

The Abnegated Self

Nellie Wieland

Penultimate version for *Virtue, Narrative, and Self: Explorations of Character in the Philosophy of Mind and Action* (eds. Joe Ulatowski and Liezl van Zyl)

Abstract: A self-abnegating person lacks contact with their agency. This can be against their will, in absence of their will, or voluntarily. This does not mean that they cannot provide reasons for or a narrative about their actions. It's just that the reasons or narrative are someone else's. People abnegate parts of their agency regularly; for example, within hierarchical institutions. In other cases, the self-abnegation is all-encompassing; for example, a victim of brainwashing. An agent in such a position can completely fail to understand themselves or be understood by others as having a self. I focus on two problems related to self-abnegation. The first is whether there is a conception of a self that can reliably discriminate between strong selves and abnegated selves. The second is whether a person with an abnegated self should be treated as a person with a strong self. I conclude that the good of respecting derived or instrumental agency comes in putting that person in a position to agentially flourish and in maintaining the structural conditions and expectations of agency.

1st Gent. *Our deeds are fetters that we forge ourselves.*

2nd Gent. *Ay, truly: but I think it is the world
That brings the iron.*¹

Sometimes a person introspects and there is nothing there. Or, whatever is there is not a single coherent self to which can be attributed a narrative, reasons, or authentic life plan. In the most extreme case, the disunified mind is a kind of Humean theater in which, “perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations...”² Oliver Sacks describes a patient with Korsakoff's condition

¹ George Eliot, *Middlemarch* (London: Penguin Classics, 2003) Book IV, Chapter IV.

² David Hume, *A Treatise of Human Nature* (Oxford: Oxford University Press, 2000) 1.4.6. Philosophers, literary authors, filmmakers, and psychologists have documented many ways in which a self can be lost. A self can be lost in utter isolation (cf. the main characters in David Markson's *Wittgenstein's Mistress* and in Lydia Davis' *The End of the Story*, and accounts of the psychiatric effects of solitary confinement, as examples). It can also be brought upon by memory loss and memory fracture, as noted since Augustine. The filmmaker, Luis Buñuel, writes poignantly about this: “You have to begin to lose your memory, if only in bits and pieces, to realise that memory is what makes our lives. Life without memory is no life at all...Our memory is our coherence, our reason, our feeling, even our action. Without it, we are nothing...” (Luis Bunuel, *My Last Sigh* (New York: Alfred A. Knopf, 2013), 4–5). In addition to isolation, memory loss, and alcoholism, similar effects are documented with other degenerative conditions.

(severe memory loss and confabulation due to alcohol abuse). This patient is characterized as lacking a past or future, and stuck in a “constantly changing, meaningless moment.” Sacks said that the patient had lost part of himself, but was in no position to know it. Unlike the loss of any other part of oneself (such as a leg), one cannot know that one has lost one’s self.³

There are more ordinary cases of disunified selves that are less severe and may allow for eventual resolution. In an ordinary case, a person may be alienated from themselves, and lack even internal coherence. In this paper, I consider the problem of self-abnegation, both descriptively and normatively. This abnegation can be to another person, to external pressures, or even to a concept or an idea. The cases that I consider involve more than just deference, or institutionally circumscribed abnegation of agency. They involve something that looks like the transfer of agency to something or someone outside of oneself. Any narrative of one’s life or reasons given for action, in such a case, are properly those of another. For example, if I believe whatever B believes, if I want whatever B wants, and if I treat the reasons for my beliefs and desires as mine just in case they are B’s, then I have abnegated my self to B.

In what follows I look at two questions. The first is this: is there an account of selfhood that can reliably discriminate between a strong self and an abnegated self? The second is this: what are the reasons, if any, for treating a person with an abnegated self as if they were a person with a strong self? To answer these questions I begin with four features of a strong self. My goal is to identify the feature of a strong self that, when missing, is evidence of abnegation. I argue that abnegation differs from other ways in which a person’s self can be compromised. For example, if a self is lost to the psychiatric effects of isolation or to neurological degeneration from alcoholism, it will fail to evince the features of strong selfhood identified here. The case of self-abnegation is different because it is a kind of borrowed or externally imposed selfhood. In this case, an abnegated self may appear to satisfy each of the four features of strong selfhood on their own. In order to resolve this, I provide more detailed descriptions of abnegation in order to identify how it diverges. I speculate that these four features are all necessary features of a strong self and must be sourced autonomously from a first-person perspective. Even if an abnegated self putatively satisfies the four features of strong selfhood, those four features must also be derived from a first-person perspective. I then turn to the normative question of why the abnegated self deserves to be treated as an intact agent. I defend what I call a ‘conversational’ and ‘ecological’ account of the conditions of agential flourishing. In doing this I explicitly reject what I call a ‘counterfactual’ account—which I take to be the primary alternative.

³ Oliver Sacks, “The Lost Mariner,” *The Man Who Mistook His Wife for a Hat* (New York: Summit Books, 2006). For other interesting descriptions of the loss of self in patients with degenerative conditions, see Anil Ananthaswamy, *The Man Who Wasn’t There: Investigations Into the Strange New Science of the Self* (New York: Dutton, 2015).

