

Desire and motivation in desire theories of well-being

Atus Mariqueo-Russell¹

Accepted: 20 April 2023 / Published online: 6 May 2023 © The Author(s) 2023

Abstract

Desire theories of well-being claim that how well our life goes for us is solely determined by the fulfilment and frustration of our desires. Several writers have argued that these theories are incorrect because they fail to capture the harms of self-sacrifice and severe depression. In this paper, I argue that desire theories of well-being can account for the harm of both phenomena by rejecting proportionalism about desire and motivation. This is the view that desires always motivate proportionally to their strength. If we reject this view, then we can explain the harm of many cases of self-sacrifice as arising when we act upon our weaker desires and thereby frustrate our stronger desires. Moreover, the harm of many cases of severe depression can be explained by its suppression of the motivational force of desires. This inevitably frustrates desires that we are left unmotivated to fulfil. I argue that this approach captures the experiential quality of self-sacrifice and severe depression better than rival views that seek to problematise these phenomena for desire theories of well-being. Moreover, these theories have sufficient conceptual resources to account for residual cases of self-sacrifice and severe depression that are less well explained by this approach.

Keywords Well-being · Welfare · Self-sacrifice · Depression · Desire Satisfactionism · Preferentism

1 Introduction

This paper considers two problems for desire theories of well-being. Firstly, Mark Overvold argues that they entail that self-sacrifice is impossible. Secondly, Ian Tully and Mark Spaid argue that these theories entail that severe depression does not dimin-



Atus Mariqueo-Russell atusmariqueorussell@gmail.com

¹ University of Southampton, Southampton, UK

ish well-being. If either argument is correct, then desire theories of well-being are undermined by the counterintuitive nature of their conclusions. I claim that solutions to both problems can be found by rejecting proportionalism about desire and motivation. This is the view that desires always motivate proportionally to their strength. Rejecting this view allows us to account for problem cases of self-sacrifice by claiming that they occur when our altruistic desires motivate us with disproportional strength. Moreover, the harm of many cases of severe depression can be explained as arising from its suppression of the motivational force of desires. This approach solves two problems for desire theories of well-being and provides independently attractive characterisations of self-sacrifice and severe depression.

The paper has the following structure: § 2 outlines desire theories of well-being; § 3 examines the problem of self-sacrifice, surveys proposed solutions, and puts forward a new approach; § 4 outlines the problem of depression and proposes solutions; § 5 highlights independent reasons to reject proportionalism about desire and motivation.

2 Desire theories of well-being

Well-being is the value that determines how intrinsically well a life goes for the subject living it. Desire theories of well-being (henceforth desire theories) are a family of views about what comprises well-being. The most plausible versions of these theories contain variations of the following three features:

- 1. Only the fulfilment of a subject's desires intrinsically increases their well-being, while only the frustration of their desires intrinsically decreases their well-being.
- 2. The extent to which a subject is made better or worse off by a desire fulfilment or frustration is proportional to the strength of their fulfilled or frustrated desire.
- 3. A subject's lifetime well-being is entirely determined by aggregating the well-being effects of instances of desire fulfilment and frustration over the course of their life.¹

While this formulation is relatively uncontroversial, it is also where consensus breaks down over how these theories ought to be constructed.² Nevertheless, the problems of self-sacrifice and depression are relevant to most popular variants.

² For instance, some writers claim that these views should exclude instrumental desires from intrinsically affecting well-being (Sarch, 2013, 223; Heathwood 2005, 489; Brandt 1979, 111). I remain neutral on this question because of the difficulty in distinguishing between instrumental and intrinsic desires (Parfit, 1984, 117), because it is possible that no instrumental desires exist (Murphy, 1999, 252–256), and because some intuitions suggest that instrumental desires can intrinsically affect well-being (Heathwood, 2019, 669).



¹ This feature is the most contestable. Some writers claim that lifetime well-being is non-additive (Velleman, 1991). On such views, the distributional shape of well-being throughout the course of a life is intrinsically relevant to the total well-being value of that life. I set aside this issue here.

3 The problem of self-sacrifice

It is often claimed that for a theory of well-being to be viable it must generate intuitive conclusions in paradigm cases of benefit and harm (Fletcher, 2016, 10). Mark Overvold argues that desire theories fail in this respect (1980).

According to Overvold, for an action to qualify as self-sacrificial it must contain three features. Firstly, it must result in an anticipated loss of well-being for the person who performs it. Secondly, it must be voluntary. Thirdly, there must be a viable alternative action that the agent correctly predicts would be more in their self-interest to perform (Overvold, 1980, 109–114). In what follows I assume that Overvold's position is broadly correct. However, not much hinges on this. Alternative formulations of self-sacrifice raise similar problems for desire theories, and other writers have put forward related problems.³ Responses to Overvold can be adapted and applied to these problems.

Overvold argues that if desire theories are correct, then voluntary actions can never contain the first or third features of self-sacrifice. His view is that we cannot anticipate losses to our well-being from our informed voluntary actions because they always improve our well-being. Moreover, he claims that there can never be alternative actions that we anticipate are more in our self-interest to perform. This is presumably because he takes our voluntary actions to always be produced by our presently strongest balance of desires. Our presently strongest balance of desires is for the action that we predict will maximise the fulfilment over frustration of our presently existing desires when adjusted for the relative strength of the desires in question. Consequently, if we always act upon our strongest balance of desires, then our informed voluntary actions are always optimal for our well-being (Overvold, 1980, 115). If an action is involuntary, then it can fulfil the first and third criteria of self-sacrifice. However, by definition, it falls foul of the second.

Self-sacrifice is a paradigm case of harm that viable theories of well-being must capture. To illustrate this point, Overvold constructs an example of a father who desires to kill himself for the life insurance payout to fund his children's education (1980, 107). Overvold takes this to be an uncontroversial case of self-sacrifice. However, desire theories seem to counterintuitively entail that the father's suicide would not diminish his well-being. If this is correct, then we have good reason to reject these theories. I will refer to this argument as 'the problem of self-sacrifice'.

The problem of self-sacrifice can be summarised as follows:

- P1 Self-sacrifice is a harm that viable theories of well-being must capture.
- **P2** Desire theories fail to capture the harm of self-sacrifice.
- **C** Therefore, desire theories are unviable.

