) and those which are the result of depressive thinking (which do not bear on SPDP) are relevant, in different ways, for the duties of those the self-saboteur is in a relationship with. Here, then, I differ from clinical opinions on self-sabotage.

harms her, this is because the odds were stacked against her, and not indicative of either an intrinsic or instrumental desire to suffer.<sup>22</sup>

Nor does an adaptive response to a terrible situation count as self-sabotage. An oppressive or abusive environment can force an agent into a pattern of self-defeating behavior.<sup>23</sup> Oppression characteristically does this by shaping oppressed agents to continually make choices that maintain their own oppression.<sup>24</sup> For instance, a woman may choose to be a housewife rather than enter the workforce, thereby remaining financially dependent on her husband. Such behavior is, in a strict sense, self-defeating as it ultimately contributes to a system that harms the agent. But properly understood, an oppressed agent does not bring harm upon herself: she is stuck in a rigged system of unfair choices *because* she can't escape harm. Her choice may be detrimental to her, but this does not mean she chose it because she wanted to harm herself. Similarly, families who live in poverty may make decisions that continue the cycle of poverty for their children, say by not saving for a college education. But it is fallacious to conclude that they thereby *chose* to be in poverty. Oppression works to perpetuate itself through oppressed agents, but this does not mean the agent desires or chooses to be oppressed.

In this same vein, some self-sabotage is an adaptive response to abuse, and is in no way identical with deliberate self-sabotage or masochism. One may stay with one's abuser for any number of reasons, but this behavior is distinct from self-sabotage driven by a punitive need to continue suffering. Feminists have argued against the problematic notion that women who

---

<sup>22</sup> These cases can be much more complicated than suggested above. Though one's options may be constrained, it may still be the case that some act is chosen because the agent's goal is to self-destruct. Thus, double binds do not entirely preclude the possibility that one is acting self-destructively. Rather, my point in bringing up this contrast is to argue that deliberate self-sabotage is not *merely* acting against one's best interest, since there are plenty of ways an agent can do this without having a motivation toward self-sabotage.

<sup>23</sup> I thank a referee for pushing me to develop this point.

<sup>24</sup> For instance, see Cudd, A. E. 2006. *Analyzing oppression*. New York: Oxford University Press.

remain in abusive relationships are suffering from masochistic personality disorder.<sup>25</sup> Behaviors that may appear self-defeating are often the result of adapting to abuse in an effort to survive, rather than the result of a personality disorder. It would thus be erroneous, and dangerous, to conclude that an abused victim wants to be abused based on the sabotaging pattern of her actions.

A third type of case that is worth distinguishing from self-sabotage is that in which agents keep making bad decisions for themselves but not because they are seeking to suffer. Rather, these agents make poor choices, misjudge what will ensure the best outcome, accidentally choose a painful path of action, or fail to learn from their mistakes.<sup>26</sup> And some folks may just get in their own way. Charlie Brown earnestly believes that Lucy will hold the football *this time*. But she never does, and because he never adjusts his expectations or behavior, he contributes to his own continued embarrassment.<sup>27</sup> Pitiful as Charlie Brown is, though, he does not *want* to suffer. Strictly speaking, his own actions do lead to his suffering, but not because they are aimed at it. What is missing from these ‘poor schmuck’ cases, and what distinguishes them from deliberate self-sabotage, is some motivation to suffer that drives the sabotage. Instead, they result from a failure to learn from mistakes or an inability to get out of one’s way. As such, these cases, too, are bracketed from the discussion.

Taking these contrast cases into consideration, we can define self-sabotage as follows.

First, these behaviors are part of a repeated pattern, an expression of a disposition or habituated

---

<sup>25</sup> See Rosewater, L. B. (1987). A critical analysis of the proposed self-defeating personality disorder. *Journal of Personality Disorders*, 1(2),190-195. Indeed, patterns of behavior that are a response to anticipated abuse are explicitly exempted from the diagnostic criteria of the proposed SDPD. Op. cite notes 2 and 16.

<sup>26</sup> I thank an anonymous referee for pushing me to clarify this point.

<sup>27</sup> Of course one can speculate about Charlie Brown’s psychology, and in general, we might suspect of someone who routinely self-sabotages out of ostensible ignorance that, on some level, they know what they are doing and are quite content to continue. But this isn’t always the case, and the contrast highlighted above is between ‘poor schmuck’ self-sabotage that results from poor judgment or failure to learn from mistakes versus deliberate self-sabotage that results from a desire to suffer. I thank a referee for bringing up this point.

action that may or may not be pathological. Second, acts of self-sabotage are chosen despite the existence of viable alternatives that would not cause (as much) suffering. Third, self-sabotage is deliberate. Self-saboteurs act with the goal of making themselves suffer either as the intrinsic or instrumental goal and know that their action will achieve this effect. Neither is acting under a false belief or accidental knowledge about the effects of their actions an instance of self-sabotage.<sup>28</sup> Rather, one is reasonably aware of the suffering that will result from their behaviors. These actions may be motivated by a conscious or non-conscious desire to suffer, which will be considered in the next section. Call this behavior of deliberately and habitually choosing self-undermining acts where better options exist *deliberate self-sabotage*.

It is worth stressing that the distinctions between deliberate self-sabotage and other cases like abusive adaptation or ill-informed decision-making are made here only for the sake of clarity. In reality, agents' self-sabotage can have many different causes or combinations of causes, deliberate and otherwise. Agents won't cleanly fit into one or the other category because causes for these behaviors are overdetermined by desires to suffer, influences from the environment, failures to learn from mistakes, or other motivations and external pressures.<sup>29</sup> So, since the motivations for deliberate self-sabotage are plural and overdetermined, this category includes cases in which self-destruction is the intrinsic or primary aim, cases in which self-destruction is a secondary or instrumental aim, cases in which self-sabotage is in some way linked to pathological behavior, and non-pathological cases.

---

<sup>28</sup> The point is just that the self-saboteur knows what they are getting into: they aren't acting under false assumptions, misinformation, mistakenly or accidentally, but rather have a reasonable ability to discern that doing X will have effect Y, and this effect is adversarial to them. This awareness can exist on a continuum, as it can be more or less acute over time, or alight on only some behaviors and overlook others. But the relevant point is that a deliberate self-saboteur does not sabotage herself out of happenstance or accident. While the  *motive*  for her self-sabotage may be conscious or non-conscious, the detrimental results of her actions are apparent to her when she acts.

<sup>29</sup> I wish to thank a referee for bringing up this point.

I have made the category of deliberate self-saboteurs intentionally wide because all of these cases complicate ethical relationships, though they may do so in slightly different ways. In particular, the different motivations, or incentives, for deliberate self-sabotage will partly determine the ethical difficulties of caring for a self-saboteur. I discuss some of these motivations in the following section.

### 3. Incentives for Deliberate Self-Sabotage

A crucial feature of deliberate self-sabotage is some investment in suffering, some reason that makes (continued) suffering worthwhile enough for the agent that she actively creates the conditions for it. As mentioned above, any given case of deliberate self-sabotage can involve overdetermined motivations, and these can include both intrinsic desires to suffer, in which self-sabotage is the primary aim, as well as instrumental desires to suffer, where self-sabotage secures some secondary aim. Since both primary and secondary motivations for self-sabotage will bear on implications for caring relationships in different ways, both sorts of case will be considered here.<sup>30</sup> I'll discuss intrinsic motivations for self-sabotage, instrumental motivations for self-sabotage, and a third, 'hybrid' case type in which personal suffering is desired as a *necessary* part of some additional end. Any of these motivations, or combinations therein, can drive deliberate self-sabotage.<sup>31</sup>

First, consider cases in which self-sabotage is the primary aim—when an agent acts against their own best interest because they want to suffer for its own sake. The notion of an

---

<sup>30</sup> I thank an anonymous referee for suggesting I clarify this point.

<sup>31</sup> Note that a given agent's motivation for self-sabotage will not be so neatly separated into these categories, nor will there necessarily be one single desire that spurs the behavior. Rather, the reason(s) an agent has to suffer can overlap and change.

intrinsic desire to suffer may seem pathological, if not outright paradoxical. The puzzle resides in *wanting* to suffer, and it is this desire—for any perceived ‘bad’—that many philosophers have decried as impossible. As a matter of psychological necessity, the argument goes, we only desire what is good for us, so any desire to suffer must rest on a misunderstanding of its badness or else is actually desired for the presumed good it will bring.<sup>32</sup> Insofar as a primary incentive to self-sabotage involves an intrinsic desire to suffer, it may likewise seem incoherent. Therefore, some claim, there can be no inherent attraction to suffering outside of pathology.

However, this explanation sacrifices an accurate account of human psychology for theoretical neatness. As Michael Stocker argues, the good need not attract us and “[a] bad thing (harm) can be the proper object of desire.”<sup>33</sup> That is, we can be unmoved to benefit ourselves, and harmfulness can be attractive. A desire for harm is thus all the explanation required for self-sabotage: “Just as there are desires and appetites directed at harming others, there are desires and appetites directed at harming oneself. In...self-directed modes of disgust, hatred, guilt, shame, I may seek to humble, abase, or harm myself.”<sup>34</sup> David Velleman explains that this is because the attitude of desire does not aim at getting things correct—at desiring what is actually good—analogueous to the way that the attitude of belief aims at getting the truth right.<sup>35</sup> Rather, we can desire things under the description that they ought not be desired, and this explains perverse desires. Sometimes, what is attractive *is* the fact that something is bad, and reducing this desire to a mistaken conception of the good would miss this point.