1. A Strong Self

We can begin to think about what it might mean to abnegate your self, or your agency, by first providing an account of what it looks like when it is intact. A strong self has at least four features:

1. It is *integrated* with one's reasons and desires.
2. It has *continuity* over time.
3. It is reflectively *endorsed* at a higher level.
4. It is *interpersonally* engaged with or recognized by others.

Each of these features has been argued for extensively. Each on their own cannot reliably discriminate between a self that is autonomously and authentically one's own and one that has been abnegated. Let's see why this is the case.

A strong self must be integrated with one's reasons and desires.⁴ These three features of a person — selfhood, reasons, and desires — are interdependent. A person's desires are authentically one's own if they are coherently integrated with one's reasons for action and if they come from one's self. This can be described as having a 'distinct perspective'⁵ — that there is some way in which one acts in the world that comes from an integrated and self-determining stance. That stance is distinct insofar as a self is integrated with its own reasons and desires for acting. On its own, this condition for a strong self is fairly weak. An agent's desires and reasons for acting could be fully integrated with one another and yet completely derived, as by a rule, from another agent. The self of that agent could be derived from the will of another and yet display full integration.

A strong self has continuity over time. This kind of continuity is typically represented as narrative continuity.⁶ The standard view of narrative identity is one that allows for multiple narrators throughout the course of a life held together; Owen Flanagan calls this a 'multiplex'

⁴ This is a traditional account of selfhood, defended in some version by just about everybody since John Locke; paradigmatically by Michael Bratman, and challenged rarely, cf. David Hume and Elijah Milgram.

⁵ See, Geoffrey Sayre-McCord and Michael Smith, "Desires...and Beliefs...of One's Own, in *Rational and Social Agency: the Philosophy of Michael Bratman*, eds. Manuel Vargas and Gideon Yaffe, (Oxford: Oxford University Press, 2014), and a fuller defense of this standard in the text below.

⁶ A classic defense of this view can be found in Owen Flanagan, *Self Expressions: Mind, Morals, and the Meaning of Life* (Oxford: Oxford University Press, 1996). For another variation, see Marya Schechtman, *The Constitution of Selves* (Ithaca: Cornell University Press, 2007).

self and contrasts it with a ‘multiple’ self — which he describes as disordered.⁷ In this latter case, there is no self that is holding the narrative together between the sub-narrators. He calls this holding together ‘narrative connectedness’. This narrative connectedness comes about by active authorial work. This is itself interesting, and probably central (and not accidental) to any correct account of the self. It’s not just that a narrator can tell a coherent story, continuous across time, and from a first-person perspective, but that this is something that is constructed by that first-person narrator. Telling a coherent story, and having a first-person perspective, is not a passive act. It requires autonomous action to tell (and be) the story of your life.

Narrative continuity may demand that an agent be able to tell a connected story about who they are and how their life is going. This in turn should be held together from a unique first-person perspective on past events and a continuous thread to a present sense of self. A demand for narrative continuity can take different forms. It can ask for a whole life narrative: a single story that holds all of the parts of one’s life and identity together. Or, it can allow for segmented narratives: carving up either different parts of one’s life, or different aspects of one’s agency.⁸ Abnegation, on the other hand, is not always a short-term state. Long-term abnegation can still have all the marks of narrative continuity. A derivative self can be narratively continuous if it is derived from an agent who evinces narrative continuity. This will lack a unique, first-person perspective, properly understood, but is otherwise possible.

A strong self is reflectively endorsed at a higher level. This is a simple statement of Harry Frankfurt’s⁹ account of second-order desires and volitions. A person is the kind of creature who possesses second-order desires and volitions, and the self of that person is the kind of entity that is endorsed at the second-order. A person’s self is constituted by their beliefs and desires, which are endorsed as their own and arranged in some such way at the second-order. The volitions of this self just are those that emerge from this constitutional ordering.¹⁰ Some version of this concept of persons, and their selves, is undoubtedly both correct and essential to the possession

⁷ Flanagan, *Self Expressions*.

⁸ While there are critics of narrative views of the self (see, Peter Lamarque, “On Not Expecting Too Much from Narrative,” *Mind & Language* 19, 2004: 393–408), there are also persuasive defenses of narrative views that do not demand a whole-life narrative arc (see, David Lumsden, “Whole Life Narratives and the Self,” *Philosophy, Psychiatry, & Psychology* 20, 2013: 1–10).

⁹ Harry Frankfurt, “Freedom of the Will and the Concept of a Person,” *Journal of Philosophy* 68, 1971: 5–20.

¹⁰ Another variation on this view comes from Donald Bruckner, “In Defense of Adaptive Preferences,” *Philosophical Studies* 142, 2009: 307–324. In his discussion of reflective endorsement in the case of adaptive preferences, Bruckner distinguishes between autonomously acquiring and autonomously retaining a preference. They are one’s own once they are retained upon reflective endorsement. I think that this is a good distinction to make. The view defended in this paper may be slightly stronger but is in a similar vein.

of a strong self. However, on its own it still fails to distinguish the strong self from the abnegated self. The reason for this is that a person can possess first- and second-order beliefs and desires and not have come by them on their own. If I first-order desire whatever B desires, and if I second-order treat the reasons for my desires as mine just in case they are B's, then I have abnegated my self to B. Second-order desires can still organize first-order desires and operate according to a rule such that they are whatever is derived from an external source. There is no particular way to ensure that they are authentically, autonomously, and internally generated.