³ Chris Heathwood (2011, 18–19) has compiled a list of related arguments by luminaries including Richard Brandt, James Griffin, Stephen Darwall, Thomas Carson, Thomas Schwartz, L. W. Summer, and Amartya Sen.



Overvold's argument is valid. Moreover, premise one strikes me as unassailable. However, premise two is contestable. If it can be shown to be incorrect, then an influential objection to desire theories will be defused.

Premise two rests upon at least four unexamined assumptions. These are:

- 1. Desire theories claim that the fulfilment or frustration of any desire affects well-being.⁴
- 2. Human psychology is structured in such a way that altruistic actions are necessarily motivated by desires.
- 3. Desire theories claim that maximising present desire fulfilment is always optimal for our well-being.
- 4. Human psychology is structured in such a way that the strength of our motivations is always proportional to the strength of our desires.

Overvold takes assumption one to be part of the definition of desire theories. Whereas assumptions two, three and four are presupposed but not explicitly articulated in his paper (Overvold, 1980). Proposed solutions to the problem of self-sacrifice can be constructed from the rejection of any one of these assumptions. The following subsections consider them sequentially.

3.1 Do all desires count?

Overvold's first assumption concerns the definition of desire theories. He takes these views to claim that the fulfilment or frustration of any desire intrinsically affects well-being. However, this is not true of all versions of the theory. For instance, some writers argue that desire theories ought to restrict which desires intrinsically affect well-being to only our self-regarding desires (Carson, 2000, 75; Parfit 1984, 494; Sidgwick 1907, 109–110). These are the desires that we have for states of affairs that involve our own lives in some way. If this theory is right, and the desires that motivate self-sacrifice are non-self-regarding, then their fulfilment does not intrinsically improve well-being. Consequently, actions motivated by non-self-regarding desires can qualify as self-sacrificial on Overvold's criteria.⁵

However, it is not obvious how to non-arbitrarily distinguish between self- and non-self- regarding desires (Adams, 1999, 88). One approach claims that self-regarding desires have an 'I' in their content.⁶ However, this view is subject to counterex-

⁶ This approach is premised on the idea that desires are, or can be accurately represented as, propositional attitudes (Sinhababu, 2015; McDaniel & Bradley, 2008, 268). Some writers have argued that this is incorrect (Brewer, 2006; Thagard, 2006). If they are right, then we cannot distinguish between desires in this way.



⁴ Following Richard Brandt (1972, 682), Overvold restricts the desires that determine well-being to only those that are fully informed (1980, 107). Nevertheless, his problem emerges regardless of whether we accept this restriction. For simplicity, I omit further mention of it.

⁵ Overvold suggests his own solution for how desire theories can handle the problem of self-sacrifice. His proposal is that only our desires that logically require our own existence as a constituent part count towards our well-being (Overvold, 1980, 117–118 n.10). This proposal is more restrictive than the one discussed herein and is susceptible to a similar critique.

amples. For example, if I desire that 'I help you attain your goal of learning to ride a bike', then this intuitively appears to have both self-regarding (that I help you) and non-self-regarding (that you attain your goal) components. However, because this desire has an 'I' in its content, this approach relegates my desire to simply being self-regarding. Nevertheless, if I fulfil my desire to help you learn to ride a bike at considerable cost to myself, for example, by buying you the bike, then it seems counterintuitive to claim that the fulfilment of this desire contributes to my well-being as much as that of any other equally strong self-regarding desire. Desires to be of service to others are counterexamples to this way of distinguishing between self- and non-self-regarding desires (Adams, 1999, 140).

Furthermore, while it is counterintuitive to conclude that self-sacrifice cannot exist, it also seems wrong to claim that many of our altruistic actions do not improve our well-being. Yet, if this restriction is correct, then those altruistic desires that do not have an 'I' in their content, do not intrinsically improve the well-being of the agent who fulfils them. However, common intuitions suggest that many altruistic actions do improve the well-being of the altruist (Wolf, 1982, 437). This version of the restricted desire theory fails to capture these intuitions. Moreover, it also counterintuitively excludes other types of non-self-regarding desires from intrinsically affecting well-being. For example, it excludes desires about our children and desires about the fortunes of our favoured sports teams from doing so (Heathwood, 2011, 25). It strikes me that any version of the restricted desire theory that relies upon a distinction between self- and non-self-regarding desires is destined to make casualties of our intuitions in these types of cases. Consequently, this approach is not an attractive response to the problem of self-sacrifice.

More recently, Chris Heathwood has put forward an argument based upon a popular distinction between two senses of the term desire (Heathwood, 2019; Schueler, 1995, 1; Vadas 1984, 273). Though Heathwood does not himself rely upon this distinction to solve the problem of self-sacrifice (§ 3.3), it might be thought that it can be useful. According to the behavioural sense, desires are simply any psychological state that motivates. Conversely, according to the genuine attraction sense, desires are only those motivational states that are characterised by enthusiasm, appeal, interest, excitement, and attraction (Heathwood, 2019, 673). Heathwood argues that desire theories ought to restrict which desires affect well-being to only the latter sense. If we accept this restriction, then we can explain why acts of self-sacrifice do not improve well-being. This is because these acts are motivated solely by desire in the behavioural sense.

There are at least three reasons to be sceptical of this solution. Firstly, it is unclear whether the distinction that undergirds it holds up. It is possible to have a desire that begins by encompassing the features of genuine attraction, progresses onto having purely behavioural features, before returning to having the features of genuine attraction. If the two senses of desire are distinct, then it is surprising that desires can oscillate between them while retaining the same content. The fact that desires can fluctuate in this way suggests that we are dealing with one sense of desire that has both motivational and phenomenological components. This point is reinforced by the observation that both the genuine attraction and behavioural senses of desire come in degrees. For instance, people frequently experience only moderate motivation



towards outcomes that they have a strong phenomenological attraction towards. The existence of gradations in these effects suggests that desire is not divisible in this way.

Secondly, even if the distinction holds up, it is unclear whether well-being ought to be based upon it. Doing so leads to counterintuitive conclusions. People experiencing anhedonia may, for instance, not feel much enthusiasm or excitement about pursuing their desires (Tully, 2017, 4). Nevertheless, at least in mild cases, we should be reluctant to accept that they do not have desires relevant to well-being. Afterall, if they still experience desire in the behavioural sense, then it seems counterintuitive that the fulfilment or frustration of these desires never intrinsically affects their well-being.

