

Desire's Own Reasons*

Forthcoming in the Journal of the American Philosophical Association. Please cite the published version

1. Introduction

It is a truism that people's desires change over time. What one preferred to eat as a child need not reflect the gustatory preferences of one's adult self. A person may lose their desire to attend rock concerts at one point and acquire it afresh at another. What one wants to achieve in life can go through dramatic changes: a desire to excel at some activity can fade away and give place to some completely different aspiration. Given that desires change, can we say that an adoption or loss of some desire is a change for the better?

* My gratitude goes out to the audiences at the Department of Philosophy Talk-Shop at Harvard University, the 2019 meeting of Central APA in Denver, the Department of Philosophy work-in-progress seminar at Tartu University and Tokyo Workshop on Agency and Rationality 2019. Special thanks to Rachel Achs, Derek Baker, Selim Berker, Alex Davies, Sandy Diehl, Paul Katsafanas, Robbie Kubala, Kengo Miyazono, Patrick Shirreff and two anonymous referees of this journal.

The research in this paper was supported by JSPS Postdoctoral Fellowship for Research in Japan (Standard), JSPS KAKENHI (19F19762).

The answerability of this question seems to depend on whether there are reasons for forming or giving up a desire. If there are, we can say that a person's desires have changed for the better if they have better reasons for the present desires than their past ones. We certainly can at least *attribute* such reasons to others and ourselves. It is a further question, however, whether any of the attributed reasons derive from the nature of desire itself or whether they are externally imposed. The normative force of reasons that we attribute to an attitude is, as some authors would say, opaque: it is not apparent *how* a reason bears on the attitude (D'Arms & Jacobson 2014, 217). Almost any contentful mental state can in principle be deemed reasonable or unreasonable, but only in some cases does this evaluability derive from the nature of the state itself. For instance, it is plausible that *belief* is the type of attitude that is by nature such that reasons apply to it (see Kelly 2002, 177). According to one popular view, these can only be evidential reasons (Adler 2006; Shah 2006).

In this paper, I will argue that there are reasons that apply to desire in virtue of its nature and in the right kind of way. I call them 'Desire's Own Reasons' (DOR in short). I will claim that DOR for having a desire are considerations that bear on whether the satisfaction of the desire would also satisfy the agent who has it. Such reasons are effective in the sense that desire-regulation mechanisms are directly sensitive to them, and they are accessible in that they appear intelligible to the agents who have them. The approach in this paper is naturalist in that I assume that an account of DOR has to respect the facts about how our desire-regulation mechanisms actually operate and how DOR can figure in their operation.

In the literature about attitudinal normativity, there is a much-discussed contrast between right and wrong kinds of reasons (see Rabinowicz & Rønnow-Rasmussen 2004; 2006; D'Arms & Jacobson 2000). It is fair to say that desire's own reasons roughly correspond to desire's right kind of

reasons: when a consideration bears on an attitude in virtue of its nature, it is also the right kind of reason for that attitude. My use of the term “DOR” instead of “right kind of reasons” mainly stems from a terminological preference. The former carries more explicitly on its sleeve the idea that these reasons constitutively apply to desires. I also want to leave open a possibility that some considerations that may intuitively seem as the right kind of reasons do not actually count as DOR because they do not play the required role in desire-regulation.

Note also that the view that reasons for an attitude can be derived from the nature of that attitude is a form of constitutivism. Constitutivist accounts of normativity have drawn a wide range of criticism, both in the context of mind and action (see Enoch 2006; Côté-Bouchard 2016). It is not within the scope of this paper to respond to this criticism. What I hope to show is that *if* reasons for desire are derivable from the nature of desire, then the view of DOR I propose is the most promising one to account for them.

My strategy is as follows. In Section 2, I will first argue that DOR should be considerations that derive from the desire’s constitutive standard of correctness, in that DOR to want something explain why such a desire would be a correct attitude to have. In Section 3, I will consider some candidates for this standard and argue that they are not satisfactory. Then, in Section 4, by relying on an empirically and phenomenologically informed understanding of how desires are regulated and updated, I will defend the claim that desire’s standard of correctness from which DOR can be derived is subjective satisfaction that is distinct from desire satisfaction, and that DOR are considerations that bear on the former. Finally, in Section 5, I will address and respond to some objections to this proposal.