A strong self is interpersonally engaged with or endorsed by others. Without taking a position on whether this is a necessary condition, we can at least see that without such interpersonal recognition — in cases of isolation — selves disintegrate. There are reasonable descriptions of selfhood and autonomy that emerge from or are realized within relations.¹¹ There are also theories of agency and selfhood that assume something about the significance of interpersonal engagement or recognition. Versions of this are embedded into compatibilist and conversational accounts of moral responsibility, and into interpersonal views of selfhood.¹² I will say more about this in the final section of the paper.

A strong self is integrated with one's reasons and desires, has continuity over time, is reflectively endorsed at a higher-order, and is interpersonally engaged with or recognized by others. These sound like pretty high demands but I think that they are the steady state for selfhood. Before continuing to develop an account of self-abnegation in light of these four features of strong selves, let's look at the conditions in which selves are abnegated.

2. Three varieties of self-abnegation

The self can be abnegated in a number of ways, some diffuse, some direct. I briefly describe three varieties of self-abnegation. The first example, the 'diffuse' abnegated self, is a case where an agent does not develop, or fully realize, a self of their own due to social circumstance. This may happen over the course of a lifetime and may result in an under-developed conception of oneself as a source of reasons, narrative, and agency. The second kind

¹¹ For a sample of descriptions and defenses of related views, see Andrea Westlund, "Deciding Together," *Philosophers' Imprint* 9, 2010; Catriona Mackenzie & Natalie Stoljar, eds., *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (New York: Oxford University Press, 2000); David Velleman, "How to Share an Intention," *Philosophy & Phenomenological Research* 57, 1997: 29–50.

¹² See P.F. Strawson, "Freedom and Resentment," *Proceedings of the British Academy* 48, 1962: 1–25; Michael McKenna, *Conversation and Responsibility* (Oxford: Oxford University Press, 2012); Manuel Vargas, *Building Better Beings* (Oxford: Oxford University Press, 2013); Andrea Westlund, "Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?" *Philosophical Review* 112, 2003: 483–523.

of case I describe is of a self abnegated as a result of domination by another. The third kind of case I describe is of a self willfully abnegated to another person or idea.

I have primarily thought about self-abnegation as a bad thing. Perhaps it is a personal bias that I regard the giving up of one's self as something that is worse for a person than anything else. But this isn't obviously true. Subjective well-being is awfully important, and an abnegated agent can still score highly on measures of subjective well-being, even though their goal-fulfillment is inauthentic. There may be ways in which adopting and holding one's own goals is antithetical to subjective well-being because that can be more difficult, isolating, and disappointing. Failing as a group, or failing at achieving goals inauthentically set, may be less painful than failing as an individual. So while I focus in this paper on the harms that accompany that abnegation of self, a discussion of the joys will have to be had another time.

In the previous section I introduced four conditions of strong self. Self-abnegation works with these four conditions. Abnegation occurs because of some failure in the development of these conditions. For example, an agent who is systematically and uniformly unrecognized as an agent (a failure of the fourth condition), or an agent who does not arrive at second-order organization of their first-order desires autonomously, is an agent who is partially or wholly abnegated. However, a self can be abnegated while one or more of the conditions is putatively being met. I return to this point below.

2.1 Abnegation from social circumstance

A person's development of self can, pretty obviously, be shaped, limited, and structured by social forces. I think of this as *diffuse* because it cannot be tied to a limiting event or person (as in 2.2), but it is the result of chronic and structural suppression of the features of agency and selfhood. These features of agency and selfhood might include having a subjective point of view, permission to form life plans of one's own, and recognition that one's reasons and rational deliberations are taken as legitimate sources of these plans. They might also include being seen as someone who is capable of interpersonal engagement with their reasons and rational deliberations, and having the resources to organize one's life around these processes. Suppression of these features of agency and selfhood take familiar forms: legal, societal, and cultural enforcements of suppressed personhood. They may be publicly instantiated in familiar ways: a lack of freedom of movement or livelihood, ineligibility to own property or represent oneself in court, denial of rights to form relationships with other people or to choose who one marries or with whom one is intimate, etc. In cases of chronic and structural oppression, these features of agency and selfhood can be suppressed, the abnegation of the self is assumed by the structure of society until it is assumed by the agent as well. The kinds of conditions in

consideration here are, for example, slavery, and societies structured around sexist, racist, ethnic, or religious control.