Thirdly, even if the distinction holds up and is attractive to base well-being upon, this approach fails to satisfactorily solve the problem of self-sacrifice. This is because not all cases of self-sacrifice are motivated purely by the behavioural sense of desire. For example, consider the parent who sacrifices themselves to prevent terrible harm to their children. When faced with the consequences of inaction, they may find that their prospective sacrifice incites enthusiasm in them. Heathwood's restriction fails to capture the intuition that such acts qualify as self-sacrificial. Therefore, we ought to look elsewhere for a solution to the problem of self-sacrifice.

3.2 Do only desires motivate?

Another approach is to reject Overvold's second assumption. This is the claim that altruistic actions are always motivated by desires. The view that voluntary actions are necessarily motivated by desires is most often referred to as the Humean Theory of Motivation. Although Overvold does not explicitly subscribe to this position, his argument is nevertheless premised upon a Humean explanation of altruistic action. This is because desire theories only entail that altruistic actions improve well-being if they fulfil desires.

If we reject the Humean position, then we can claim that some altruistic actions are not motivated by desires. They may instead be motivated by alternative mental states, such as moral beliefs (Shafer-Landau, 2003, 122). On this view, the fact that we sometimes choose to act self-sacrificially does not mean that we are motivated by a desire to do so (Baber, 2007, 107). If altruistic actions are instead motivated by moral beliefs, then they do not necessarily fulfil desires. Consequently, desire theories can account for the existence of self-sacrifice by claiming that self-sacrificial actions are not motivated by desires.

Adopting a non-Humean explanation of altruistic action can harmoniously combine desire theories with the existence of self-sacrifice. However, this approach involves weighty theoretical commitments. The Humean Theory of Motivation is popular (Sinhababu, 2017; Smith, 1994). To reject it commits defenders of desire theories to a controversial position within moral psychology. Moreover, this approach means that the viability of desire theories is contingent upon whichever theory of

⁷ This argument fails if we hold the view that moral beliefs do motivate but that desire is always somehow involved in this process. For instance, it has been argued that a logical consequence of having a motivating belief is that it produces a desire (Nagel, 1970a, 30). If this is correct, then an appeal to the motivational effects of moral beliefs will not solve the problem of self-sacrifice.



motivation happens to be correct (Heathwood, 2011, 33). This is precarious grounding for any theory of well-being.

Furthermore, while this argument makes self-sacrifice possible, it nevertheless leaves us in the unenviable position of conceding that we can never sacrifice ourselves if we are motivated solely by desires. This seems wrong. Even if we accept that moral beliefs can motivate, there are still cases of self-sacrifice that are more intuitively explained as motivated by desire. For instance, sacrificing oneself to save the life of a loved one is often more intuitively explained as motivated by desire, rather than moral belief. This is especially the case when our moral beliefs conflict with our self-sacrificial action. For example, if we know that our loved one will do great harm to others if saved but we choose to save them anyway, then this seems to be a case where the motivational force of our desires outweighs that of our moral beliefs. Therefore, even if we accept that moral beliefs can motivate, self-sacrifice is nevertheless sometimes more intuitively explained as motivated by desire. Consequently, it is better to remain neutral on the independent viability of the Humean Theory of Motivation and look elsewhere for solutions to the problem of self-sacrifice.

3.3 Is well-being best served by maximising present desire fulfilment?

Another approach to solving the problem of self-sacrifice involves rejecting Overvold's third assumption. This is the idea that desire theories claim that maximising present desire fulfilment is always optimal for well-being. Chris Heathwood challenges this view. He argues that only a specific type of desire theory accepts this claim. He calls this view Life Preferentism (Heathwood, 2011, 22).

Like other desire theories, Life Preferentism claims that well-being is solely determined by the fulfilment and frustration of desires. However, it is distinctive for the additional claim that fulfilling our presently strongest balance of desires is always optimal for our well-being. Consequently, to take Overvold's example, if a father's presently strongest balance of desires leads him to kill himself for the life insurance to fund his children's education, then this is the best outcome available to him. This view entails that actions motivated by our presently strongest balance of desires cannot be self-sacrificial.

Life Preferentism is not an attractive theory. Not only does it have counterintuitive implications for self-sacrifice, but it also fails to recognise that sometimes fulfilling our presently strongest balance of desires is not in our best interests (Heathwood, 2011, 26). For example, my strongest balance of desires may lead me to socialise late into the night on a weekday evening. Nevertheless, my fatigue the following day may frustrate more and stronger desires than those fulfilled by the late-night revelry. In this case, acting on weaker desires would have been better for my well-being. The fact that Life Preferentism fails to capture this intuition means that it should be rejected.

⁸ Heathwood (2011, 22) cites John Rawls as advancing a similar position to this. The difference is that Rawls requires desires to be adequately informed and rational to count towards well-being (1971, 417).



Heathwood argues that more plausible versions of the desire theory fall under the rubric of Desire Satisfactionism (2011, 24). This view claims that well-being is determined by the total balance of desire fulfilment over frustration, when adjusted for strength, over the course of a life. This captures the intuition that sometimes we are made worse off by fulfilling our presently strongest balance of desires. In such cases, our well-being is improved by the fulfilment of our desires, but this improvement is outweighed by the desires that our action frustrates. If we accept Desire Satisfactionism, then the problem of self-sacrifice is surmountable.

Heathwood discusses two ways in which an action can be self-sacrificial. The first occurs when our action prevents future desire fulfilments. This is how Heathwood explains Overvold's case of the father who sacrifices his life to fund the education of his children (Heathwood, 2011, 27). In this case, the bulk of the harm arises from the deprivation of future desire fulfilments that are prevented by his suicide. ¹⁰ However, not all cases of self-sacrifice are well explained by an appeal to the deprivation of future desire fulfilments. In cases where the subject's life does not end as a consequence of their action, it appears far more common that the harm of self-sacrifice arises primarily as a result of frustrating desires. Heathwood writes of such cases that 'it is possible for a person to know, even vividly, that he will desire certain things in the future, and yet fail to be moved in the present to behave in such a way that those future desires will be satisfied' (2011, 28). On this view, sometimes we act self-sacrificially in the knowledge that our action will frustrate our future desires. A feature of Heathwood's position is that there must be a temporal gap between the self-sacrificial action and the frustration of at least some future desires. This is because he tacitly accepts Overvold's view that we always act upon our strongest balance of desires (Heathwood, 2011, 32–35).