Before moving on, I should clarify what I mean by desire. I take desires to be attitudes whose paradigmatic instances involve motivational, affective, attentional and (possibly) evaluative

components (see Goldman 2017). To have a desire for x is to have an attitude toward x which involves being motivated to act in ways that promote x, being disposed to feel positively about the prospect of x and having thoughts about x occupy one's attention. I also follow Goldman in understanding desire in terms of the cluster of those components, without identifying desires with their neural basis. For the latter view, see Schroeder (2004) and Arpaly & Schroeder (2014). According to Arpaly and Schroeder, to have a desire with some content is for one's reward system to constitute that content as a reward. Although I do not commit myself to their account, I am fairly confident that what is said in this paper about desire-regulation can also be said when one adopts the reductive view.

By using "x" as a tag for whatever is the content of desire, I want to leave open what the exact nature of that content is. What matters here is that x constitutes the satisfaction condition of desire: desire for x is satisfied just in case x obtains. Finally, I am here focusing solely on intrinsic, as opposed to instrumental, desires. It is plausible that the latter are governed by the norms of instrumental rationality, and the respective reasons that apply to them are derivable from those norms.

2. Desiderata for DOR

In this section, I am going to look more closely at what is required from a satisfactory account of DOR. As said in the introduction, DOR are considerations that count in favor of or against having a desire in virtue of the kind of attitude it is. What does that mean exactly?

I take it that reasons can bear on an attitude in virtue of its nature if the attitude can be correct or incorrect, where its standard of correctness is not something we just attribute to it but derives from

the kind of attitude it is. By admitting a constitutive standard of correctness, an attitude is fit to be regulated in response to considerations that bear on whether its standard is met or not. In the case of desire, DOR are considerations that, if they exist, play such a role. Compare this requirement with the right kind of reasons for belief. It has been suggested that these consist of truth-relevant considerations because the latter bear on whether beliefs meet their constitutive standard of correctness (Sharadin 2015, 388).

I propose the following schema for a satisfactory account of DOR, where S is a subject, p is a fact (true proposition) and x is the content of desire (presumably a state of affairs).

SCHEMA: p is a DOR for S to want x just in case p explains why wanting x meets desire's constitutive standard of correctness.

In order to rule out considerations that do not apply to desires in the right kind of way, there are further desiderata that an account of DOR should satisfy. First, I take it that DOR should make a substantial difference to the desire-regulation mechanisms of an individual. Analogously, it has been pointed out that right kind of reasons, as opposed to those of the wrong kind, exhibit motivational asymmetry, in that it is easier to believe on the basis of the right kind of reasons (Schroeder 2012, 459). For instance, it is easier to believe that p on the basis of evidence for p's truth, as opposed to the practical value of having that belief. I think that DOR should also be such that they can directly bear on an agent's desire-regulation, unlike considerations that do not constitutively bear on desires. DOR should be directly effective with respect to one's desires, so that one doesn't have to take any intermediate deliberative steps to adjust one's desire accordingly

in response to DOR. I call this requirement Effectiveness. Effectiveness reflects the fact that DOR are considerations to which desires are naturally sensitive.

It is important to see how the idea of DOR bearing on the constitutive standard of desire and their Effectiveness fit together. That an attitude type has a constitutive standard means, among other things, that the way in which tokens of that type are acquired, updated and given up is geared toward meeting that standard. In order for the regulation of attitudes to be geared toward meeting the standard, they have to be robustly sensitive to considerations that bear on whether the standard is met or not, i.e., to the attitude's own reasons. And for an attitude to be robustly sensitive to a consideration is for the latter to be Effective with respect to the former. Applying this to desires, since DOR, if they exist, govern desires in a way that they are apt to satisfy their constitutive standard of correctness, DOR should play a role in desire-regulation in a way that satisfies the Effectiveness requirement.