So, an agent may self-sabotage simply because she wants to make herself miserable. Just as no further explanation is needed to explain a desire for some pleasure, so some cases can be

---

<sup>32</sup> For a survey of views of this sort, see Stocker, M. (1979). *Desiring the bad: an essay in moral psychology. The Journal of Philosophy*, 76, 738—753.

<sup>33</sup> Op. cite note 31, p. 748.

<sup>34</sup> Op. cite note 32.

<sup>35</sup> Velleman, J. D. 1992. “The Guise of the Good,” *Nous*, 26(1), 3-26. p. 15-17.

motivated by a desire for what's bad for oneself. Moreover, such a desire is by no means categorically pathological. As opposed to psychologists like Baumeister, who holds that *all* self-sabotage is really motivated by some secondary gain, I instead hold that there are a plurality of motivations that can drive self-sabotage.<sup>36</sup> This includes an intrinsic motivation for self-sabotage which, again contra Baumeister, I do not think is only possible in the presence of mental illness.<sup>37</sup>

While some self-sabotage because they are (mainly) driven by an intrinsic desire to suffer, others use self-sabotage to secure separate ends. The latter are cases of instrumental desires to suffer, whereby self-sabotage is a secondary aim that guarantees some other, primary aim. Psychologists Stanley Rosner and Patricia Hermes argue that self-sabotage can yield instrumental benefits to the agent, what they call “secondary gains.”<sup>38</sup> These secondary gains explain why sabotaging behaviors are so prone to repetition.<sup>39</sup> The risk of losing these gains is too high: “... more gratification is derived by perpetuating this state of affairs than by giving them up,” even if this would mean giving up suffering.<sup>40</sup> Some of these secondary gains are relational, concerning responses the self-saboteurs can bring about from others. I'll consider two such secondary gains here: securing a sympathetic connection with others and taking revenge on others (often under the guise of exacting punishment) for some wrong done to the self-saboteur.

The following case is a characteristic example of sabotaging oneself in order to elicit a particular reaction from others. Berglas and Baumeister recount the case of an adult patient,

---

<sup>36</sup> Op. cite note 9, p. 39.

<sup>37</sup> Op. cite note 35. For more on this debate over whether pain can be desired intrinsically or only as an instrumental end, see Warren, V. L. (1985) Explaining masochism. *Journal for the Theory of Social Behavior*, 15, 103—129.

<sup>38</sup> Op. cite note 18, p. 29.

<sup>39</sup> Warren argues that appealing to secondary gains is insufficient to explain masochism, and thus argues that one may come to desire pain as an essential part of some other end, like a stable self-conception (discussed below). Op. cite note 35.

<sup>40</sup> Op. cite note 37.

Jerry, who worked with his father at the family business.<sup>41</sup> As recounted by Berglas and Baumister,

Jerry reported that his dad never attended a single one of his choir recitals or art exhibitions from elementary school through college, and despite constant pleas, Jerry's father would never spend "nonbusiness" time with him. In fact, when Jerry attempted to interact with his dad outside the corporate offices, he was typically rebuked for wasting time and "acting like a child." As Jerry put it, "When I was performing for the company, I'd hear, 'Good body'; otherwise, not a word."<sup>42</sup>

In response, Jerry would repeatedly put himself in debt through risky gambling in order to gain attention from his otherwise neglectful father, who would then pay his debts. The sabotaging spiral was ultimately "psychologically rewarding" for Jerry since it gave him the attention of his otherwise neglectful father.<sup>43</sup> Jerry "discovered that by getting into gambling-related trouble—particularly with large outlays of money, his father's 'life preoccupation'—he could get the attention he yearned for from this man."<sup>44</sup> By putting himself in this troubling situation, Jerry achieved the attentive response from his father he desired.

In this vein, a self-saboteur may want to gain those responses that are predictably forthcoming from others when he is facing some difficulty and in pain: attention from others as they tend to his situation, sympathy as they comfort him, and care as they attempt to meet his needs.<sup>45</sup> In other words, self-sabotage can be a way to bring others close, to guarantee the attention and connection that one craves (even if this connection *also* involves anger towards the

---

<sup>41</sup> Berglas, S. and Baumeister, R.F. (1993). *Your Own Worst Enemy: Understanding the Paradox of Self-Defeating Behavior*. New York: Basic Books. p. 151-154.

<sup>42</sup> Op. cite note 40, p. 154. Interestingly, Beglas and Baumeister make a point to note that Jerry did not meet the criterion for a gambling addiction, arguing instead that Jerry's behavior is better explained as self-defeating desire for attention and revenge (p. 153). Though, this is not to suggest that, in general, those with addictions never engage in ethically significant self-sabotage.

<sup>43</sup> Op. cite note 41.

<sup>44</sup> Op. cite note 41.

<sup>45</sup> There are appropriate moral responses to suffering that involve the obligation to alleviate that suffering, but the benefit of concern here rather involves the socially sanctioned responses to suffering independent of their moral standing.

self-saboteur, as in Jerry’s case).<sup>46</sup> Some of these responses can be appealing to the self-saboteur in their own right. One resource for understanding this desire to connect with others is found in motivations for factitious disorders. Factitious disorders are “characterized by physical or psychological symptoms that are intentionally induced or feigned.”<sup>47</sup> They involve the creation or feigning of illness in oneself for the sake of being ill: the “presumed goal of factitious behavior is to assume the ‘sick role,’” and the attention and care that comes with it.<sup>48</sup>

To be clear, the appeal to factitious disorders is not to suggest an equivalency between self-sabotage and factitious disorders. Rather, it suggests only that the same gains that can motivate someone to feign an illness can also motivate someone to self-sabotage (though the latter acts are by no means feigned). One explanation for factitious disorders is “the powerful reward in the form of identity and recognition that the patient obtains in taking on the sick role.”<sup>49</sup> The reward of being recognized as an ill patient involves attention from doctors and caretakers, and while this attention can serve as its own reward, it may involve other perks.<sup>50</sup>

Lauren Slater describes the allure of illness in terms of being cared for. She writes: “When I was a girl I loved fevers and flues and the muzzy feeling of a head cold, all these states carrying with them the special counterments [sic] of illness...best of all, a distant mother coming to your bedside with tea.”<sup>51</sup> For Slater, illness represents an exception to parental neglect, a sure way to gain attention. Illness can be alluring when the care of a doctor represents more than health. A

---

<sup>46</sup> I thank an anonymous referee for suggesting I clarify this idea with the notion of ‘connection with others.’

<sup>47</sup> Iezzi, T., Duckworth, M. P., & Adams, H. E. (2002). Somatoform and factitious disorders. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive Handbook of Psychopathology*, Third Ed. (pp. 211—258). New York: Kluwer Academic Publishers. p. 211. Munchausen’s by proxy, in which someone induces illness in another for attention from healthcare workers (as when a parent makes their child ill) is another particularly disturbing factitious disorder, though it is worth mentioning to show the strength of the incentive to be an object of compassion, even indirectly.

<sup>48</sup> Op. cite note 46.

<sup>49</sup> Op. cite note 46, p. 247.

<sup>50</sup> Kropp, R. P. & Rogers, R. (1993). Understanding malingering: motivation, method, and detection. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 201—216). Guilford Press.

<sup>51</sup> Slater, L. (1998). *Prozac Diary*. Penguin Books, p. 21. I use Slater’s account as an example of an incentive to suffer – here, a desire to assume a ‘sick role’ – not as an instance of self-sabotage.

similar appeal in being an ill patient applies to being treated as a sufferer: gaining from others attention, sympathy, and efforts to help and heal. An agent can gain this sort of connection with others by displaying their self-sabotage in front of others and, anticipating their response, leveraging their suffering to that effect.<sup>52</sup>

An additional secondary gain from self-sabotage arises when agents believe they have been wronged by another and want to exact revenge on them. Self-sabotage can be an effective tool of revenge when wielded against those close to the saboteur: one hurts oneself in order to hurt another, in order to retaliate for the perceived wrong done to them.<sup>53</sup> Jerry's motives for self-sabotage again apply here (emphasizing the overdetermined nature of many cases of self-sabotage). In addition to gaining the attention he craved, Jerry's destructive gambling was also a way of retaliating against his father. According to Berglas and Baumeister, Jerry's father "was cold, rebuking, and hostile to his son...he was cruel to Jerry, who, as the youngest of five children, seemed perennially to get little of his father's time and attention."<sup>54</sup> Not only did Jerry's troubles force his father to acknowledge him, but they did so in a way that hurt his father. Jerry recounts of the effects of his gambling, "...it's not my pain that hurts...it's seeing the tortured look on my Dad's face and seeing him cry that gets me. I know I'm stabbing him in the gut with a knife when I gamble."<sup>55</sup> Putting his father through this pain was a way for Jerry to retaliate against him for past cruelty and neglect. Seeing another suffer can, to some, be worth the pain of suffering oneself.

---

<sup>52</sup> The same incentives that motivates one to make themselves suffer publicly can similarly motivate one to make their suffering public, though these are distinct cases and I leave the latter to the side for this discussion.

<sup>53</sup> Berglas and Baumeister call this type of self-defeating behavior "Pyrrhic revenge," in which a self-saboteur takes retaliatory actions against another at a "ruinous" cost to himself: "...Pyrrhic-revenge strategies provide psychological reinforcement or relief that in some way compensates significantly for the overt harm to the self." Op. cite note 40, p. 151.