Heathwood's view intuitively explains many cases of self-sacrifice. For example, one can know that large charitable donations in youth will mean sacrifices to quality of living in middle age and nevertheless not experience strong frustrations until well after the sacrifice is made. In this type of case, it seems right to claim that the action is motivated by our presently strongest balance of desires and that its harms emerge through the frustration of future desires. However, there are counterexamples that are ill-described by this psychology. Many acts of self-sacrifice feel as if we are acting against our presently strongest balance of desires at the time of action. Consider volunteering your time in a dull but important awareness-raising leafleting campaign about the pernicious health effects of air pollution, rather than spending the day in the park basking in the sun. This experience feels like you are frustrating stronger desires than you are fulfilling. This phenomenology is unaccounted for by Heathwood's position.

Due to the unviability of Life Preferentism, we ought to reject Overvold's third assumption and accept Heathwood's Desire Satisfactionism. However, without fur-

¹⁰ This approach to the harm of death has substantial philosophical precedent (Feldman, 1991; Nagel, 1970b).



⁹ This discussion of Heathwood is based on his paper 'Preferentism and self-sacrifice' (2011). More recently, Heathwood has advanced a separate argument based on a distinction between the behavioural and genuine attraction senses of desire (2019). I discuss why we ought to reject this argument in § 3.1 of this paper.

ther modification, our resultant theory is committed to a counterintuitive explanation of the psychology of some cases of self-sacrifice. While Heathwood's approach makes self-sacrifice possible, it does so by putting forward an incomplete picture of the psychological structure to which acts of self-sacrifice must conform. According to this picture, it is only through reference to future desire frustrations or the deprivation of future desire fulfilments that we can account for the harm of self-sacrifice. However, as we have seen, this explanation does not account for all cases of self-sacrifice in an intuitive way. Consequently, we need a supplementary explanation to account for residual cases of self-sacrifice.

3.4 Do desires always motivate proportionally to their strength?

An intuitive explanation of self-sacrifice should account for the experiential quality of acting against our presently strongest balance of desires. Recall that our presently strongest balance of desires is for the action that we predict will maximise the fulfilment over frustration of our presently existing desires when adjusted for the relative strength of the desires in question. It is directed at the action that we overall most desire to perform. Desire theories can capture the phenomenology of acting against our presently strongest balance of desires by rejecting Overvold's fourth assumption. This claims that human psychology is structured in such a way that the strength of our motivations is always proportional to the strength of our desires. We can term this view 'proportionalism about desire and motivation' (henceforth proportionalism). ¹¹ A commitment to proportionalism explains why Overvold assumes that we always act upon our presently strongest balance of desires.

If we reject proportionalism, then we can account for cases of self-sacrifice that are ill-described by Heathwood's Desire Satisfactionism. In such cases, our presently strongest balance of desires does not cause our action. Instead, our action is caused by weaker desires that motivate us with disproportional strength. This approach captures the experiential quality of self-sacrificial actions in cases such as that of the leaflet deliverer. Our feelings of sacrifice arise because, while we do desire to undertake the action, we have other stronger desires that are frustrated by doing so. Consequently, while the action does benefit us, its benefit is outweighed by the countervailing harms that we incur from the desires that it frustrates. If we accept this view, then there is no need to appeal to the frustration of future desires or the deprivation of future desire fulfilments to explain the harm of all cases of self-sacrifice. This approach allows desire theories to put forward a psychologically intuitive picture of self-sacrifice that accounts for cases that are unconvincingly described by Heathwood's explanation. 12

¹² Neil Sinhababu provides a supplementary explanation of the phenomenology of obligation. He points out that aversions may explain motivation in these cases rather than desires (Sinhababu, 2017, 48). If this is correct, then the leaflet deliverer's aversion to falling short morally may be what motivates their self-sacrificial action. Avoiding the frustrations that aversions produce may not generate the same feelings of satisfaction that usually emerge from fulfilling a desire. This explains why self-sacrificial actions have a different phenomenological quality. While this approach can explain the phenomenology of some cases



¹¹ Mark Schroeder uses the term 'proportionalism' to express the view that the strength of our reasons is always proportional to the strength of our desires (2007, 164–170). I am appropriating that term and applying it to the claim that the strength of our motivations is always proportional to the strength of our desires.

Rejecting proportionalism requires making a distinction between motivation and desire. Fortunately, this distinction has ample precedent in moral psychology. For instance, one view claims that motivation is a disposition towards an action. This may include a disposition to do and to think about doing an action (Gregory, 2021, 30). This means that under the right conditions we would think about and undertake that action (Firth, 1952, 320). In contrast, desires are not always connected with actions. It is possible to desire something that no possible action could affect. For instance, I may desire that the English football team wins a major trophy again while being in no position to affect that outcome. Moreover, while it is conceivable that all motivation is explained in terms of desire, it is at least conceptually possible that things other than desires could motivate us. In § 5 I argue that there are strong independent reasons to reject proportionalism.

3.5 Section summary

Heathwood's argument makes self-sacrifice possible for desire theories. Nevertheless, he puts forward a picture of self-sacrifice that is incomplete. If we reject proportionalism, then we do not need to appeal to the deprivation of future desire fulfilments or to the frustration of future desires to explain how every act of self-sacrifice harms us. Instead, we can explain some cases of self-sacrifice as involving weaker desires that motivate us with disproportional strength. This means that these actions are aggregate harms. This explanation better captures the experiential quality of some cases of self-sacrifice. Therefore, it marks an improvement on Heathwood's position.