Second, I take DOR to be such that a person who is aware of them in the right kind of way is able to see that having a desire for x in response to them is fitting. Agents should be able to see that responding to DOR is a proper response to have. Only then can DOR figure in personal-level desire regulation. I call this requirement Accessibility. Without Accessibility, one can reasonably ask why an agent should follow DOR in their desire-formation. If DOR were Effective in regulating desires in accordance with their Standard, but without agents being in a position to acknowledge their normative significance, they could be disregarded in our conscious deliberation because they provided no recognizably normative guidance.

But what does the right kind of awareness consist in? The agent should be properly situated with respect DOR. For instance, an agent who is under cognitive load or distracted need not see the intelligible connection between a consideration and a desire, without this excluding that

consideration from being a DOR for that desire. On the other hand, we cannot be too demanding with respect to the right kind of awareness either. For instance, requiring rational reflection in order to arrive at the proper kind of awareness would be too demanding. People who do not engage in explicit reflection about reasons can be equally responsive to DOR in their desire-formation as people who do. It suffices when it appears to the agent that having a desire in response to DOR is a fitting response to have, where this appearance is cashed out in terms of nonconceptual content (Hawkins 2008, 257) or know-how (Sylvan 2015, 604). I do not thus think that Accessibility condition assumes that the agent to whom the consideration is Accessible has to be ideally rational. Assuming the latter would run counter to the overall naturalist approach of this paper in which case we are interested in reasons that govern our actual desire-regulating mechanisms (I thank an anonymous reviewer for pressing me on this.)

To sum up the desiderata for an account of DOR, DOR for someone to want something are considerations that explain why wanting it satisfies the constitutive standard of desire. Furthermore, DOR are considerations that are fit to directly regulate desires (Effectiveness) in a way that is intelligible to the subject (Accessibility).

3. Desire's standard: some candidates

What could be the constitutive standard of correctness for desire from which one can derive DOR? In this section, I will consider three candidates, none of which is satisfactory. Note that the aim is not to argue that these standards do not apply to desires, I am willing to embrace pluralism about standards; the claim is just that they do not allow us to derive DOR from them.

The most immediate candidate that comes to mind is the satisfaction of desire itself. It seems to follow from the basic understanding of desire that the latter is successful when it is satisfied, i.e., when its content obtains. Take, for instance, David Papineau's account, according to which the satisfaction condition of a desire is its effect, the production of which is the desire's biological selected function (Papineau 1993, 58; 1998, 10). Although Papineau doesn't put it in those terms, we could then say that desire meets its constitutive standard of correctness when its content obtains. If we take that route and try to derive DOR from desire satisfaction as the standard, the resulting view would be:

DSR (desire-satisfaction reasons): p is a DOR for S to want x just in case p explains why the desire for x is satisfied

But this gives us an awkward result: DSR would not give any guidance as to whether one should want x or not. If there is some fact that explains why x obtains, this does not have any bearing on whether one should have the desire for x. If anything, it is a reason to *believe* that x obtains, not a reason to want it to obtain. Thus, even if desire satisfaction is (quite plausibly) desire's correctness condition, DOR cannot be derived from it.

Another candidate for the standard is the satisfaction of some second-order desire. According to this proposal, a desire for x meets its constitutive standard when S has a second-order desire toward having it. If having the former satisfies the latter, the former would be a correct desire to have. The appropriateness of having a desire is thereby evaluated in terms of some other desire. The resulting

DOR would then be considerations that explained why having a desire would satisfy some second-order desire.

DDR (second order desire reasons): p is a DOR for S to want x just in case p explains why the desire for x satisfies a second-order desire to have the desire for x.