<sup>54</sup> Op. cite note 40, p. 154.

<sup>55</sup> Op. cite note 40, p. 153

Note though that exacting revenge for a wrong is distinct from morally significant punishment for that wrong.<sup>56</sup> While one colloquial sense of the term ‘punishment’ seems to equate a wrongdoer’s pain with the redress of that wrong (as one might withhold affection to ‘punish’ a partner for some slight), morally significant punishment, which serves as a penalty for wrongdoing, seems to require more than the pain of the perpetrator. For instance, a public acknowledgment of the offense may also be necessary for legitimate moral punishment. It is an open question whether the offender’s pain is necessary for punishing them, but at the very least, revenge seems insufficient for punishment.

While ‘taking justice into one’s own hands’ by hurting another may not constitute genuine punishment (and further, may be inappropriate even if it *did* serve as morally significant punishment), it may seem that way to the self-saboteur. That is, a self-saboteur may conflate revenge and punishment, believing that hurting another (by hurting themselves) *is* fair and just punishment for the wrong committed. This sort of desire for revenge may be better construed as a desire for revenge-under-the-guise-of-punishment. Cooper argues that this is a particularly strong motivation for masochistic patients since letting go of the desire for perceived retribution can *feel* like a gross injustice.<sup>57</sup> This is evident in his patient’s experience of visiting their parents’ home:

...[I]t never occurs to [the patient’s parents] that they owe me an apology for what they put me through...I can’t bear the thought that they are getting away with it. They don’t even feel guilty. It made me very depressed and I ended up moping around. I could tell it bothered them, and I was glad...Finally, I picked a fight over what my mother served for lunch...[i]t was worth it to me to ruin the visit if it hurt them.<sup>58</sup>

For this patient, self-sabotage is desirable for the purported justice it serves, even though this ‘justice’ is really a thinly veiled revenge (note that this self-saboteur believes the wrong has not

---

<sup>56</sup> I thank an anonymous referee for suggesting that I explore this distinction.

<sup>57</sup> Op. cite note 1, p. 61.

<sup>58</sup> Op. cite note 1, p. 61

yet been addressed since they are still owed an apology). ‘Ruining the visit’ is this self-saboteur’s way of expressing their belief that a wrong has been done and repair is necessary, even if their actions do not actually serve as morally significant punishment for that wrong.<sup>59</sup>

Both secondary gains of securing attention and sympathy from others and taking revenge-under-the-guise-of-punishment on others involve using self-sabotage to guarantee some further desired end from other people. But other motivations for self-sabotage involve a personal stake or satisfaction in suffering.<sup>60</sup> Specifically, a self-saboteur may desire to suffer in the irreplaceable service of securing some additional, personal end. Warren characterizes the desire thusly:

...[O]ne may desire something *as an essential (necessary) part or aspect of a more encompassing intrinsic end*. The part is not desired merely as a means of obtaining the whole, since (by virtue of being ‘essential’ to the whole) no other means could be substituted for the part. Rather...one desires the end as a whole, *including* its part or aspect.<sup>61</sup>

Warren gives the examples of wanting to run a marathon or express oneself in writing. Desires for these things entail the desire for frustration and pain that comes as part of the package. If we *really* want to run a marathon with a full understanding of what this entails, Warren claims, then we want some physical pain. Likewise, some suffering is so closely bound up with other ends, that to desire those ends is *necessarily* to desire that suffering. This is particularly likely when suffering is in some way connected with an agent’s personal identity. To the extent that one’s self-conception relies on suffering, the benefits that arise from viewing oneself as a sufferer—and ensuring continued suffering—are bountiful. This sort of desire to suffer is perhaps best

---

<sup>59</sup> This is not to say that this sort of revenge-as-punishment could never factor into a genuine reconciliation between parties (or indicate and partly address a real wrong), but just that it can often be a case of mishandling punishment. I don’t have room to unpack this issue further, but it is worth wondering whether and to what extent personal revenge in the form of self-sabotage can legitimately factor into punishment for some actual wrong.

<sup>60</sup> Though of course, there may be personal satisfaction in taking revenge on another and gaining their attention; again, relational and personal incentives for self-sabotage are not hermetically separable. Warren characterizes this distinction as that between external and internal rewards, or the benefits from other people compared to psychological benefits to the sufferer. Op. cite note 36.

<sup>61</sup> Op. cite note 36, p. 112.

understood as a sort of hybrid instrumental and intrinsic desire to suffer: the primary aim is so closely tied up with suffering that suffering is both a means *and* an ends for that agent. Since I hold that self-sabotage can be either intrinsically or instrumentally motivated (or both), Warren’s distinction of a desires that is *a necessary part of an intrinsic end* fits well within the scope of self-sabotage I’m concerned with.

We can identify a few types of these desires to suffer as a necessary part of an intrinsic end, which are often private incentives to suffer. First, an agent may thrive off of suffering because it reinforces their view of themselves as a victim, and it is satisfying to victimize themselves.<sup>62</sup> Just as someone may be pleased by a compliment, gratification can be “derived by being limited and crippled physically or psychologically.”<sup>63</sup> This satisfaction is prominent in individuals with a masochistic personality type mixed with a bit of narcissism. Cooper writes of this patient:

We should...not underestimate the exquisite pleasures of self-pity [derived] from the repeated unconscious demonstration to an imagined audience of ‘look how badly I am treated.’<sup>64</sup>

A self-saboteur may experience these ‘secret satisfactions’ as a sort of vindication, at once affirming her suffering and proclaiming its perceived injustice.<sup>65</sup> This is further supported by a belief that suffering makes one special. An agent may “view suffering as a compatible, and possibly exceptional, part of their personality,” a testament to “one’s special plight in life” that

---

<sup>62</sup> This does not entail that all agents that have been victimized feel or act this way.

<sup>63</sup> Op. cite note 18, p. 29. Notice that though both involve an identity as a sufferer, the gain that comes from this incentive is a personal satisfaction, not the public benefits of attention noted above.

<sup>64</sup> Cooper, A. M. (2009). The narcissistic-masochistic character. *Psychiatric Annals*, 39, 904—912. p. 910-911. Cooper argues that the masochistic-narcissistic personality type should be recognized as an independent clinical entity.

<sup>65</sup> Though again, there may be a conflation here between revenge and morally legitimate punishment. Though in this case, the goal seems to be to prove to oneself and others *that* one has been mistreated, rather than to take revenge on whoever did the mistreating.

makes them remarkable.<sup>66, 67</sup> Having suffered, the thought goes, they are importantly unlike others who could not know what they have gone through. This satisfaction in being a victim is somewhat similar to the public satisfaction of taking revenge on someone else in that it affirms that a self-saboteur believes they have been wronged.<sup>68</sup> Instead of using self-sabotage to hurt someone who has hurt them (as in the case of taking revenge on another),<sup>69</sup> here the self-saboteur maintains their status as someone-who-has-been-hurt or victimized by continually making themselves suffer. By self-sabotaging, agents continue to suffer, which supports the idea that they are victims, which feeds the satisfaction of having-been-wronged and being exceptional. The motivational story runs in the other direction: in order to maintain the gratification of being a victim, an agent self-sabotages.

A second personal incentive to suffer is the safety and stability an agent may find in continued suffering. This helps explain the repetition of self-sabotage: agents may continue to sabotage themselves since “the prospect of upsetting the stable and the familiar is too frightening.”<sup>70</sup> This is true even though ‘stability’ means emotional turmoil, and ‘familiarity’ means suffering. Cooper writes, “With the feeling of victimization, the drama has come to what feels to the [agent] like an appropriate and expected ending...and there is a sense of familiarity, relief, and closure.”<sup>71</sup> There is a comfort in familiarity and a corresponding anxiety in uncertainty, and self-sabotage may be employed as a way to manage this fear. Instead of waiting

---

<sup>66</sup> Op. cite note 17, p. 991.

<sup>67</sup> Op. cite note 63, p. 909. Warren identifies the belief that one is superior because they suffer and a sense of justice or entitlement to compassion as some of the internal rewards of masochism. Op. cite note 36.

<sup>68</sup> This is distinct from self-sabotage that is an attempt to atone for a wrong one has committed; see below.

<sup>69</sup> Again, this doesn’t necessarily serve as morally legitimate punishment of another for their wrongdoing.

<sup>70</sup> Op. cite note 18, p. 29. Warren also notes the internal reward of feeling in control of one’s life by controlling their suffering. Op. cite note 36.

<sup>71</sup> Op. cite note 1, p. 54.

for the frightening unknown, some precipitate familiar pain so that it is no longer uncertain and out of their control.<sup>72</sup>

A particularly safe stability that can motivate suffering concerns one's very sense of self. Recall that the allure of illness can include enacting the identity of a sick person.<sup>73</sup> If an agent's identity fundamentally revolves around suffering, then continuing to suffer offers the benefit of continuing to *be oneself*. Whatever it is founded on, a secure sense of self is a comfort for it grounds the agent in an uncertain world and offers some predictability for the future. An agent may be unwilling to refrain from self-sabotage if doing so poses a threat to this security: the risk of conceiving of oneself without suffering is too great. If an agent has formed their identity around suffering, then the promise of happiness is also a threat of "yourself unmoored," a risk of loss that might not be worth the gain.<sup>74</sup> Slater poignantly recounts this fear of receiving treatment for her mental illness: "I was concerned that Prozac, and the health it spawned, could take away...my very identity...I was a different person now, both more and less like me, fulfilling one possibility while swerving from another. There is loss in that swerving."<sup>75</sup> Though Slater's is not a case of self-sabotage, she expresses a personal connection to her suffering in which the abatement of symptoms would be, for her, a 'loss.' To define oneself as a victim, a wretched person, a worthless loser, requires that supporting data of continued suffering. Warren adds another dimension to this story, arguing that the particular feature of a self-conception built on suffering is that of "being at war with oneself."<sup>76</sup> This is a self that thrives on the constant

---

<sup>72</sup> Op. cite note 18, p. 122.