4 The problem of depression

The problem of depression for desire theories was introduced in a 2017 paper by Ian Tully, and later developed in Andrew Spaid's 2020 PhD thesis. In his paper, Tully argues that desire theories are unable to capture the intuition that severe depression is a state of harm.¹³ To make his case, Tully appeals to a distinction between two species of depression: motivational and consummatory anhedonia (Tully, 2017, 4). The former occurs when we lose interest in pursuing many of the things that we previously enjoyed. Whereas the latter occurs when we consistently fail to experience pleasure from satisfying our desires. He argues that experiencing both concurrently indicates that our desires have been weakened or eliminated. On this view, motivational anhe-

¹³ Tully refers to states of ill-being rather than states of harm. However, for simplicity, I will refer to states of harm instead. Some people conceptualise ill-being as a distinct value to well-being. This fits well with views that claim that well-being has multiple non-commensurable components (Keller, 2009, 664–665). However, as Tully points out, it counterintuitively entails that an individual could have high well-being *and* high ill-being (2017, 3–4). Moreover, the practice of deliberation in response to prospective trade-offs between benefits and harms suggests that these values are commensurable (Kelly, 2008, 372).



of self-sacrifice, I nevertheless think that we ought to supplement it by rejecting proportionalism (which Sinhababu does). If aversions were the sole motivator of self-sacrifice, then we ought to expect feelings of relief to accompany acts of self-sacrifice. This is seldom the case.

donia is taken as evidence that we lack desire; whereas consummatory anhedonia is taken as evidence that we did not desire the outcome in the first place. ¹⁴

Undoubtedly, severe depression is a paradigm case of harm. However, Tully argues that desire theories fail to endorse this conclusion. This is because, on these views, for something to diminish well-being it must frustrate desires. If severe depression weakens or eliminates desires, rather than frustrates them, then it is not a state of harm. I will refer to this argument as 'the problem of depression'.

The problem of depression can be summarised as follows:

P1 Severe depression is a state of harm that viable theories of well-being must capture.

P2 Desire theories entail that severe depression is not a state of harm.

C Therefore, desire theories are unviable.

This argument is valid. Moreover, premise one is overwhelmingly intuitive. However, premise two is contestable. If it can be shown to be incorrect, then desire theories can survive this objection.

This section is structured in the following way: § 4.1 argues that an effect of severe depression is that it decreases motivation while leaving some desires intact; § 4.2 examines counterarguments to this claim; § 4.3 explains how desire theories can deal with residual cases that are less well explained by this approach; § 4.4 concludes with a summary.

4.1 Depression and the dampening of motivation

One way of answering the problem of depression involves challenging the characterisation of severe depression that it is premised upon. Tully takes the coexistence of consummatory and motivational anhedonia to indicate lack of desire. There is something intuitive about this explanation. If both pleasure and motivation are lacking, and there is no countervailing evidence, then it seems reasonable to assume that desire is absent. If he is right, then desire theories struggle to account for the harm of severe depression.

However, there is good reason to think that this characterisation of severe depression is incomplete. This is because it fails to capture some aspects of its experiential quality. The testimony of people with severe depression illuminates this. Severe depression is often described as decreasing motivation while leaving underlying desires intact. Depressed people commonly report having desires but being unable to motivate themselves to fulfil them. In the words of Steven Swartzer, 'This is part of what is so <u>frustrating</u> about such experiences. That one is unable to engage in activi-

¹⁴ The view that depression weakens or eliminates desires has philosophical precedent (Smith, 1994, 135; Stocker 1979, 744).



ties that one cares strongly about is part of why such situations are so heartbreaking' (Swartzer, 2015, 9, underlining added). 15

To illustrate this point, Swartzer appeals to the testimony of a man who is unable to motivate himself to attend his son's wedding:

'I knew that my son's wedding would be emotional ... and that anything emotional, good or bad, sets me off. I wanted to be prepared. I'd always hated the idea of electroshock therapy, but I went and had it anyway. But it didn't do any good. By the time the wedding came, I couldn't even get out of bed. It broke my heart, but there was no way that I could get there' (Swartzer, 2015, 8).

In this case, the man's desire to attend the wedding appears to persist, even though it fails to motivate him. Given his testimony, it seems perverse to insist that the man did not *really* desire to attend his son's wedding. Consequently, there are good reasons to think that severe depression has harmed him by dampening the motivational force of his desires and thereby frustrating them. On this view, motivational anhedonia is explained as an aspect of severe depression that affects motivation directly.

Nevertheless, one might concede a strong desire on the man's behalf but postulate stronger countervailing aversions. If aversions have a different phenomenological character to desires (Sinhababu, 2017, 48), then the turmoil reported in Swartzer's example may be explained by them. However, it strikes me that an appeal to aversions less convincingly captures the phenomenology of severe depression than the dampening of motivation. This is starker in other cases of severe depression that involve no such turmoil. For instance, severely depressed people sometimes struggle to motivate themselves to get out of bed despite having many desires whose fulfilment is dependent upon doing so. People in this situation tend to report a lack of motivation for action, rather than aversions to action. In such cases, the most philosophically parsimonious explanation is that desires are failing to motivate, rather than being outweighed by unreported countervailing aversions.

Cases like this are counterexamples to Tully's characterisation of severe depression. The most intuitive way of rectifying his position is to claim that severe depression leaves some desires intact while dampening their motivational force. On this view, severe depression harms us by making us less able to fulfil our desires. Consequently, it is a state laden with desire frustrations. This explanation is premised on the view that desires sometimes fail to motivate proportionally to their strength. Therefore, it requires the rejection of proportionalism (§ 5).

4.2 Counterarguments to the view that depression dampens motivation

Tully anticipates the argument that severe depression may dampen motivation while leaving desires intact. He considers the possibility that it may simply mask our desires (Tully, 2017, 10–12). Masked desires are undetectable through introspection and do not motivate. If severe depression masks desires in this way, then desire theories can recognise it as a state of harm. However, Tully finds this suggestion implausible for two reasons. Firstly, he points out that if desires are masked by depression, then desire

¹⁵ Other writers characterise depression similarly (Arpaly & Schroeder, 2014, 126; Schroeder 2004, 31–32).



theories implausibly entail that their fulfilment improves well-being (Tully, 2017, 12). To illustrate this point, he constructs an example of having a favourite meal prepared for dinner while severely depressed (Tully, 2017, 11–12). He takes the absence of feelings of satisfaction to be evidence that well-being is not improved. Secondly, he points out that when depression dissipates, not all desires return. He considers this to be evidence that severe depression dampens desires rather than motivation.

The first thing to say about this argument is that the experiential case for severe depression dampening motivation does not rest upon the postulation of masked desires. Whereas Tully suggests that masked desires are imperceptible through introspection, severe depression often leaves us painfully aware of our desires. I am not seeking to defend the idea that severe depression masks desire. It is enough to show that it dampens motivation and leaves some desires intact and frustrated. Nevertheless, I will address Tully's argument on his own terms.