It is important to draw distinctions between the self-abnegation from social pressure that I have in mind and two other phenomena. Adaptive preferences are changes in one's attitudes towards realizing one's desires given environmental limits.¹³ These can be conscious or subconscious. The kind of self-abnegating person I have in mind is not merely a person adapting their preferences, since adaptive preferences are compatible with self-affirming agents who are simply working within a world that doesn't always give them what they want. The kind of structurally diffuse, oppressive abnegation I am describing is something that strips the authenticity and self-ownership of an agent's reasons, rational deliberations, and interpersonal engagement. So while adjacent to adaptive preferences, I am trying to describe a much stronger attack on the self. It's also important to distinguish between the self-abnegation I have in mind and shared autonomy.¹⁴ The latter can be a positive phenomenon that does not identify an agent as having diminished or inauthentic agency, but rather as having agency or autonomy that is not located in a single self, isolated from their relationships with others.

In cases of abnegation from social circumstances, a person is systematically immersed in the idea that their own agency and selfhood lacks value, authenticity, and autonomy, and ought to be substituted with the agency and selfhood of another person or group of people. This is the case when, for example, women or racialized persons are regarded as having deficient agency and selfhood and are not allowed to make decisions or have standing with respect to marriage, reproduction, finances, property ownership, legal representation, religious belief, enfranchisement, etc. The claim being made here is not that self-abnegation is determinate in these cases (obviously, it is not, which explains resistance against these very forms of oppression) but that it is not only possible, but likely, that self-abnegation will be internalized in the oppressed group.

2.2 Abnegation under domination

The second kind of case I have in mind is of an agent who never develops a self of their own due to domination by a particular other. These are cases of self-abnegation to the point of internalizing the domination or oppression and becoming pathologically deferential. Illustrative examples include brainwash, abuse, and other chronic and isolating oppressive circumstances. It

¹³ See Bruckner, "Adaptive Preferences"; Dean Zimmerman, "Sour Grapes, Self-Abnegation and Character Building: Non-Responsibility for Self-Induced Preferences," *The Monist* 86, 2003: 220–241; Jon Elster, *Sour Grapes: Studies in the Subversion of Rationality* (New York: Cambridge University Press, 1983).

¹⁴ See Mackenzie and Stoljar, *Relational Autonomy*.

may be the case that some cases of this kind of concentrated (in contrast with ‘diffuse’ in **2.1**) oppression do not result in self-abnegation. The kinds of cases I am interested in are those that do — those cases where a person objectifies themselves, and comes to believe that they lack subjectivity, rationality, and individuality as a result of persistent mistreatment. A person in this position cannot identify values, aspirations, or a self-reflective identity of their own and this can be seen as directly tied to delimited circumstances of autonomy-denial. For example, an isolated community or family that fosters an ideology of domination by one group over another, extreme deference, and provides no models or access to alternative conceptions of moral relationships is in a position to shape every aspect of a person’s self-conception. People may exist in such communities with no conception of themselves as oppressed, since they don’t conceive of themselves as having unrealized selves of their own. This can also be found in some abusive relationships, long-term kidnap victims, and, perhaps especially, in children raised in similar conditions.

2.3 Willful abnegation

The third kind of case is of the agent who willfully abnegates to another person or conception. Dean Zimmerman¹⁵ describes an agent in this circumstance who, in order to escape an overbearing work environment, renounces his self entirely (he might renounce the very idea of self) in order to be unflappable in the face of threats to his desires. He might simply reject the idea of his having a self and the beliefs and desires that come along with it. He might do this as a way of escaping negative experiences. But, as a result, he doesn’t have an authentic self to whom can be attributed reasons or interpersonal engagement. Or, consider a case of religious devotion that is so robust that an adherent regards themselves as ‘vehicles’ of the will of God, or as ‘married’ to Jesus, or as having ‘renounced the world’ and themselves to their devotion. This kind of renunciation moves along several planes. The adherent may give up on their own self and its reasons, beliefs, and desires, but they may also give up on (moral) responsibility for themselves and for others. These cases differ from those in **2.2** because they don’t require express control or domination by another. They may not require the presence of another person— certainly renouncement of self to God in such a case. This doesn’t mean that there is always a clear line between willful abnegation and domination by another. Some of the time we willfully seek out those leaders, rulers, dominators, prophets who we enthusiastically follow until our selves are erased by control or abuse. However, these cases can come apart too, and it’s worth attending to the differences in these varieties of self-abnegation.

¹⁵ Dean Zimmerman, “Sour Grapes, Self-Abnegation and Character Building: Non-Responsibility and Responsibility for Self-Induced Preferences,” *The Monist* 86 (2003): 220–241.

3. Identifying an abnegated self

How can an abnegated self be identified, and what do the above accounts of self-abnegation have in common? In each of the cases, the four criteria for a strong self may putatively be met, or be met in various degrees, but what is lacking is an integration of the four criteria from a first-person, autonomous position. This is why I argue that none of these criteria for selfhood is sufficient on its own. What does it mean for these four criteria to be integrated from a first-person, autonomous position? The remainder of this section is devoted to understanding this.