<sup>73</sup> Op. cite note 46.

<sup>74</sup> Op. cite note 50, p. 11.

<sup>75</sup> Op. cite note 50, p. 49.

<sup>76</sup> Op. cite note 36, p. 118.

tension between conflicting desires.<sup>77</sup> Should suffering abate, the agent would seek out conflict to incite the war anew.

A third private incentive to suffer involves a self-saboteur seeking to atone for some particular wrong they (believe they) have committed – though whether this counts as morally legitimate atonement is, again, an open question. Sigmund Freud’s notion of ‘moral masochism’ embodies this idea of self-inflicted pain motivated by an agent’s unconscious guilt, or their “need for punishment.”<sup>78</sup> Unlike other forms of masochism, which are connected to the sexual enjoyment of pain, with moral masochism, “[t]he suffering itself is what matters...the true masochist always turns his cheek whenever he has a chance of receiving a blow,” because this is what, he feels, he deserves.<sup>79</sup> Masochistic behaviors thus provide one way to address this guilt (guilt one may be quite conscious of, contra Freud). The logic holds up: if the self-saboteur believes suffering is required to atone for a wrong and assuage their guilt, and also believes they have committed some wrong, then, they believe, they must suffer.<sup>80</sup> Again, though, it is not obvious that such self-induced suffering actually counts as morally legitimate atonement; the self-saboteur may at least need to apologize or publicly acknowledge their wrongdoing, and it is an open question whether suffering is ever a necessary component of atonement. As with self-sabotaging to hurt others who have committed a wrong against them, the contrite self-saboteur may collapse suffering and atonement. The point is rather that the agent believes that have done wrong and further believes that causing themselves to suffer is the appropriate answer to having committed this wrong, even if it is more clearly aligned with a sort of self-regarding revenge.

---

<sup>77</sup> Op. cite note 34.

<sup>78</sup> Freud, S. (1995). The economic problem of masochism. In M.A.F. Hanly (Ed.), *Essential papers on masochism* (pp. 274—285). New York: New York University Press. (Original work published 1924). p. 280. Freud later explained masochism as arising from the death drive. Op. cite note 36. Freud’s use of ‘punishment’ seems to fit the colloquial sense of word (as mentioned above) rather than a morally significant atonement.

<sup>79</sup> Op. cite note 77, p. 279.

<sup>80</sup> My thanks to an anonymous reviewer for suggesting this clarification.

Wanting to suffer as a form of desert (whether this is morally legitimate or not) need not be localized to a particular wrong; it can apply to the agent globally. A fourth private incentive to suffer is that a self-loathing agent may believe it is what they deserve *as a person*. Psychiatrist Roy Krawitz discusses agents who “describe intense experiences of severe chronic self-hating, self-disgust and self-contempt,” expressed, for instance, by speaking “of themselves offensively as an animal (e.g. pig), as offensive matter (e.g. scum) or as deserving of punishment.”<sup>81</sup> Chronic, pervasive feelings of contempt towards oneself can be grounded in the belief that one is “fundamentally flawed,” the problem residing in the agent herself, not in some particular act.<sup>82</sup> Suffering becomes unobjectionable, even justified if an agent believes that they, as a person, are somehow *wrong*. This is part of the reason why treating a chronically self-loathing patient is so difficult: by the very nature of their self-loathing, they believe they do not *deserve* to be happy.<sup>83</sup> Suffering is thus consistent with the self-loather’s conception of their worth, resonating as it does with his believed wretchedness. A self-loather may seek out suffering as a way to regulate their hatred towards themselves.<sup>84</sup> This private incentive may be particularly likely to overlap with an intrinsic desire to suffer, since there is a desire to suffer *because* it is painful.

Taken together, there is a plurality of incentives for self-sabotage, whether this is intrinsic (involving a primary motivation to suffer for its own sake), instrumental (securing some additional secondary gain, like the relational gains of connection and attention from others or taking revenge on others), or a necessary part of an intrinsic end (private satisfactions of self-sabotaging in order to view oneself as a victim or satiate a feeling of self-loathing). All of these various motivations, and their intersections, count as deliberate self-sabotage. As we’ll see, this

---

<sup>81</sup> Krawitz, R. (2015). Behavioural treatment of chronic, severe self-loathing in people with borderline personality disorder, part 1: interrupting the self-loathing cycle. *Australasian Psychiatry*, 20, 419—424. p. 419.

<sup>82</sup> Op. cite note 80, p. 419-420.

<sup>83</sup> Op. cite note 80.

<sup>84</sup> Op. cite note 80, p. 420.

is because the motivations for a given instance of self-sabotage bear on the ethical relationships between the self-saboteur and the person who cares about them: whether a self-saboteur is self-loathing or simply wants to connect with others, for instance, will influence what is ethically required of a friend or family member. I turn to these issues below.

#### 4. Ethical Implications of Self-Sabotage

Obligations to care for others are not always straightforward. Moral conflicts, inexhaustible duties, and the threat of excessive self-sacrifice are among the factors that complicate care. But caring for a self-saboteur compounds whatever complications may already surround care since these cases present a tension between something the self-saboteur wants—to suffer—and the alleviation from suffering that those in close relationships may be obliged to bring them (call these agents *carers*).<sup>85</sup> Caring for self-saboteurs thus presents atypical ethical difficulties while also calling into question the assumption that suffering always ought to be relieved.

What follows is a discussion of some of the central ethical issues that arise in these cases: the risk of becoming complicit in another's self-sabotage, the acceptable limits of caring for a self-saboteur, and the permissibility of paternalistic interference. I argue that because these relationships can be imperfect but valuable, some asymmetry can be permissibly tolerated even if it involves a degree of manipulative treatment. Yet this leeway is limited, and I gesture toward considerations that make care too demanding in these cases and in which scaling back care is permissible. Additionally, I argue that caring for a self-saboteur involves more than simply making them better off at any cost; it ought to involve sensitivity to their desires and the

---

<sup>85</sup> Though I focus on a dyadic relationship for the ease of exposition, these relationships can of course have more than two members.

possibility that their reasons to suffer are important to them. Thus, rather than complying with or paternalistically interfering with the self-saboteur, a carer ought to help the self-saboteur by maintaining an honest relationship built on mutually shared and recognized reasons, communicating their own concerns, and acknowledging the self-saboteur's agency.

Consider first the epistemic uncertainty regarding a saboteur's motivations and how this relates to an ethical uncertainty in responding to the saboteur.<sup>86</sup> Given that self-saboteurs' motivations to suffer may not always be transparent, their aim in asking others for help may be difficult to discern. A request for help may be a good faith expression of need and consent to another's effort of care. Alternately, that same request can be a maneuver toward self-sabotage if the saboteur asks for help with the expectation that this help will fail. That is, a saboteur may make a disingenuous request for help in order to undermine it as part of their self-sabotage. This epistemic uncertainty—does the saboteur *really* want help, or is this just something else to ruin?—can make approaching care difficult.

To put this issue in context, return to the case of Jerry, who ran up debt in order to gain attention from his father. Jerry repeatedly put himself in situations that required his father's help, assuming that the consequences of defaulting on his debts were significant. From his father's perspective, Jerry's request may be a genuine way to avoid further distress. But, it might also be a way to make things worse, say if Jerry asks for money with the intention of spending it rather than repaying his debt, thereby defaulting and also incurring his father's anger. Given Jerry's past behavior, this may not be an unfair assumption. Thus it may be difficult for the father to ascertain Jerry's motives: does he want his father's help as another tool for self-sabotage? Or does he genuinely want that help in order to protect himself from harm—is this time different? This uncertainty can create pervasive strain on the relationship as a self-saboteur's pattern of

---

<sup>86</sup> I wish to thank an anonymous referee for pushing me to discuss this issue.

sabotage (perhaps interspersed with times of remission) means his motives may *always* be unclear to the carer.<sup>87</sup> A carer who is uncertain whether or not he is being used may feel disrespected and resentful, adding further stress to the relationship.<sup>88</sup>

A parallel ethical uncertainty arises around what sort of help (if any) a carer should provide. Now, since the saboteur has asked for help, and supposing that the carer is actually in a position to offer the requested help (e.g. Jerry's father has the money to spare), it seems *prima facie* permissible to offer that help. When the recipient has a genuine desire to be better off motivating their request for help, their consent and the carer's ability to help seem to straightforwardly speak in favor of performing that action.<sup>89</sup> However, such a desire cannot be taken for granted in these cases. Rather, there are two distinct, and opposing, senses of 'help' at play.<sup>90</sup> First, there is the *ostensible* help that the saboteur asks for, which, if disingenuously requested, would ultimately make the saboteur worse off. Second, there is help that would *effectively* make the saboteur better off rather than playing into their plan of self-sabotage—exactly what the disingenuous saboteur does not want.<sup>91</sup> The conditions that make help permissible are split between these two senses: a carer only has the saboteur's consent to further their suffering, not make them better off. And this pulls the carer in opposing ethical directions.<sup>92</sup>

Since the saboteur's desire to be better off is exactly what is at issue in cases of self-sabotage, and since it may be unclear what sort of 'help' the saboteur is asking for, the

---

<sup>87</sup> I wish to thank [name removed for review] for suggesting this point.