My main objection to DDR is that it leaves unexplained what the standard of the relevant second-order desire is. This suggests that if DOR were to derive from second-order desires, the latter would not be subject to DOR themselves. Here we have a dilemma: the possible ways in which to account for the standard of second-order desires either leads us to a dead end or away from the main point of the proposal. On the one hand, if the standard of the second-order desire is the satisfaction of some third-order desire, we face the question what the standard of that third-order desire is, and we are on our way to an infinite regress. On the other hand, if it is some other feature of the second-order desire, then it is presumably that other feature that grounds the standard of desire, and not the second-order desire itself. In principle, one could just accept the regress and settle with the idea that higher-order desires make a normative difference to lower-order desires. However, we should choose this option only if it turns out that there is no other, more fundamental, standard that can regulate desires across the board and from which DOR can be derived. In the next section, I will try to show that there is such a standard.

The third candidate for the constitutive standard of desire locates it in something else than desire-satisfaction. According to a view with a reputable ancestry, desires are for the good or represent their object/content as good (see Moss 2012; Stampe 1987; Oddie 2005; Tenenbaum 2007). The

so-called Guise of the Good view can be taken to imply that the constitutive standard of desire is the good and that desires that fail to be for it also fail to meet that standard. Take, for instance, Allan Hazlett who formulates Guise of the Good view in terms of correctness conditions:

GG It is correct to desire x if and only if x is good (and incorrect otherwise). (Hazlett 2019, 853)

Hazlett understands the goodness in question in terms of absolute goodness (Hazlett 2019, 857). If we understand correctness conditions as conditions for satisfying the constitutive standard, DOR will be those considerations that explain why the content/object of desire is good, i.e., why wanting it is correct.

GR (goodness reasons): p is a DOR for S to want x just in case p explains why x is good.

However, it is questionable if GR satisfies Effectiveness. The constitutive standard of desire should apply to it in virtue of the kind of psychological state that it is and the kinds of psychological processes it is implicated in. The processes of desire-regulation do not seem to be robustly sensitive to what appears as absolutely good. At least the idea that they are sensitive would require empirical support and can't be taken for granted. If it seems to an agent that p is a respect in which x is good in the absolute sense, this need not yet have any positive effect on their desire for x. It is possible that there are virtuous agents whose mechanisms of desire-regulation are robustly sensitive to

considerations regarding absolute goodness. However, I do not think that ordinary human animals are like that.

Furthermore, it is unclear if GR satisfies Accessibility. If absolute goodness is a real property, it is plausible to think that it is instantiated by many things and situations that have no relation to particular agents. For any S, there are numerous x which appear good in the absolute sense, but which do not bear any relation of relevance to S. It is implausible to think that, for all those x, an agent is in a position to see that it is fitting for her to have a desire for x. She might realize that having a desire for x is fitting for *someone* for whom x is relevant, but not that she herself should want x. In order for this kind of realization to be possible, we would need to idealize S to an extent that S loses any similarity with actual human beings with their subjective evaluative points of view. I thus take it that only agent-relative goodness has a chance of being a fitting candidate for reasons that could satisfy Accessibility. This also suggests that DOR themselves are agent-relative (for the distinction between agent-relative and agent-neutral reasons, see Nagel 1970).

What about relative goodness as the standard of desire? Hazlett considers it but rejects it because it delivers unintuitive predictions regarding the conditions in which a desire is correct (Hazlett 2019, 856). For instance, in a situation wherein I would take pleasure in causing some harm, my desire to cause harm would be correct because my pleasure is a good relative to me. This objection is not successful because it is possible to accept that some desires that are highly problematic for moral or other kinds of reasons are still correct in terms of their constitutive standard. In such cases, it makes sense to say that although the desire is correct, given its constitutive standard, one should not have it, all things considered.

That being said, I do not think that understanding the standard in terms of relative goodness passes muster. The main problem with relative goodness as the constitutive standard is that it doesn't

sufficiently specify what that goodness consists in. There are many ways in which x can be good for S: by being instrumentally good for some goal of S, by contributing to S's flourishing, by having use value, etc. (for an account of various good-for relations, see Rosati 2009) To identify the kind of relative goodness that grounds Effective reasons we need to consider how our desires are actually regulated. In this determinable form, the connection between relative goodness and desire-regulation remains unclear. It leaves open the possibility that some considerations that bear on some kinds of relative goodness are actually not Effective. There might be kinds of relative goodness that fail to engage desires in the right kind of way (i.e., by generating Effective reasons). Notice that the criticism of the goodness-based proposal, at least when goodness is understood in relative terms, differs from the concerns about others. I am not claiming that relative goodness cannot provide us with DOR, I am only saying that it is not sufficiently specified. One cannot avoid axiological questions when one wants to understand DOR in terms of relative goodness because it is not clear if all kinds of relative goodness can regulate desires in Effective way.