Here I think it is helpful to think about moral responsibility. Accounts of moral responsibility have arisen out of two related problems: the (lack of) freedom of the will, and the problem of degrees of (diminished) agency. Very simply understood, if it's the case that the freedom of the will is an illusion, then it seems that it must follow that there is no place for moral responsibility. One can only be responsible for that which one can control. Instead, we can look to how we hold one another responsible. We can look for how we take one another to be responsible, depending on our capacities, and that *taking* (or reacting) is what structures the norms of moral responsibility.¹⁶

The taking of one another as responsible is a matter of assessment of the 'quality of will' of the agent in question.¹⁷ This can be minimally understood as the giving and taking of reasons. If someone is taken to be morally responsible, then presumably they are in a position to offer reasons for their behavior, and to be responsive to questions. This is itself an interpersonal exchange, one where agents are taken to be responsible to the extent that they can provide reasons for their behavior in exchange with other agents. This is a corollary to the criteria I use above of (1) integration with one's reasons and desires, and (4) interpersonal engagement or recognition by others. Theories of moral responsibility can demand a richer understanding of criterion (4). They can also require the presumption of competence in the person taken to be responsible. For example, I take *X* reasons to be the kinds of reasons a competent agent takes to be governing. Insofar as I offer *X* reasons to you to explain your behavior or to justify mine, I am claiming that each of us is such an agent. This can also be applied to agents who have a diminished set of competencies. It can make sense of agents with limitations in their competence (or perhaps willingness) to accept as governing those reasons that a strong agent accepts. These

¹⁶ Cf. Strawson, "Freedom and Resentment"; and J.L. Austin, "A Plea for Excuses," *Proceedings of the Aristotelian Society* 57 (1957): 1–30.

¹⁷ Cf. McKenna, *Conversation and Responsibility*, for a 'conversational theory' of agency.

agents lack some quality of their will and could be considered morally responsible to some degree, even if not fully morally responsible.¹⁸

This enriched understanding, borrowed from accounts of moral responsibility, of the integration of one's will with one's reasons and desires, interpersonal engagement and recognition, narrative continuity, and second-order endorsement, still don't seem to capture all that has gone wrong in the case of self-abnegation. What is missing is the source of the self that performs the integrative function. By definition, self-abnegation means that an agent's self is borrowed from or derivative from another. The strong self must be sourced first-personally. Here I find it helpful to think about what it means to have beliefs of one's own and to extrapolate from that to thinking about what it means to have a self of one's own. The idea of borrowing the beliefs of another person is easier to understand. If I believe whatever my husband, boss, religious leader, or political party believes, then those beliefs are not mine. Or at least they aren't mine in the same sense as if I arrived at them independently. Of course it's the case that all agents rely on authorities or external sources for their beliefs, so the question to ask is whether these borrowed beliefs are otherwise integrated with one's reasons and desires, and endorsed at a higher-level. But this too is still inadequate. What's missing is the process whereby these beliefs are adopted autonomously.¹⁹ The autonomous holding of beliefs, as understood here, means that the belief can be identified with the agent, and the belief can be identified by the agent as their own.²⁰ This bi-directional identification is what constitutes the agent's perspective on the world. The self is just the integrated set of autonomously derived and integrated beliefs and desires which are second-order endorsed (and result in second-order volitions), have narrative continuity, and are responsive to interpersonal engagement with others.

At this point, it's worth stepping back and looking at how far we've come. It seems at first that any one of the four criteria are adequate for a strong self (in fact, many people have argued just this). The problem of self-abnegation has led me to argue that all four criteria are — in the strongest version of the argument — necessary, but only jointly sufficient if they are sourced autonomously from a first-person perspective. This is perhaps too strong since a strong self is not a single end state: there can be degrees of achievement of a strong self, and degrees of waning. Who among us has such a strong self? Perhaps most of us do, but there are certainly grounds for doubt.

¹⁸ McKenna, *Conversation and Responsibility*, 84. Although this seems compatible with McKenna's view, it's not entirely clear that he endorses it. He mentions it in a footnote and claims that this articulation of the challenge comes from Michael Zimmerman. McKenna, *Conversation and Responsibility*, 82, fn. 3.

¹⁹ This is the conclusion of Geoffrey Sayre-McCord and Michael Smith, "Desires...and Beliefs...of One's Own," which I have found very helpful in thinking about the problem of self-abnegation.

²⁰ McCord & Smith, 135–136.

The account of selves and agents defended here, and all mainstream theories of agency, assume that there is someone in charge. They assume that each of us has a ‘subjective motivational set’, or a ‘practical identity’. It is on this basis that we make decisions, fulfill our desires, and override this or that errant agency. What should we think, then, about the person who is not an ‘abnegated agent’ but is experiencing a momentary lapse? We say that people like this are ‘not themselves’ today or ‘out of character’. This assumes that there is some self underlying the errant agent being projected today, one who is, for some reason, merely resting? Hiding? Then who is the self that is here today? You are well enough constituted to direct your will, to organize your intentions and efforts, but that is not *who you are*. But perhaps it is the case that there is no *you* underneath the projection; there are just different projections (‘projections from nowhere’) with varying endurance and recognizability.²¹ We *are* whoever we experience most of the time and whomever others take us to be. This might just be right; stable agency and strong selfhood might be illusions. If this is the case, then there may be no difference between the self that has been abnegated and the one who has not. But, rather than undermining the concern of this paper, it amplifies it.²² What I will consider in the next section is the nature of our obligations toward abnegated selves. If it turns out to be true that there’s no meaningful contrast with the abnegated self, then this worry just applies to everybody.