<sup>88</sup> Blumenthal-Barby outlines some potential damage to a relationship when manipulation is vicious. Blumenthal-Barby, J.S. (2014). A framework for assessing the moral status of 'manipulation.' In C. Coons & M. Weber (Eds.), *Manipulation: Theory and Practice*, (pp. 121—134). Cambridge University Press. p. 128-129.

<sup>89</sup> Though, of course, there may be other ethical complications: is the carer actually able to help? Would the cost of this help be too great? Would it contribute to an immoral end? And so on.

<sup>90</sup> I wish to thank [name removed for review] for suggesting this distinction.

<sup>91</sup> I mean to remain neutral between theories of well-being with the distinction of what would make one *actually* better off; this category is just meant to serve as the contrast to the saboteur's intention to continue suffering.

<sup>92</sup> To be clear, this ethical confusion arises only when the saboteur asks for help anticipating that it will fail, not when they genuinely want to be better off. But due to the epistemic uncertainty of helping someone who continually sabotages themselves, it can be difficult even to determine *if* this ethical question has arisen.

considerations in favor of simply doing as the saboteur asks are undermined. Abiding by the saboteur's request would only provide fodder for further sabotage, which stands as a *prima facie* reason *not* to do what the saboteur asks, even though one is able. Notice the implications: in these cases, doing as another requests under the guise of helping them is ethically dubious at best. This is a strange result if positive duties hold any sway over our moral lives, as it means denying (ostensible) help to others, help that is achievable and explicitly welcome. And executing this restraint may be difficult, especially in close relationships. Feelings of guilt and accusations of betrayal may proliferate, adding further strain to the relationship and weakening the carer's resolve to avoid complicity in another's self-sabotage. These considerations speak against playing into another's self-sabotaging hand, thereby refraining from offering ostensible help.

Further, by doing as the self-saboteur asks the carer is submitting herself to the other's manipulation. This provides yet another reason to refrain from offering ostensible help. Consider cases in which self-sabotage is (at least in part) a way for the saboteur to connect with the carer by gaining their sympathetic attention. Return again to Jerry and his father. Jerry's self-sabotage was a way to compensate for his father's neglect by ensuring his attention. Jerry knows how his father will respond to Jerry's distress; indeed, he counts on it. Leveraging one's suffering over another in this way is typically viewed as morally problematic manipulation.<sup>93</sup> This may be the case even if duplicity is not present; a saboteur may seek attention out of a real need, but just find that manipulation is the easiest way to achieve it. Thus, the carer who gives the self-saboteur the desired attention, and which is appropriate given the saboteur's suffering, exposes herself to manipulation that risks straining the relationship.

---

<sup>93</sup> For a description and ultimate rejection of this position, see *op. cit.* 87, p. 121-123.

Manipulation, or what J.S. Blumenthal-Barby calls “non-argumentative influence,” is “influence that operates either by bypassing a person’s awareness or by relying on facts about the subject’s psychology such as knowledge about his emotions, how he perceives things, how he makes judgments and decisions, and what he desires.”<sup>94</sup> Knowledge of how the carer ‘works’ psychologically provides the leverage the self-saboteur can use to gain the desired response. Intuitively, we condemn manipulative behavior, and tend to think we should avoid situations or relationships in which we may be manipulated, and some have argued that manipulation is always morally problematic.<sup>95</sup> Given these considerations, the response to self-saboteurs may seem clear: since we should generally avoid exposing ourselves to manipulation, we should refrain from giving in to a self-saboteur’s manipulative ploy for help.

However, the reality of these cases is not so straightforward precisely because they concern close relationships and genuine needs from others, even in dysfunctional contexts. This suggests that even if manipulation is always morally problematic (a contestable claim on its own), it may sometimes be worth enduring given the stakes involved in these cases. Leveraging one’s suffering may not be an ideal way to connect with those one loves, but it can be an effective way to ensure that connection. And insofar as this manipulation provides the self-saboteur with the attention and care they need and fosters a connection between the members of the relationship, then accepting some manipulative treatment from another may sometimes be permissible.

Notice that there are two questions here: whether it is permissible for a self-saboteur to use manipulative means to achieve their end, and whether the carer permissibly exposes himself

---

<sup>94</sup> Op. cite note 87, p. 123.

<sup>95</sup> See, for instance, Cave, E.M. (2007). What’s wrong with motive manipulation? *Ethical Theory ad Moral Practice*. 10(2): 129-144.

to this manipulation.<sup>96</sup> The considerations of a close relationship apply to the latter question. It may be the case that it is always impermissible to manipulate another (a claim I won't argue for here), yet a separate question concerns whether allowing oneself to succumb to that manipulation may sometimes be acceptable. Exposure to manipulative force is inherent in any close relationship, after all, since knowing a person well means knowing what moves them. Sarah Conly explains this point in regards to someone who may use emotional pressure on a family member: "They are using the strength of family ties to their own ends, but one downside of having family is that they are allowed to do that. It gives them an advantage because we are vulnerable to our families, but that vulnerability is the price you pay for having an emotional relationship with your family, which on the whole we think is a good thing."<sup>97</sup> Knowing how to influence a loved one's behaviors is the inevitable result of being close with them. In instances like these, allowing oneself to be exposed to some ultimately harmless, low-stakes or short-term plying may be unproblematic, especially since doing so meets the saboteur's real needs.

If vulnerability to manipulation is one risk of close relationships, then so is the potential for unequal demands. Care ethicists have discussed the asymmetry in dependence relationships, in which a cared-for person is dependent on the carer to meet their basic needs.<sup>98</sup> Because of the human reality that we are dependent on others at some points in our lives (at least), everyone experiences being dependent and being depended on. But a less extreme asymmetry also exists in close relationships. Some people may simply require more from us than we do of them. This asymmetry is not just a difference of kind—each of us has unique needs and so requires different

---

<sup>96</sup> I wish to thank [name removed for review] for suggesting this distinction.

<sup>97</sup> Conly, S. (2004). Seduction, rape, and coercion. *Ethics*, 115: 96-121. P. 115.

<sup>98</sup> See Kittay, E. F. (2011). The ethics of care, dependence, and disability. *Ratio Juris*, 24, 49—58; Kittay (2002). When caring is just and justice is caring: justice and mental retardation. In E. F. Kittay & E. K. Feder (Eds.), *The Subject of Care: Feminist Perspectives on Dependency* (pp. 257—277). Lanham, MD: Rowman and Littlefield Publishers.

things from others—but also a difference of degree. Some friends or family may need more comfort, more reassurance, more sensitive treatment than we do, and if these relationships are valuable, then accommodating them in these ways, even if it is unreciprocated, need not in itself be a problem. This is the asymmetry involved in accommodating another’s shortcomings and limitations for the sake of the relationship—‘meeting someone halfway,’ so to speak. Some instances of relationships with self-saboteurs may be like this: a self-saboteur requires certain attention from the carer, but since this is one of their particular needs, it may be acceptable to allow the saboteur to exercise some manipulative behavior to achieve it (again, supposing the stakes are relatively low and harmless). If perfectly symmetrical treatment were a requirement of relationships, we wouldn’t have very many.

However, this is not to say atypical needs warrant exposure to any and all treatment. To extend Conly’s point, even if it is permissible for a family member to induce vulnerability in another, this doesn’t mean they will be happy about having that vulnerability exploited. Even permissible behaviors can strain a relationship, and while this doesn’t make these behaviors suddenly impermissible, it can create a consideration of how much a carer is willing to put up with. Interacting with a self-saboteur can become a balancing act between recognizing another’s needs and the frustration of constantly having to meet those needs under manipulative influence. So, even though the carer may recognize and accept the asymmetrical dynamic, the relationship can *nonetheless* become fraught, especially if the pattern is repeated over months or years.<sup>99</sup>

This is a relevant ethical consideration since the history of a relationship will impact the moral expectations involved. If help is routinely given (money lent, job interviews set up, appointments made, etc.) and yet the self-sabotage and manipulation continue such that the

---

<sup>99</sup> Compassion fatigue can compound this effect, as the caretaker may feel additional guilt over their resentment and frustration at being a caretaker.

relationship comes to depend on this dynamic, then a new consideration of excessive effort comes into play, as does the concern of being complicit in another's self-sabotage. A buildup of resentment from feeling used may turn an otherwise mild instance of manipulation into the final offense that breaks the relationship. Even accounting for the acceptable sacrifices we make for the sake of our relationships, it may be unreasonable for a carer to continually cater to a self-saboteur's manipulations. The strain on the relationship can become so great that the carer may become exhausted by their ceaseless care, frustrated at being constantly used, and hopeless that any change is possible. They also risk neglecting their own self-care for the sake of caring for the self-saboteur.<sup>100</sup> Continued care can thus cause further harm to the relationship and the carer, even though the saboteur may still need that help.