There are at least three responses to his first objection. Firstly, we can point out that desire theories do not require feelings of satisfaction to be present for well-being to be improved. Desires can be fulfilled without generating these feelings. For instance, this happens when we do not find out about our desire's fulfilment. Nevertheless, while feelings of satisfaction are not required for desire fulfilments to improve well-being, the lack of these feelings is often a good indicator that we have not fulfilled our desire. Consequently, without supplementation, this response is unconvincing.

A second response is to argue that Tully's example does not describe genuine desire fulfilment, but rather a mere simulacrum. Desires may be more fine-grained than is sometimes assumed (Graff Fara, 2013; William, 2016, 213). In Tully's example, it may be that his desire is to *enjoy* his favourite meal. This can be expressed as a proposition with two distinct components. Tully desires to have his favoured meal prepared for him *and* to take pleasure in eating it. If his desire is fine-grained in this way, then both components need to be present for it to be fulfilled. If severe depression prevents pleasure from emerging, then this inevitably frustrates his desire. This explains why we are not always benefitted by the apparent fulfilment of masked desires. Some desires cannot be fulfilled without pleasure being present, and consummatory anhedonia prevents pleasure from emerging in the normal way. On this view, pleasure is built into the fulfilment conditions of some desires.

However, if we cannot introspectively detect the existence of a desire, then often the most intuitive explanation is to accept that no desire is present. Consequently, a third response is to accept that severe depression has eliminated the desire. This means acknowledging that severe depression sometimes weakens or eliminates desires. This also explains why not all desires return when depression lifts. Although, we should note that desires do change over time anyway. The fact that people emerge from severe depression with different desires does not undermine the view that severe depression masks desires. It may simply have masked desires that gradually changed throughout the course of depression in the normal way.

Accepting that desires can be weakened or eliminated by severe depression does not mean that desire theories cannot recognise it as a state of harm. Providing that severe depression also frustrates desires, then these theories are perfectly able to recognise its harm. On this view, the removal of some desires is a benign effect of severe depression, while its harm emerges from the desires inevitably frustrated through



their inability to motivate and by severe depression's prevention of pleasure from emerging in the normal way.

4.3 Remaining problem cases

I have argued that desire theories can recognise severe depression as a state of harm. Nevertheless, there remain problem cases that appear to be less well explained by these arguments. Andrew Spaid provides a fictionalised example:

'Jane is diagnosed with clinical depression, and understands that she is depressed. She also understands that an effective treatment for her depression is available. In other words, she understands that with treatment she would come to have the desires and the joys most non-depressed people have—in short, a normal life. Nevertheless, Jane refuses treatment for her current episode of depression, claiming that she does not care about the treatment outcome—she sees no point in regaining the desire to live because she believes nothing is worth doing' (Spaid, 2020, 28–29).

Spaid supplements this fictionalisation with a wealth of supportive testimony from sufferers of depression (2020, 38–39). These cases suggest that sometimes the primary effect of depression is the weakening or elimination of desires. If this is right, then desire theories may struggle to recognise these cases as states of harm.

One way of responding to cases like this is to claim that people like Jane have desires to be badly off. This may be in response to feelings and beliefs of low self-worth that are sometimes experienced by depressed people. On this view, the fact that Jane refuses treatment is not taken as evidence of an absence of desire. Rather, it is taken as evidence of a desire to remain badly off. This explains why Jane actively refuses, rather than simply expresses indifference to, treatment. If this is right, then desire theories can recognise Jane's severe depression as a state of harm. This is because Jane has a desire to remain badly off that motivates her with disproportional strength to frustrate her other desires. Accordingly, she has many frustrated desires. The example of having a general desire to be badly off is somewhat confected. It seems far more common that severe depression facilitates the growth of self-destructive desires that obstruct the fulfilment of other stronger desires. Nevertheless, we can bracket these types of cases under the umbrella of desires to be badly off. The acquisition of such desires seems to be an effect of severe depression that some people experience.

The dampening of motivation can account for the harm of many cases of severe depression. Other harms are explained by depression's prevention of pleasure from emerging in the normal way. Still others can be explained by desires to be badly off. However, some residual cases may remain. Sufferers of complete conative collapse

¹⁶ Readers may be concerned that if Jane fulfils her desire to be badly off, then according to desire theories her well-being is improved by this desire's fulfilment. Heathwood responds to this by accepting that the fulfilment of desires to be badly off provides some benefit. However, this is outweighed by the harm of frustrating other desires. Consequently, fulfilling a desire to be badly off is an aggregate harm (Heathwood, 2005, 501–502). Nevertheless, a paradox for desire theories can emerge when the fulfilment of a desire to be badly off generates enough well-being to make a person's life overall good for them (Bradley, 2007). I set aside this issue here, as it is a separate problem. Other writers have proposed potential solutions (Dorsey, 2012; Skow, 2009).



may fall into this category (Tully, 2017, 6). Tully claims that people in this condition are entirely devoid of motivation and largely unresponsive. Nevertheless, they remain conscious and aware of their surroundings. It is possible that such people lack all, or almost all, desires.

To account for this group, desire theories can appeal to a deprivation account. Some views about the harm of death take this approach (Feldman, 1991; Nagel, 1970b). On these views, death is bad because it deprives us of future goods. Consequently, death is not an intrinsic harm but rather an instrumental or extrinsic harm (Bradley, 2008, 300). We can apply a similar argument to severe depression. On this view, some cases of severe depression instrumentally harm us by making us less able to generate and fulfil new desires.

The deprivation of future goods is undoubtedly an effect of severe depression. However, deprivation is not an intrinsic harm. Consequently, it is difficult to see how severe depression makes us badly off, rather than simply worse off. While deprivation accounts can explain why severely depressed people lack well-being, they cannot characterise severe depression as a state of intrinsic harm. Therefore, this approach does not adequately account for the harm of most cases of severe depression.

Nevertheless, I think that it goes some way toward explaining residual cases. When we think of severe depression as a paradigm case of harm, we tend to imagine someone unmotivated to pursue their desires, someone who fails to experience pleasure when pursuing their desires, or someone plagued by self-destructive desires that keep them badly off. Desire theories can explain why people who experience any of these effects are harmed by severe depression. Conversely, if we imagine someone truly devoid of desires, then it seems less clear that they are in a state of harm. Afterall, someone in this condition appears to be in a better position than someone with many frustrated desires. Even if our intuitions persist that desirelessness is a state of harm, rather than one of deprivation, then I think that these intuitions should not be considered decisive. Our criterion for adequate theories of well-being is that they provide intuitive results in paradigm cases of benefit and harm (Fletcher, 2016, 10). The atypical case of complete conative collapse is by no means a paradigm case. Consequently, if this is a bullet to bite, then it is a far smaller one than Tully and Spaid's arguments suggest.