We have now considered three candidates for the standard of desire and found them all lacking in their capacity to provide DOR. Desire-satisfaction wasn't usable as a standard for evaluating which desires to have, satisfaction of a second-order desire didn't account for the desire-independent regulation of desire, and goodness-standard was not specific enough. In the next section, I am going to suggest another candidate for the standard and DOR that are derivable from it, which, as I will argue, make better sense of how our desire-regulating mechanisms actually operate.

4. Subjective satisfaction as the standard of desire

To see that there is a fourth candidate for the constitutive standard of desire, we have to notice a distinction between two senses of satisfaction. It is one thing for a desire to be satisfied, it is another for the subject of the desire to be satisfied (Goldie 2000: 25). That I get to eat the type of ice cream I wanted entails that my desire is satisfied. This does not yet mean that I am satisfied with that ice cream. Correspondingly, if I don't get to eat the ice cream that I want, this entails that my desire is frustrated, but this need not mean that I am frustrated or disappointed with my situation. I can have desires in whose satisfaction I don't find any fulfilment and I also can have desires whose frustration leaves me cold. Desire satisfaction (D-satisfaction) and subjective satisfaction (S-satisfaction) are both conceptually and psychologically distinguishable states but they both can have the same content as their target. I take it that it is a common experience that desires that we have often do not meet our expectations when they are satisfied, and we are left disappointed: there is D-satisfaction but no S-satisfaction. In other words, there is a phenomenal contrast between D-satisfaction that is accompanied by S-satisfaction and D-satisfaction that isn't. The distinction between a desire being satisfied and a subject being satisfied has also been noticed by Braun (2015) who distinguishes between desire satisfaction and agent satisfaction, de Sousa (1998) who has contrasted semantic satisfaction and emotional satisfaction, and Lycan (2012).

Intuitively, there seems to be something lacking in a desire whose satisfaction does not deliver any subjective satisfaction. We generally expect the satisfaction of our desires to be satisfying. If a person keeps on pursuing something, although the thing pursued continues to cause nothing but disappointment, it would have been better to give up the desire instead of clinging on to it. On the other hand, if a satisfaction of a desire is also subjectively satisfying, this suggests that having this desire was appropriate. Given these normative intuitions, I suggest S-satisfaction as another candidate for the standard of desire.

What does it mean for a subject to be satisfied in the sense of S-satisfaction? I think that a reductive analysis of it is neither needed nor even possible. Identifying S-satisfaction with pleasure, for instance, would not do. One can be subjectively satisfied with x when one is just relieved that x obtains, where the state of being relieved is not necessarily experienced as positively pleasurable. That S finds x pleasurable does not suffice for S to be subjectively satisfied with x either: surely, a chemical stimulation of the brain's hedonic hotspots is pleasurable, but the subject need not find it satisfying, especially when the stimulation overstays its welcome. Pleasures can be satisfying but they can also disappoint. In the context of the present paper, "S-satisfaction" is treated as a placeholder for what is missing in cases where a desire for x is D-satisfied but S is disappointed in x. There are interesting connections between S-satisfaction and some other psychological phenomena, though, that help us specify what the former amounts to.

First, although the relation between pleasure and S-satisfaction is not that of reduction, I take it that both are positively valenced states where valence is understood either as a primitive evaluative representation (Carruthers 2018) or as a reward marker (Prinz 2010). A positive valence of an experience signals that it is worth pursuing, thereby indicating that it was correct to have the desire. That S-satisfaction is a positively valenced state helps us understand how it could play a role in desire-regulation because positive valence of a state functions as a reinforcer.