Though there will be no definitive point at which the duty to care 'switches off,' there are nonetheless considerations that speak in favor of scaling back or modifying care. These considerations limit the reasonable amount of leeway in meeting asymmetrical needs in close relationships. The dynamic and history between the self-saboteur and carer are relevant moral considerations in determining this limit. Scaling back care may be permissible even if the relationship is valuable and another is suffering when asymmetrical needs tip over into unilateral impositions. When a relationship dissolves into nothing more than a manipulative dance of seeking and giving attention, this may be an indication that the scales have been tipped and that normally appropriate accommodations for a loved one should be reevaluated.

Another consideration concerns the carer's well-being and his own limitations: his tolerance, resilience, and resources for care, the fact that he may be too exhausted and frustrated to continue catering to the self-saboteur, and the dangers of neglecting his own self-care or

---

<sup>100</sup> For concerns about excessive care, see Pettersen, T. (2011). The ethics of care: normative structures and empirical implications. *Health Care Analysis* 19, 51-64.

becoming excessively self-sacrificial.<sup>101</sup> This is not a question of when the duty to help a loved one is formally discharged so much as a question of how much one can be expected to do for others while still maintaining care and respect for themselves, as well as a question of how much they can tolerate. The potential toxicity of the relationship and the inexhaustible expectations on the carer speak to the limitations of this care.

Even if a carer does feel they have done everything that can be reasonably expected of them, tempering their care may feel very wrong. Because of their close relationship, a carer may feel they are acting cruelly or neglectfully by scaling back their care. Guilt may remain even if the carer recognizes that they cannot continue helping as they have been.<sup>102</sup> A peculiar feature of these cases is that the carer who recognizes the limitations of their care may be accused of being the ‘bad guy.’ As such, the carer must also contend with the knowledge that by scaling back care, he may become complicit in the saboteur’s pain in a different way than he may have been before: a self-saboteur may interpret the carer as abandoning them, which can play into a narrative of misfortune and deserved suffering. An unfortunate consequence of caring for a self-saboteur is that doing what is ethically best can often feel painful and induce guilt. This is part of what makes care so difficult in these cases.

Given the real and often asymmetrical needs of self-saboteurs and the value of relationships with them, exposing oneself to some of their manipulation for the sake of these needs can sometimes be permissible—provided the stakes and potential for damage are low. There is a certain amount of leeway in the behaviors that are morally acceptable in close relationships. Yet, this leeway is not without limitation: factors that tax the carer and add further

---

<sup>101</sup> On the moral dangers of self-sacrifice, see Hampton, Jean. (1993). Selflessness and the loss of self. *Social Philosophy and Policy*, 10, 135—165.

<sup>102</sup> I suspect guilt of this sort is part of the explanation for why so many carers will continue to cater to self-saboteurs, but this is only speculation.

strain to the relationship, like resentment, exhaustion, and the carer's ability to meet their own needs, can outweigh even reasonable demands of asymmetry. This suggests that succumbing to another's manipulation is not a sustainable long-term solution, but rather only deployable in due measure. Further, exposing oneself to this manipulation by doing as the self-saboteur asks is not ultimately a good way to help them, as it effectively makes the carer a party to the self-sabotage and furthers that behavior.

All this suggests that close relationships ought to involve offering *effective* help when it is needed, rather than simply doing what a self-saboteur asks when this involves manipulative requests. The question then arises as to whether it is ethically permissible to go *against* a self-saboteur's consent because one ought to effectively help them, regardless of how painful it may feel to do so, and even though the saboteur does not want this type of help. The question can be framed in terms of paternalism: is it permissible to go against the wishes of a self-saboteur in order to effectively help them?

Paternalism is generally defined as interference with an agent's autonomy without their consent and for the sake of the agent's own good.<sup>103</sup> Paradigm cases include laws mandating seatbelt use and criminalizing drug use, and while these exemplify the legal paternalism that gets much philosophical attention, interpersonal paternalism gives rise to similar ethical issues. To act paternalistically towards someone is to treat them as a child, and when this treatment is inappropriate (not towards actual children, e.g.), it is usually condemned.<sup>104</sup> It is not difficult to find fault with paternalism if autonomy is considered important. We make choices that make us worse off all the time, yet our freedom to make these choices is often more valuable than the benefit we would gain from others' interference in them. But in some cases, preventing another

---

<sup>103</sup> Dworkin, G. (1972). Paternalism. *The Monist*, 56: 64—84.

<sup>104</sup> Feinberg, J. (1986). *Harm to Self*. Oxford: Oxford University Press. P. 2-4.

from harming themselves is so important that the agent's autonomy is permissibly overridden. As Gerald Dworkin puts it: "The usual justification for paternalism refers to the interests of the person being interfered with. These interests are defined in terms of the things that make a person's life go better; in particular their physical and psychological condition."<sup>105</sup> When paternalism is justified, it is through an appeal to the agent's interests that is significant enough to override their autonomy.

An agent's consent to some harm does not factor into justified paternalism, as Joel Feinberg emphasizes. Preventing another from harming themselves means preventing "simple damage to a person's interest, whether consented to or not."<sup>106</sup> In other words, whether or not an agent consented to the damage he is liable to incur does not matter, for even if he did consent to it, the principle that justifies interfering with the agent's autonomy is just that doing so would protect his interests, "for his own good."<sup>107</sup> Where paternalistic action is justified, it is justified whether the agent agrees to it or not precisely because the agent's desires do not hold enough sway to be respected. So, if paternalistic intervention in self-sabotage were justified, it would be justified even though the agent has brought their own sabotage upon themselves and regardless of whether or not they consent to another's interference.

Applied to the cases of concern here, the question is whether the reasons in favor of interfering with an agent's self-sabotage outweigh respect for their autonomy in choosing to harm themselves. Put another way: is it permissible to effectively help the saboteur, though the saboteur has not consented to this type of help and, in fact, does not want it? An initial point to make is that it is not obvious what effective help amounts to. This is not simply because every

---

<sup>105</sup> Dworkin, G. Paternalism. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2017/entries/paternalism/>>.

<sup>106</sup> Op. cite note 103, p. 10.

<sup>107</sup> Op. cite note 103. This is in the case of indirect paternalism, in which the liberty of persons other than the agent are restricted in order to avoid harm to the agent. Dworkin calls this "impure" paternalism (Op. cite note 102, p. 68).

case will be different, but because it can be hard to know what would be best for someone. Would Jerry *actually* be better off if his father cut him off completely, or coerced him into rehab? How to effectively help another is a matter of practical concern made no less complicated by a desire to suffer.

Whatever the particular acts of help amount to, whether or not to implement that help against the wishes of the saboteur remains an unresolved ethical issue. Suppose some act that impinges a saboteur's autonomy would effectively help them (perhaps Jerry would lose his desire to suffer, or at least resist acting on it, were his father to force him into gambling rehabilitation).<sup>108</sup> Given the precedent in legal paternalism, an initial presumption may be against interference: committing your adult son to rehab is a constraint on his autonomy, even though he would exercise that autonomy to harm himself, and so is impermissible. However, one might think that self-sabotage is *just* the sort of case in which paternalistic interference is permissible precisely because the saboteur intends harm towards himself.

An argument may be made against refraining from interference by appeal to the obligations generated by the relationship. A history of standing by while another self-sabotages, even if in the interests of respecting their wishes, may eventually morph into complicity. Failing to interfere when someone hurts himself is a way of allowing that harm, and this may outweigh the fact that this harm was consented to precisely because of the history of a relationship. When self-sabotage is repeatedly allowed, complicity becomes an issue. Respect for another's choices is an important part of a close relationship, yet we can still harm another by failing to interfere

---

<sup>108</sup> The moral maturity of the self-saboteur and the particular type of relationship she is in with the carer will further impact this issue. For instance, if a parent is caring for his minor child, the limits of permissible interference will likely widen. In this case, paternalistic interference is arguably not interference in autonomous action, but rather an expression of parent's duty is to promote their child's best interest until the child can do so on their own. Unfortunately I don't have room for a fuller treatment of this issue, and have therefore restricted my discussion to caring relationships between adults.

when they harm themselves. The duties that arise in a close relationship may therefore speak against refraining from interference since this may constitute complicity in the self-saboteur's suffering. In this way, paternalistic action may be indirectly justified.

One may object at this point that a positive argument can be made for paternalistic interference, and in fact, that cases of caring for a self-saboteur provide paradigm conditions for justified interventions. Consider a parallel case: if the agent and recipient of harm were two separate people—suppose M asks S to harm them, under no deceptions or errors—would it be wrong to allow that harm? Surely some cases exist in which one ought to respect M's choice. Interference with a choice to be harmed made in a BDSM community would not be justified, for instance. Yet in other cases, the fact that M has consented to their own harm may be taken to demand interference on the grounds that harmful desires undermine autonomy by their nature. Recall that the saboteur is not a 'poor schmuck' who makes himself suffer as a result of bad reasoning. Rather, he seeks out his own harm, and perhaps this difference is crucial in justifying interference. A proponent of this view may reason that it is the desire to suffer *itself* that warrants interference since it signals that an agent cannot or should not be allowed to act for himself. A saboteur's inability or unwillingness to want what is best for himself revokes his claim to autonomy in matters concerning his well-being: one cannot autonomously desire to harm oneself, the argument goes.<sup>109</sup> So, the saboteur's desire to suffer can be effectively ignored as an ethical consideration. Simply having a reason to suffer does not bear any ethical weight given that it fundamentally rests on a mistake, so need not be seriously considered or respected when judging how to help someone. At most, it may inform a practical approach to care. Thus, the

---

<sup>109</sup> This position illustrates a substantive account of autonomy according to which to be autonomous an agent's desires must have a certain normative content—here, they must be in the agent's actual best interest.

proponent may conclude, interfering with effective help is justified as it is in the best interests of the saboteur.