4.4 Section summary

I have argued that if we reject proportionalism, then we can conceptualise severe depression as dampening motivation, rather than solely dampening desire. This allows desire theories to account for a large part of its harm. On this view, severe depression harms us by making us less able to fulfil our desires. It is consequently a state laden with desire frustrations. This view captures the experiential quality of severe depression more accurately than Tully's characterisation. Nevertheless, it explains some residual cases of severe depression less well. For these cases, I have argued that we can explain its harm as arising from its prevention of pleasure. This inevitably frustrates desires that specify the presence of pleasure in their fulfilment conditions. I have also argued that severe depression sometimes gives rise to self-



destructive desires to be badly off. For residual cases of complete desirelessness, I have appealed to a deprivation account to explain how this state affects well-being.

5 Proportionalism about desire and motivation

I have argued that rejecting proportionalism can serve as the basis of a response to the problems of self-sacrifice and depression. Our consequent understandings of self-sacrifice and depression make them intelligible as harms to desire theories. This makes these theories robust to two problems that are sometimes said to undermine them. I turn now to consider the viability of rejecting proportionalism. Making a comprehensive case for this would involve a far wider discussion of the nature of desire and motivation than this paper permits. Instead, I limit myself to claiming that rejecting proportionalism is a plausible view that is independently motivated and compatible with popular views within moral psychology. § 5.1 examines how rejecting proportionalism can enrich our understanding of other psychological phenomena. § 5.2 alleviates the worry that rejecting proportionalism commits us to controversies within our wider moral psychology.

5.1 Additional counterexamples to proportionalism

Alongside self-sacrifice and depression, there are other psychological phenomena where strength of motivation does not seem to track strength of desire. Weakness of will is one such example (Gregory, 2021, 34–35). Perhaps the most common variety of this occurs when we fail to defer gratification (Sinhababu, 2017, 38). In such cases, we appear to pursue weaker desires that motivate us with disproportional strength. After experiencing weakness of will, we often reflectively feel that we have acted against our own strongest balance of desires. Moreover, in cases of 'clear-eyed' weakness of will, we may even feel that we are undermining those desires as we act. This phenomenology can be captured by rejecting proportionalism.

Another case emerges when we forget our desires and they thereby fail to motivate us (Gregory, 2021, 34). For instance, we may forget our desire to pay off a credit card before incurring fines. In this type of case, it appears that our desire has failed to motivate us because we did not keep it in mind when making decisions. The existence of non-occurrent beliefs is well-established. Given that desires share a similar structure to beliefs (Gregory, 2012), we should not be surprised that non-occurrent desires also exist. If this is right, then proportionalism is incorrect because forgotten desires do not motivate.

Additionally, some desires are unusually strongly felt. This is often the case with desires that are accompanied by strong emotions (Raibley, 2010, 598–599). Anger often seems to amplify the motivational force of desires (Nussbaum, 2016, 96). For instance, road rage can cause fleeting but strongly felt desires to shout at other motorists. In such cases, we are reticent to describe our desires as strong, as they quickly dissipate when our attention shifts. A more intuitive description of the psychology of these cases claims that they involve relatively minor desires that motivate with disproportional strength.



The existence of self-sacrifice, depression, weakness of will, forgotten desires, and unusually strongly felt desires undermine the attractiveness of proportionalism. None of these cases when taken independently are decisive arguments. There may well be plausible characterisations of some of them that do not require the rejection of proportionalism. Nevertheless, these examples collectively illustrate the extent to which rejecting proportionalism can enrich our understanding of a range of psychological phenomena.

5.2 How rejecting proportionalism affects our wider understanding of desire

Some readers may worry that rejecting proportionalism requires taking controversial positions within philosophy of desire. If this is the case, then my solutions to the problems of self-sacrifice and depression will depend upon whichever theory of desire happens to be correct. Earlier, I dismissed the idea of rejecting the Human Theory of Motivation to solve the problem of self-sacrifice partly because it would commit us to taking strong positions on controversies within moral psychology (§ 3.2).

It is true that ultimately my solutions to problems within philosophy of well-being depend upon findings within moral psychology. Nevertheless, the rejection of proportionalism is far less controversial than the rejection of the Human Theory of Motivation. Rejecting proportionalism is compatible with a range of different theories of desire. If one takes desire to be characterised by a disposition to experience pleasure when imagining fulfilling desires (Strawson, 2010), then there is no reason why we need to commit to proportionalism. If one takes desire to be primarily characterised by reward and learning responses (Schroeder, 2004), then there is no reason why we need to commit to proportionalism. If one takes desire to be a subset of our normative beliefs (Gregory, 2021), then there is no reason why we need to commit to proportionalism. Indeed, the only views of desire that are in tension with the rejection of proportionalism are a subset of motivational theories of desire.

According to motivational theories of desire, desires are defined by their effects on motivation. The most reductive of these theories claim that desires simply *are* motivation (Dancy, 2000, 85). If this is correct, then proportionalism is trivially true. However, the reductive view is plainly incorrect. We often desire things that fail to motivate us because we lack a means-end belief to attain them. Moreover, our motivation is often dampened when we predict that our actions have only a limited chance of fulfilling our desires. In both these cases, desires remain constant, while motivation fluctuates according to our beliefs. Yet the reductive view must counterintuitively claim that in such cases our desires are weakened or eliminated. Consequently, the eliding of desire and motivation is implausible. Alternative motivational accounts of desire claim that desires are dispositions to be motivated toward outcomes. A disposition to be motivated does not entail that every instance of desire will motivate us. Therefore, this position is compatible with the rejection of proportionalism.



6 Conclusion

This paper has addressed two problems for desire theories of well-being. These problems claim that these theories are unable to account for the harm of self-sacrifice and severe depression. Far from this being the case, I have argued that suitably enriched understandings of these phenomena are well-explained by desire theories of well-being that reject proportionalism about desire and motivation. I then argued that rejecting proportionalism about desire and motivation is independently motivated. If I have been successful, then discoveries within moral psychology will have resolved problems within philosophy of well-being. This is an interesting finding. There may be other problems thought to bedevil otherwise attractive theories of well-being that can similarly be solved by investigation into moral psychology.