4. Respecting an abnegated self

In this section I turn to questions about moral engagement with self-abnegating agents: should a person be treated as if their reasons and narratives are their own even though they are derived from the will of another? What kind of obligation exists to cultivate self-affirmation in the self-abnegating agent?

Imagine a person who is profoundly deferential to her spouse now, and who was profoundly deferential to her parents before she was married. Imagine that she has never formulated a life plan of her own in any recognizable way. Her life plans were formed for her, and were not, in any substantive way even identified as hers. Instead, she has been told to take, and has always taken, the goal of her agency to be to serve the life plans of someone else. We

²¹ This argument is influenced by Elijah Milgram and his description of ‘segmented agents’. I take him to be arguing that all agents are segmented agents, or that there is nothing but “interfaces you conjure up to meet the needs of the moment,” Elijah Milgram, *The Great Endarkenment* (Oxford: Oxford University Press, 2015), 263.

²² I take it that the challenge of segmented agency to the view I’m developing here parallels the challenge of determinism for moral responsibility. If the response to the challenge of determinism is to cast the norms of moral responsibility in the structure of how to act *as if* we are responsible — the reactive attitudes — then the response to the challenge of segmented agency is to cast the norms of mutual regard in the structure of the conditions of agency as argued in the final section of this paper.

could think of her situation in two ways: we could say (a) that she has a life plan, but it is a plan that has been determined in generalities and specifics by someone else, or (b) that she does not even have a life plan since she is assumed to be a wholly instrumental agent, there to serve the life plan of someone else. Typically, the reason why it is wrong to undermine the life plan (or desires, autonomy, etc.) of an agent, is because that life plan is derived from her process of reason-giving and her values. When these reasons and values, and the goals and plans they produce, are derivative, or can only be understood instrumentally, does the obligation to respect those reasons and values change? I think that the answer to this is tricky, in part because it assumes quite a bit of confidence in ordinary agents to authentically determine and construct their selves. It also assumes a neat separation between different ways of harming the same person. For example, if the person in the example above were to desire to undergo a medical procedure, but that desire was derivative on the will of her spouse, we might want to distinguish between harming her by blocking her authentic-desire and by blocking her derivative-desire. But, in both cases, this might involve us doing something that brings about pain or bodily control.

Again, the reason why it is harmful to undermine or obstruct a person's life plan is because it comes from that person in some meaningful way; it is *one's own* in a way that matters. If this is not the case, then the basis of that harm changes in all of those cases that matter deeply to us — those which evinces values, beliefs, plans, and aspirations.

4.1 Counter-factual arguments

What do we owe the self-abnegating agent? We might think that this question can be answered by simply considering what we would owe them had they never abnegated their self. Consider the counter-factual status of this agent. The claim to consider is that respect for the self-abnegating agent should be grounded in the respect that would be owed that agent if they had lived life without the oppressive circumstances that brought about their self-abnegation. That is, ostensible agents are owed the regard of well-functioning agents because of the counter-factual possibility that they could have been well-functioning agents in the absence of mistreatment.²³ This counter-factual argument is offered as a way to contrast the agentic qualities we have 'naturally' and those we have that are socially imposed. We can acknowledge that we will have some difficulty separating natural and socially imposed agentic qualities. Still, when we apply

²³ I identify this kind of counterfactual argument with Rae Langton, *Sexual Solipsism: Philosophical Essays on Pornography and Objectification* (Oxford: Oxford University Press, 2009); and with Sally Haslanger, "On Being Objective and Being Objectified," in *A Mind of One's Own*, edited by Louise Antony and Charlotte Witt (Boulder, CO: Westview Press, 2002). I don't think either of them ever make this explicit argument, so my attribution should be taken with a grain of salt. Nevertheless, the counter-factual argument discussed here is still something that I think is at least implicit in their views.

the counter-factual argument in each of the above cases, the outcomes are uneven. In the first case, the causal structure is too diffuse to locate the counter-factual reality. In the third case, the counter-factual argument is question-begging since the renunciation of the self, and its aspirations, was deliberate. Only in the second will the counter-factual argument be tractable. Even in this case it requires assumptions about ideal agency that are tricky. For example, return to the example of people living in the isolated community enforcing an ideology of male-domination. What is the counter-factual person like? How do we imagine her life? What is the model of agency that she realizes in this counter-factual case? Perhaps she is self-aware and prone to interpersonal engagement; perhaps she is well-educated and checking all the boxes on indices of flourishing; perhaps she is respected and admired. But why? The model for this counter-factual person is herself a unicorn among us. This makes her no less a source of norms and expectations, but it means her magical qualities might be a matter of wishful thinking.