But recall Stocker's and Velleman's assessments of perverse desires. Self-saboteurs do not mistakenly desire their own good, but rather know full well that what they choose is against their interests—that is the point. A desire to suffer can therefore be a coherent part of a saboteur's motivations and actions. If, for instance, a saboteur believes she has done wrong, it is at least understandable that she could autonomously desire to suffer in order to atone for this wrong (whether this atonement is thereby achieved is a further question). In this case, desiring to suffer is at least plausibly an autonomous desire, and the self-sabotage it motivates does not rest on a mistake. If we can properly desire something bad for ourselves, then a desire to suffer cannot be considered non-autonomous merely on the grounds that it is a misguided object of desire.

If a desire to suffer can sometimes be autonomous, the paternalist will have to explain why self-sabotage motivated by this desire is justifiably interfered with. The paternalist may claim, 'regardless of what an individual wants, their suffering should be eliminated because it is inherently bad for the agent.' Since self-sabotage is always against a saboteur's interests, concern for these very interests may outweigh respect for their decision.<sup>110</sup> Yet this approach summarily dismisses the importance to an agent of certain reasons to suffer. If we are trying to figure out the appropriate ethical response to someone who has a real stake in suffering, then these attachments and motivations matter, especially in the context of a close relationship. And while certain extreme conditions may justify interference, like serious or irreversible (potential) harm, a

---

<sup>110</sup> If one is an objectivist about well-being, for instance, this argument will be easy to make.

categorical justification for paternalistic intervention in self-sabotage seems far too quick, as does an absolute deference to someone's desire to suffer.<sup>111</sup>

So, paternalistically interfering with a self-saboteur's autonomy, thereby offering effective help against their stated wishes, and complicit aiding in their effort of self-sabotage by acquiescing to their request for ostensible help, thereby exposing oneself to manipulation, are both problematic options for the carer. However, this bad choice only exists when the carer tries to determine what she ought to do for the self-saboteur in absence of considerations of the health of the relationship. That is, rather than simply capitulating to the self-saboteur or defying their consent, a carer can instead endeavor to help the self-saboteur while also preserving (or establishing) an honest, genuine relationship. She can do this by having a forthright conversation with the self-saboteur about the very concerns that threaten to harm the relationship. This suggests that in order to fulfill one's duties as a friend or family member, a carer ought to engage honestly with the self-saboteur about his actions, the carer's response, and their combined effect on the relationship in an effort to reach a common ground that both the carer and the self-saboteur can accept.<sup>112</sup>

As is evident in our relationships, caring for someone in the context of a friendship, family, or companionship involves more than simply making sure another's needs are met and acting in their best interest. Rather, these caring relationships entail a certain reciprocity: roughly, regarding another in a certain way, expecting a certain regard in turn, and recognizing the other's needs and desires alongside one's own (among other things). One way to understand this feature of close relationships is through the sharing and recognition of reasons that both

---

<sup>111</sup> Intervening in another's suicide attempt may come to mind here, but I take this to be a separate issue with its own set of considerations.

<sup>112</sup> I am indebted to an anonymous referee for suggesting I incorporate this discussion into the paper.

parties can accept.<sup>113</sup> Being a friend (for example) partly involves the sharing of mutually acceptable reasons: being able to openly communicate with another about the projects, values, and decisions that matter to a person, and expecting that a friend offers their own honest take on what is important to them. Voicing and having one's own reasons taken seriously, and reciprocating that recognition, is essential to forming an authentic friendship.

A relationship in which one party's perspective is continually discounted or ignored, or in which one party acts according to reasons the other party cannot accept, lacks this reciprocal quality of a healthy friendship. Indeed, circumstances in which one is unable to openly share their perspective with another or does not accept another's reasons suggest a detachment that can prevent legitimate relationships from forming or undermine healthy connections.<sup>114</sup> Authentic relationships involve some degree of acting according to reasons that both parties accept, especially with regard to how one acts in the relationship: if S behaves towards P in ways that P considers unjust or otherwise unacceptable, the relationship will likely become strained. All this suggests that the sharing and mutual recognition of reasons is especially crucial to preserve when

---

<sup>113</sup> One way to understand the sharing and recognizing of reasons as essential to an authentic friendship is through a contractualist perspective (though I make no commitment to this; this aspect of friendship could be grounded in other ways). The thought is that part of what it means to be a friend is to be able to share one's perspective on what is important with another, and to do so in a genuine and open way. In order to achieve this genuine and open environment for sharing perspectives, a friend will have to recognize the other's reasons, and will expect their reasons to be recognized. This basis in mutually sharing and recognizing reasons that we expect others to accept is laid out in T.M. Scanlon's take on mutual recognition (though note that he is talking about a relation more fundamental than friendship here): "The contractualist idea of acting in accord with principles that others (similarly motivated) could not reasonably reject is meant to characterize the relation with others the value and appeal of which underlies our reasons to do what morality requires. This relation, much less personal than friendship, might be called a relation of mutual recognition. Standing in this relation to others is appealing in itself—worth seeking for its own sake." (Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, p. 162). So, being in a friendship with another involves this mutual recognition (at least), and this means it matters to the friendship that one can share their reasons with a friend and expect them to recognize and accept those reasons *as reasons*. I thank an anonymous referee for pushing me to explore this point.

<sup>114</sup> For Scanlon, living according to principles that others could not reasonably accept constitutes an "estrangement" from others, a way "of having violated the requirements of a valuable relation with others" (Op. cite note 112, p. 162). Specifically, these 'requirements' speak to the "to the value people set on the belief that their lives and institutions are justifiable to others" (p. 163), a belief that is challenged (for instance) by the realization that one's principles are unjust or immoral. Scanlon cites the civil rights movement and the backlash against the Vietnam war as examples of a society confronting the fact that the principles they live by are *not* justifiable to others, a realization that results in an estrangement from others (p. 163).

manipulation from one party threatens the relationship, and that caring for a self-saboteur will involve an effort to move towards mutually acceptable reasons with regards to the self-saboteur's behavior.

Recall the epistemic uncertainty that arises when a self-saboteur asks a carer for help: a carer may be unsure whether the self-saboteur really wants help or only wants to further his self-destruction. This uncertainty about the self-saboteur's motivations coupled with the consideration of maintaining an authentic relationship suggest the carer ought to have an honest conversation with the self-saboteur. Talking to the self-saboteur about the strain his requests puts on the relationship and the burden it puts on the carer is a way of voicing one's own needs while taking the other seriously as an agent. Before merely submitting to being manipulated again, a carer ought to express to the saboteur the strain his requests puts on their relationship, and her unwillingness to aid in his project of self-sabotage: "I know that you are using me, and it is unfair of you to ask this of me." Further, the carer should recognize and ask about the self-saboteur's motives: "Look, you've asked me to do this before and you didn't follow through. Do you *really* want my help this time?" If the carer suspects the self-saboteur is not genuine in his request, then she should communicate the strain that this puts on the relationship: "This puts me in a really tough place because I don't want to keep helping you hurt yourself"; "This is the last time I'm doing this because you're being manipulative and it has to stop if we are going to continue being friends," etc. While it may be difficult to have this conversation, it is morally necessary for the sake of maintaining a relationship built on the honest sharing of mutually acceptable reasons.

Specifically, the carer's responsibility to communicate with the self-saboteur about their strained relationship serves a few related ends. It expresses the carer's needs (e.g. not being

manipulated into furthering another's self-sabotage) which is itself crucial to maintaining an authentic relationship; it appeals to the self-saboteur's reasons instead of silently undermining or complying with them; it maintains the honest communication that is the *basis for* an authentic relationship; and, crucially, it moves the carer and the self-saboteur closer to a place of mutually agreed-upon behaviors with regard to the self-sabotage. To this last point, the thought is that both parties need to be on board with the features of the relationship. For example, a self-saboteur may need the carer to stop patronizing her when she self-sabotages, while the carer may need the manipulation to stop. If both parties can accept these reasons, they have made a step towards preserving their relationship. A single conversation may not be enough to achieve this end. Rather, the carer should endeavor to maintain the conditions for honest conversation for the sake of the relationship and in order to form a relationship that both parties can accept, which will likely be an ongoing effort.<sup>115</sup>

The self-saboteur's particular motivations for their behavior will influence the direction of these conversations. Helping someone whose self-sabotage is primarily an instrumental means to some other goal may involve finding alternative, less destructive methods to achieve this goal. In some cases, honestly talking about the self-saboteur's motivations may even go some distance to obviate the behavior. If the self-saboteur is primarily motivated by the instrumental aim of connecting with the carer, for example, an open conversation about this can dissipate the need for self-sabotage. The carer and self-saboteur can instead agree that the self-saboteur will tell the carer explicitly when they need some company, for instance. Or, if the self-saboteur is hurting herself in order to take revenge on the carer, clearly a conversation about the perceived wrong is in order. An honest determination of whether and to what extent a wrong was committed can

---

<sup>115</sup> Though I don't have space to discuss it here, a more complete treatment of the ethical issues surrounding how to care for a self-saboteur will necessarily involve the reciprocal duties of the self-saboteur with respect to the caring relationship.

make way for an apology and forgiveness (if that is appropriate), or at least the acknowledgement of the self-saboteur's hurt. These efforts, again, are in the service of maintaining a relationship that is acceptable to both parties.