Acknowledgements A debt of gratitude is owed to Alex Gregory for comments on multiple full drafts of this paper. Thanks are also due to Brian McElwee, Giulia Felappi, Jonathan Way, Elliot Porter, Mariana Guerra, and two anonymous reviewers for comments on full drafts. Thanks also to St.John Lambert and Charlotte Unruh for feedback on sections of this paper. This paper has also benefitted from audience feedback from two postgraduate seminars at the University of Southampton and from the 2022 PhiGS Colloquium at the University of Kent. It was also made possible by generous funding from the Royal Institute of Philosophy's Jacobsen scholarship.

Ethics Declaration The author declares that he has no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adams, R. (1999). Finite and infinite goods: A framework for ethics. New York: Oxford University Press.

Arpaly, N., & Schroeder, T. (2014). In praise of desire. New York: Oxford University Press.

Baber, H. E. (2007). Adaptive preference. Social Theory and Practice, 33(1), 105-126.

Bradley, B. (2007). A paradox for some theories of welfare. Philosophical Studies, 133(1), 45-53.

Bradley, B. (2008). The worst time to die. Ethics, 118(2), 291–314.

Brandt, R. (1972). Rationality, egoism, and morality. Journal of Philosophy, 69(20), 681-697.

Brandt, R. (1979). A theory of the good and the right. Oxford: Clarendon Press.

Brewer, T. (2006). Three dogmas of desire. In T. Chappell (Ed.), *Values and virtues: Aristotelianism in contemporary ethics* (pp. 253–285). Oxford: Clarendon Press.

Carson, T. (2000). Value and the good life. Notre Dame: University of Notre Dame Press.

Dancy, J. (2000). Practical reality. Oxford: Oxford University Press.

Dorsey, D. (2012). Subjectivism without desire. The Philosophical Review, 121(3), 407-442.

Feldman, F. (1991). Some puzzles about the evil of death. The Philosophical Review, 100(2), 205-227.

Firth, R. (1952). Ethical absolutism and the ideal observer. *Philosophy and Phenomenological Research*, 12(3), 317–345.

Fletcher, G. (2016). The philosophy of well-being: An introduction. Oxon: Routledge.



Graff Fara, D. (2013). Specifying desires. Noûs, 47(2), 250-272.

Gregory, A. (2012). Changing direction on direction of fit. Ethical Theory and Moral Practice, 15, 603–614.

Gregory, A. (2021). Desire and belief: A theory about motivation and rationality. Oxford: Oxford University Press.

Heathwood, C. (2005). The problem of defective desires. Australasian Journal of Philosophy, 83(4), 487–504.

Heathwood, C. (2011). Preferentism and self-sacrifice. Pacific Philosophical Quarterly, 92(1), 18-38.

Heathwood, C. (2019). Which desires are relevant to well-being? Noûs, 53(3), 664-688.

Keller, S. (2009). Welfare as success. Noûs, 43(4), 656-683.

Kelly, C. (2008). The impossibility of incommensurable values. Philosophical Studies, 137(3), 369–382.

McDaniel, K., & Bradley, B. (2008). Desires. Mind New Series, 117(466), 267-302.

Murphy, M. (1999). The simple desire-fulfillment theory. Noûs, 33(2), 247–272.

Nagel, T. (1970a). The possibility of Altruism. Princeton, NJ: Princeton University Press.

Nagel, T. (1970b). Death. Noûs, 4(1), 73-80.

Nussbaum, M. (2016). Anger and forgiveness: Resentment, generosity, justice. New York: Oxford University Press.

Overvold, M. (1980). Self-interest and the concept of self-sacrifice. *Canadian Journal of Philosophy*, 10(1), 105–118.

Parfit, D. (1984). Reasons and persons. New York: Oxford University Press.

Raibley, J. (2010). Well-being and the priority of values. Social Theory and Practice, 36(4), 593-620.

Rawls, J. (1971). A theory of justice. Cambridge, MA: Belknap Press.

Sarch, A. (2013). Desire satisfaction and time. *Utilitas*, 25(2), 221–245.

Schroeder, T. (2004). Three faces of desire. New York: Oxford University Press.

Schroeder, M. (2007). Slaves of the passions. Oxford: Oxford University Press.

Schueler, G. F. (1995). Desire: Its role in practical reason and the explanation of action. Cambridge, MA: MIT Press.

Shafer-Landau, R. (2003). Moral realism: A defence. Oxford: Clarendon Press.

Sidgwick, H. (1981). [1907]. The methods of ethics (7th ed.). Indianapolis, IN: Hackett Publishing Company

Sinhababu, N. (2015). Advantages of propositionalism. *Pacific Philosophical Quarterly*, 96(2), 165–180.

Sinhababu, N. (2017). *Humean nature: How desire explains action, thought, and feeling*. Oxford: Oxford University Press.

Skow, B. (2009). Preferentism and the paradox of desire. *Journal of Ethics and Social Philosophy*, 3(3), 1–16.

Smith, M. (1994). The moral problem. Oxford: Blackwell Publishing.

Spaid, A. (2020). Desire satisfaction theories and the problem of depression. PhD. University of Nebraska. Stocker, M. (1979). Desiring the bad: An essay in moral psychology. Journal of Philosophy, 76(12), 738–753.

Strawson, G. (2010). Mental reality (2nd ed.). Cambridge, MA & London: MIT Press.

Swartzer, S. (2015). Humean externalism and the argument from depression. *Journal of Ethics and Social Philosophy*, 9(2), 1–16.

Thagard, P. (2006). Desires are not propositional attitudes. *Dialogue*, 45(1), 151–156.

Tully, I. (2017). Depression and the problem of absent desires. *Journal of Ethics and Social Philosophy*, 11(2), 1–13.

Vadas, M. (1984). Affective and non-affective desire. Philosophy and Phenomenological Research, 45(2), 273–279.

Velleman, D. (1991). Well-being and time. Pacific Philosophical Quarterly, 72(1), 48-77.

William, E. (2016). Preferences' significance does not depend on their content. *Journal of Moral Philoso-phy*, 13(2), 211–234.

Wolf, S. (1982). Moral saints. Journal of Philosophy, 79(8), 419-439.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted



manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