4.2 Self-affirmation for the self-abnegated

What would it look like to create the conditions of affirmation, and to forestall the conditions of abnegation? This is either to prevent self-abnegation or to recover agency in the abnegated self. It might involve conceiving of a place for treating abnegated selves as strong selves. And it might be virtuous to treat abnegated selves as strong selves (however that ends up looking). If so, are these virtuous acts for the performing agent, for the beneficiary, or are they part of structure of virtue that creates the conditions of agency? This final option is the most coherent. The obligation to affirm the life plans, respect the reasons, and interpersonally engage with the self-abnegated agent is itself weak when understood as a simple moral relationship. Instead, something like a structure of virtue surrounds agency among all of us — realized, potential, and counter-factual. The good of respecting even a derived or instrumental agency comes in putting that individual in a position to flourish, but more importantly it maintains the structure and expectation of agential flourishing that is a necessary condition of overcoming the oppressive conditions that lead to self-abnegation to begin with.

The model for our responsibility toward self-abnegated agents cannot be predicated on a one-to-one relationship with the individual with, for example, derivative desires. Assuming that they may, counter-factually, have been a person who could have chosen for themselves, has to be grounded in another round of desire-derivation. If we regard the person who derivatively-desires a medical procedure as someone who does not authentically desire that medical procedure, then we are imposing a desire on them yet again. We can claim that it is counter-factual but we don't know what that counter-factual looks like: we might be assuming that it looks something like us. Instead, the best model for responsibility toward self-abnegated agents should be thought of

structurally, whereby each of us acts in such a way that create the conditions of agential flourishing, interpersonal engagement, and authenticity. We do not know how things *would have been* in any given case, and we cannot always find the joints between interfering with derivative agency and other kinds of harm. Perhaps the best we can do is to expect the selves around us to be self-generated as a means of creating the structural conditions of overcoming the oppression that lead to self-abnegation to begin with.

This puts us in an uncomfortable spot. On one hand, acting according to counter-factual assumptions of who the self-abnegating agent *would have been* had they retained or developed an authentic self of their own will not work. But the alternative cannot be to reinforce the abnegation. Return to the woman undergoing the medical procedure. We are imagining that she wants the medical procedure, and that she wants the medical procedure because her husband wants her to have the medical procedure. What would she want if her desires were not derivative? What would she want counter-factually? We cannot know this. There is not sufficient reason for us to think that *she* would not want the medical procedure. If, in our counterfactual imaginations, we ask what similarly situated women would want, we are again, abnegating her agency to those similarly situated women. The alternative is to respect her current desires — go forward with the medical procedure! We do this knowing that this desire is not authentically her own, even if it is one she is stating now. Modifying Michael McKenna’s golf example, all golfers play with handicaps. It’s just that, in order to have a level playing field, each golfer is allowed more or fewer strokes relative to the players they are playing with. A weak self is still a self, they are just acting given limitations, with compromised abilities to form beliefs, desires, and reasons of their own. Each golfer has different abilities, and has their strokes adjusted to match their abilities, but they are all still allowed to play. The weak self is not expected to have a unique, first-person stance from which they autonomously and authentically construct the narrative of their life, but they are not overlooked entirely. The woman’s request for the medical procedure is a minor act of agency, even if it is founded on a major substructure of abnegation. To respect it is unsatisfying, but it is still what we ought to do.

4.3 Moral ecology²⁴

I close with a slightly different consideration from that which is discussed in 4.2. If a person is accurately described as abnegated to the point of quasi-agency, then what obligations exist to cultivate or recover their agency? It is clear that we have obligations to respect agents, but do we have corresponding reasons to create the conditions of agency (or a structure of

²⁴ This is a term used by Manuel Vargas, *Building*. I may be using it in a different way, but it is meant to have the same spirit.

virtue), and the conditions of affirmation for that agency? If this is the case, it cannot be derived from respect for the abnegated self alone; instead it is derived from an understanding of the relationship between the abnegated self and others. For example, it could be grounded in whatever reasons we have for respecting one another's existing agency. If A has an obligation to create the conditions of B's self-flourishing under a scheme of flourishing decided by B, then A may, for the same reason, have an obligation to bring about the conditions where C could even formulate a scheme of flourishing to begin with. Seeing your life's choices as aligned with a sense of self requires first that you have some sense of self, and so it is not unreasonable to suppose that we ought to help you realize that as a starting point.

The ecological system within which the abnegated self lives is one where each of us in the position of acting as if we are all strong selves, and authentically integrated across the features of our agency. In doing so we project outward to others that they are the same. Conversely, we treat others — through our laws, our practices, our stances toward one another — as strong selves, and this is a way of affirming our own status within this moral community. While I do not think that we are 'segmented' all the way down, I recognize that we all abnegate some of the time and in some spheres and that occasional deference is a normal part of social functioning. We are all in compromised states and these are usually opaque from the inside. The best we can do, for ourselves and others, is to treat ourselves and others as if our selves are intact through and through as our best hope of making it so.

5.0 Conclusions

In this paper I have focused on two related problems. The first is whether there is a descriptive account of the self that can distinguish between what I have called a strong self and an abnegated self. I described four features of selfhood: (i) integration with one's reasons and desires; (2) continuity over time; (3) higher-order reflective endorsement; and (iv) interpersonal engagement and recognition from others. Calling all four of these necessary conditions is too strong since a strong self is not all-or-nothing. Rather it is something that comes in degrees throughout one's life. These four conditions provide a standard for determining the strength of a self and the extent to which it is lacking or distorted. I have argued that these four features themselves require autonomous integration from a first-person stance, or from a 'distinct perspective'.