The more difficult conversations may be those in which a self-saboteur has some personal, intrinsic stake in suffering. Here the carer's obligations and the saboteur's desires are most strikingly at odds since the choice to suffer can sometimes be singularly important to the saboteur. Consider the agent who self-sabotages in order to maintain a coherent self-conception. Here, suffering is required for the agent to hold onto her very orientation to the world, to *be who she is*. Yes, she would be better off in one sense if she were prevented from harming herself, but the cost to her may be enormous, amounting to a profound dis-orientation to herself and the world. Though Warren does not go so far to say this, her argument of the lengths one would go to in order to continue 'being at war with oneself' suggests the gravity of this self-conception.<sup>116</sup> As such, the issue cannot be simply a matter of determining what is good or bad for the saboteur, but rather concerns what the saboteur deems to be vital *to them* independent of its hedonistic (dis)value. Endeavoring to understand the self-saboteur's reasons to suffer should therefore factor into a carer's response. This consideration extends beyond objective interests, harms, and benefits to the saboteur to include what the saboteur takes to be significant to them, even if this sometimes involves suffering.

Where self-sabotage is primarily desired for its own sake or as part of a personal end that is essentially bound up with suffering, a carer's responsibility is to address these ends with the self-saboteur and determine, together, how worthwhile and healthy they are. Again, establishing open, honest communication is crucial here, though made no less difficult by the often taboo and uncomfortable topics of victimhood or self-hatred. But also crucial is the carer's willingness to

---

<sup>116</sup> Op. cite note 36.

consider that these ends are important to the self-saboteur. That is, rather than simply subverting these ends in order to make the self-saboteur objectively better off, the carer ought to acknowledge that a self-saboteur's reasons for suffering may be legitimate to her. This acknowledgement doesn't necessarily mean the carer should capitulate to these reasons and let a loved one do whatever they want—after all, the self-saboteur may have some *bad* reasons, and the carer should tell them so. But crucially, this approach is consonant with maintaining an authentic, honest relationship in that it displays a willingness to listen to another's reasons and *reason with them* to get to a mutually shared understanding. Meeting the self-saboteur where they are, so to speak, is a way of treating her as an agent with legitimate desires and needs rather than as someone with disordered thinking who needs to be corrected. And it is this approach that endeavors to maintain a healthy relationship while actually helping the self-saboteur.

It is especially important to try to understand and communicate with self-saboteurs who are intrinsically motivated to suffer *because* their behaviors are likely to be written off as unintelligible, irrational, or pathological. Self-saboteurs may understandably shut down at patronizing attempts to 'do what's best for them' or admonitions to simply 'cut it out.' When self-saboteurs are regarded as incapable or unentitled to make decisions about their own lives, their agency is eroded. By contrast, giving self-saboteurs the opportunity to express their reasons instead respects their agency while also voicing the carer's concerns. This, again, is just what is called for when helping a self-saboteur while maintaining the integrity of a relationship that both parties can accept. A carer should take the other's project of self-sabotage seriously and endeavor to understand their reasons for acting so that she can openly communicate her concerns about the behavior and reach common ground together.<sup>117</sup>

---

<sup>117</sup> I am indebted to an anonymous referee for the term 'project of self-sabotage,' which was too good not to borrow.

The crux of the issue here is that however irrational self-sabotaging behavior may *seem* to the carer, the self-saboteur is not their patient. She is a friend, a family member, a companion, who is making choices that hurt her. Caring for a self-saboteur must therefore involve treating her as an agent who is making these choices and motivated by reasons, not as someone who must be ‘healed’ or ‘fixed’ at any cost. Recall that the cases of deliberate self-sabotage of concern here run along a spectrum of pathology, and that both non-pathological self-sabotage and cases that are symptomatic of mental illness are equally considered. This is because both types of case give rise to the ethical questions under consideration here. The fact that one is a friend to someone whose self-sabotage is symptomatic of a mental illness does not mean their duties become those of a doctor. Rather, their duties are just those of someone in a close relationship with a self-saboteur, and as such, they are partly shaped by that relationship.

Hanna Pickard makes this point at the end of her discussion about holding service users with “disorders of agency,” like addiction or eating disorders, responsible.<sup>118</sup> She writes:

...although the clinical aim is to care and to help, that is not the only aim of friends, family, or others in society at large. Their aims will likely be different and various. For instance, at least one typical and significant aim of friends and family is to have real and genuine relationships...Hence the possibility of real and genuine relationships, and of equal standing between service users and others, may be lost if the latter are too careful to act in the former’s interest at the expense of how they naturally feel. Outside of clinical contexts, equality, respect, and belongingness may be best expressed through ordinary as opposed to special treatment.<sup>119</sup>

Pickard’s emphasis on equal standing between parties is crucial: whatever the mental health of a service user, they ought to be treated in their personal relationships with the same respect and recognition as anyone else.<sup>120</sup> The same goes for relationships with self-saboteurs, whether their

---

<sup>118</sup> Pickard, Hanna. (2013). Responsibility without blame: philosophical reflections on clinical practice in KWM Fulford et al., eds. Oxford Handbook of Philosophy and Psychiatry: 1134-1152. p. 1134.

<sup>119</sup> Op. cite note 117, p. 1149.

<sup>120</sup> Pickard’s larger argument concerns the importance of holding service users responsible and the mechanics of this sort of responsibility from clinicians. In general, she writes, “...clinicians, family, friends, and others need to hold

behavior is pathological or not. Treating a self-saboteur as anything other than a friend who is making some questionable decisions that are self-destructive and straining the relationship risks further damage to the relationship itself, to say nothing of the agency-denying harm it poses to the self-saboteur.<sup>121</sup>

Still, trying to understand self-saboteur's reasons does not mean a carer necessarily accepts them. A carer can express their concerns and disapproval while also recognizing that a self-saboteur's reasons for acting are important to them. Indeed, this recognition may be essential for figuring out how to help a self-saboteur. If a self-saboteur is motivated by self-hatred, for instance, a carer who realizes this may say, "I understand that you think you are unworthy of love, and that must be so painful. But I don't believe that's true, so I won't do anything to help you hurt yourself. It's my job to try to convince you that you're wrong." In the case of self-sabotage sought for personal atonement, a carer should address the perceived wrong and identifying ways to effectively atone once and for all, or determine whether atonement is appropriate at all. If instead the self-saboteur is motivated by the belief that suffering is part of her self-conception, the carer can provide a needed perspective while still recognizing the self-saboteur's subjective authority: 'I respect that this is an important part of who you are, but I'm here to remind you it's not the only part...' These conversations must involve honestly voicing one's own reasons and listening, respectfully, to another's in order to arrive at a resolution that both parties accept.

---

service users with disorders of agency responsible for their behavior, and ask that they change it, when it causes harm to self or others" (Op cite note 117, p. 1150). I am sympathetic to Pickard's conclusion that service users should be held responsible by clinicians, friends, and family members alike, and think the same holds of self-saboteurs (though these groups are not coextensive). I don't think it will always be as simple as asking self-saboteurs to change their behavior, though, and that this can sometimes be the wrong approach.

<sup>121</sup> Of course, if a friend or family member has a diagnosed mental illness, this will impact some features of the carer's responsibilities. For instance, if they suspect a relapse or if the self-saboteur poses serious harm to themselves, they should contact their clinician or other emergency personnel. But the point is just that this duty arises in their capacity as a carer, not as a doctor; one doesn't stop treating another like a friend just because they are self-destructive or have a mental illness.

Though at first glance, paternalistic interference may seem permissible in caring for a self-saboteur, considerations of maintaining an authentic relationship and recognizing the self-saboteur's reasons reveal the problems with this approach. However, avoiding paternalism should not amount to complicity: though interference by doing what would effectively make a saboteur better off is morally problematic, deferring to the saboteur's requests for ostensible help is similarly wrong. In other words, even if a carer should not interfere with the self-saboteur, this does not mean it is permissible to be complicit in their suffering. Instead, given the strain that self-sabotage can put on a relationship, a carer should endeavor to help the self-saboteur while maintaining an authentic relationship built on honesty and the mutual recognition of reasons. This is especially important since some self-sabotage can cause damage to relationships *because* it involves dishonesty or manipulation. Crucial to this project is taking another's project of self-sabotage seriously while at the same time honestly communicating one's own needs and concerns and working towards a mutually acceptable resolution.

Caring for a self-saboteur is nothing if not complicated. And the nature of these complications means these cases do not have simple or straightforward verdicts. But teasing out some of the factors at play can help us make headway in answering questions about these cases and, perhaps, about caring relationships in general. To that end, analyzing additional factors that further complicate care is the subject for future work. Such factors involve, for instance, teasing out how the particular type of relationship (e.g. between friends, romantic companions, or between parent and child) may influence care and ethical concerns. Attending to the complexities of such cases can help us better understand them.

## 5. Conclusion

We are not simple creatures. We make mistakes, we act irrationally, and we do exactly what is against our own best interests. And our desires are complicated, involving much more than a want for comfort and happiness. We can have reasons to suffer, investments in harm, and motivations for self-sabotage—and relationships with those who are so motivated. Paying attention to the ethical questions these behaviors give rise to can hopefully encourage understanding and empathy for self-saboteurs and the people that care for them.