I have compared the treatment of abnegated agency to certain ways of thinking about moral responsibility. If someone is regarded as having a strong self then we presume that they are in a position to demonstrate the integration of their self, demonstrate their higher-order endorsement, and demonstrate its narrative continuity with their previous self. This is similar to the giving and taking of reasons for action with those whom we take to be morally responsible.

Finally, I turned from descriptive problems to normative problems. I considered what is owed to the abnegated self and whether those persons should be treated *as if* they have strong selves. I answered that, yes, they should be treated as if they have strong selves. I rejected that the reason for this is because of counter-factual arguments about what their selves could have been like had they not been oppressed. The counter-factual arguments introduce epistemic quandaries about how we could we know what an agent counter-factually would have wanted in other circumstances, and how that forces us into abnegating that self further by projecting idealized beliefs and desires onto the abnegated self. Instead I argue that we should treat the person with the abnegated self as if they had a strong self because that is how the conditions of agency are cultivated. We both make the world a place where each person can become a strong self, and we project our own selves into that world, through this kind of mutual respect. I acknowledge that counter-intuitiveness of this conclusion since it means respecting the will of a person knowing that that will is not her own.

References

- Ananthaswamy, Anil. *The Man Who Wasn't There: Investigations Into the Strange New Science of the Self*. New York: Dutton, 2015.
- Austin, J.L. "A Plea for Excuses." *Proceedings of the Aristotelian Society* 57 (1957): 1–30.
- Bruckner, Donald. "In Defense of Adaptive Preferences." *Philosophical Studies* 142 (2009): 307–324.
- Buñuel, Luis. *My Last Sigh*. New York: Alfred A. Knopf, 2013.
- Eliot, George. *Middlemarch*. London: Penguin Classics, 2003.
- Flanagan, Owen. *Self Expressions: Mind, Morals, and the Meaning of Life*. Oxford: Oxford University Press, 1996.
- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68 (1971): 5–20.
- Haslanger, Sally. "On Being Objective and Being Objectified." In *A Mind of One's Own*, edited by Louise Antony and Charlotte Witt, 209–253. Boulder, CO: Westview Press, 2002.
- Hume, David. *A Treatise of Human Nature*. Oxford: Oxford University Press, 2000.
- Jaworska, Agnieszka. "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value." *Philosophy & Public Affairs* 28 (1999): 105–138.
- Lamarque, Peter. "On Not Expecting Too Much from Narrative." *Mind & Language* 19 (2004): 393–408.
- Langton, Rae. *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*. Oxford: Oxford University Press, 2009.

- Lumsden, David. "Whole Life Narratives and the Self." *Philosophy, Psychiatry, & Psychology* 20 (2013): 1–10.
- Lumsden, David and Joseph Ulatowski. "One Self Per Customer? From Disunified Agency to Disunified Self." *The Southern Journal of Philosophy* 55 (2017): 314–335.
- Mackenzie, Catriona and Natalie Stoljar, eds. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. New York: Oxford University Press, 2000.
- McKenna, Michael. *Conversation and Responsibility*. Oxford: Oxford University Press, 2012.
- Milgram, Elijah. *The Great Endarkenment*. Oxford: Oxford University Press, 2015.
- Nussbaum, Martha. "Objectification." *Philosophy and Public Affairs* 24 (1995): 249–291.
- Sacks, Oliver. "The Lost Mariner." In *The Man Who Mistook His Wife for a Hat*. New York: Simon & Schuster, 2006.
- Saul, Jennifer. "On Treating Things As People." *Hypatia* 21 (2006): 45–61.
- Sayre-McCord, Geoffrey and Michael Smith. "Desires...and Beliefs...of One's Own." In *Rational and Social Agency: the Philosophy of Michael Bratman*, edited by Manuel Vargas and Gideon Yaffe, 129–150. Oxford: Oxford University Press, 2014.
- Schechtman, Marya. *The Constitution of Selves*. Ithaca: Cornell University Press, 2007.
- Strawson, P.F. "Freedom and Resentment." *Proceedings of the British Academy* 48 (1962): 1–25.
- Vargas, Manuel. *Building Better Beings*. Oxford: Oxford University Press, 2013.
- Velleman, David. "How to Share an Intention." *Philosophy & Phenomenological Research* 57 (1997): 29–50.
- Westlund, Andrea. "Deciding Together." *Philosophers' Imprint* 9 (2010): 1–17.
- Westlund, Andrea. "Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?" *Philosophical Review* 112 (2003): 483–523.
- Zimmerman, Dean. "Sour Grapes, Self-Abnegation and Character Building: Non-Responsibility and Responsibility for Self-Induced Preferences." *The Monist* 86 (2003): 220–241.
- Zimmerman, Michael. "Another Plea for Excuses." *American Philosophical Quarterly* 41 (2004): 259–266.