Second, S-satisfaction can be fruitfully compared with the notion of life satisfaction which is an acknowledged measure in the study of well-being and happiness (Realo et al. 2017). Life satisfaction is about *feeling* satisfied with one's life, not merely just judging that one is satisfied (see Sumner 1996, 145; Haybron 2007). Likewise, S-satisfaction is comparable to life satisfaction in its intentional and phenomenological aspects, but it targets a specific content, not the agent's life in total. While an assessment of life satisfaction takes the whole life as its object, the judgment

about S-satisfaction is ‘localized’ to a particular content, x . Note that this does not constitute an analysis of S-satisfaction; it only points toward the possibility of modeling S-satisfaction on life satisfaction.

Insofar as well-being is understood in terms of life-satisfaction, then S-satisfaction with x could also be seen as a state of one’s well-being being positively affected by x . However, well-being is a contested concept which has been theorized about in various ways, some of which are quite distant from S-satisfaction. For instance, if well-being is taken to consist in the possession of a number of goods (e.g., friendship, achievement, pleasure, knowledge etc.), then S-satisfaction is not so closely comparable with it, although it might belong to the list of goods that well-being requires. Since I do not want to take a stand on which conception of well-being is correct, I also leave open a possibility that S-satisfaction could be reduced to (changes in) well-being or that S-satisfaction could be in service of well-being as the ultimate standard. Since S-satisfaction can be specified independently of well-being, these questions can be bracketed.

Third, I take it that S-satisfaction has its precedent in (positive) homeostatic signals. The latter are indicators of changes in homeostatic regulation, which aims at optimized organismic functioning (Damasio 2019, 83). Since normally functioning desires serve our homeostatic needs, it is also reasonable to assume that desire-regulation is intertwined with homeostatic regulation, at least in normal conditions. This might (partially) explain, for instance, why a mere stimulation of pleasure centers in the brain need not be S-satisfying: such stimulation presumably does not make a positive contribution to homeostasis. While the original understanding of homeostatic regulation takes it to concern keeping the values of certain internal variables within a limited range, where this kind of regulation is not consciously experienced, there are also conscious feelings of improvement or

decline of one's homeostasis (Damasio & Damasio 2016). S-satisfaction plays the role in desire-regulation that is analogous to the role of homeostatic feelings in homeostatic regulation.

Having S-satisfaction as the standard of desire, we can fill in the SCHEMA as follows:

SSR (S-satisfaction reasons): p is a DOR for S to want x just in case p explains why x would be S-satisfying.

For instance, that mushroom soup is flavorful by having dill in it, is warm, etc., is a DOR to want to eat it just in case it explains why eating mushroom soup would be subjectively satisfying. It is the respects in which a content is S-satisfying that constitute SSR to have a desire for it.

How do SSR fare with the desiderata for DOR? They do seem to satisfy Accessibility: if p explains why x is S-satisfying then it indicates why x is worth wanting, i.e., why x is desirable. Unlike a case in which x's obtaining causes only D-satisfaction, having a desire for x in this case looks like a fitting response to have. It is S-satisfaction that makes a normative difference to those two situations. Admittedly, making sense of Accessibility is not that remarkable achievement. Accessibility is basically a prophylactic constraint, meant to secure that DOR would remain within the limits of intelligibility and be practicable. But what about Effectiveness? Here, SSR enjoy an advantage over their rivals because human desire-regulation mechanisms function in a way that allows for SSR to guide them.

Consider the way in which desires are amenable to learning from experience. When an agent wants something, they have favorable expectations regarding the content of desire. If the expectation is not met and the satisfaction of the desire turns out to be disappointing, this tends to weaken the

desire. If the content of desire keeps being disappointing over those occasions, then this increases the likelihood of the agent's giving up the desire. On the other hand, confirmation of one's expectations of S-satisfaction in experiencing the content of desire has an effect of making the desire stronger. There can also be positive discrepancies, that is, when the experience exceeds the expectation, in which case an existing desire is strengthened or perhaps even a new desire is formed. These constraints are not watertight: there are recalcitrant desires in which case a person persists in having a desire despite the experience of deep disappointment when the desire is satisfied. There are also situations wherein a person fails to strengthen a desire for a state of affairs, despite of experiencing that state of affairs as deeply satisfying. In those conditions, however, desire-updating fails to function properly. Whether the satisfaction of desire turns out to be satisfying has a predictable effect on how one's desire is updated, at least in normal conditions. As Peter Railton puts it:

In supporting feed-forward action-guidance through expectation and reliance, and in supporting thereby a process of feedback from experience by assessing discrepancy with expectation, desire exhibits an inherent learning dynamic. The structure of this dynamic is essentially similar to belief: a sentiment toward *p* underwrites an expectation with respect to *p* that is compared with actual outcomes, and, when discrepancy is detected, the sentiment strengthens or weakens to reduce this discrepancy. (Railton 2017, 260)

One difference between my account and Railton's is that he takes wanting to involve expecting its content to be satisfactory *or* beneficial (2017, 262) while I don't see any need for this kind of disjunctive formulation. What is crucial in the present context, however, is that the way in which desires are updated is constrained by how their content appears in experience. Whether or not it appears S-satisfying is a crucial factor in determining how to update the desire. In virtue of this,

SSR as considerations that explain why a content would be S-satisfying are fit to guide this kind of desire-regulation. As S-satisfying aspects of the content of desire, they affect the strength of that desire.

Against this, an obvious complaint could be raised that, generally, when a desire is satisfied, it is not strengthened. Instead, it disappears. Here, we should keep in mind the distinction between standing and occurrent desires. It is true that if I want to eat ice cream and then eat it, the desire as an occurrent state usually fades upon satisfaction, independently of whether I feel S-satisfaction or not. But the desire to eat ice cream as a standing state is strengthened, which is indicated, among other things, by the fact that upon future occasions it is more likely that I form an occurrent desire to eat ice cream. S-satisfaction plays a regulative role in making this happen through considerations that explain why a desire would be S-satisfying.

Is this sufficient for SSR to be Effective, however? It might seem that if SSR guide desire-regulation through indicating what is S-satisfying about x, then they can influence the strength of desire only when the latter is satisfied. If that were the case, the occasions for regulating one's desire would be quite rare. It is not so, however, because desires do not have to be actually satisfied in order to be updated. Instead, a *simulation* of D-satisfaction provides a way of considering whether a desire is worth having before actually satisfying the desire.

Imagining or simulating D-satisfaction and considering whether it would be S-satisfying is enough to have an effect on desire strength. If I want something and imagine what it would be like to have it, depending on how the desire-satisfaction appears to me in imagination, my desire is either strengthened or weakened. This is because in imagining the content of desire I simulate the experience of satisfying the desire. In imagination I can already check if that experience would be satisfying or disappointing. People can play out what the fulfilment of desire would be like to

evaluate whether the desire is worth having. It is true that imaginings of future satisfaction can often be inaccurate, as the research on affective forecasting indicates (see Wilson and Gilbert 2005). However, this only suggests that desire-regulation through imagination is often error-prone, not that apparent subjective satisfaction does not guide it.

There is evidence that vivid imaginings of desire-satisfaction can modulate the strength of desire, indicated by increases in motivation, anticipated reward and anticipated pleasure, both in and outside the lab (Renner et al. 2019). The most extensive data in support of imaginings' effect on desire comes from research on pro-social motivation and food cravings. Regarding the first, Gaesser and colleagues have investigated how people are more motivated to help someone in need after episodically simulating the helping behavior. It turns out that vivid simulations of pro-social actions have a robust effect on the motivation to actually help others (Gaesser, Horn, & Young 2015; Gaesser & Schachter 2014). For instance, Gaesser, Shimura & Cikara investigated if episodic simulation of intergroup interaction has a positive effect on one's willingness to take prosocial actions (e.g., writing to a person in need, donating money). What they found was that willingness to help, indicated by self-reports, increases both toward in-group *and* out-group members (Gaesser, Shimura & Cikara 2020). This suggests that, insofar as prosocial motivation involves having a desire to help, vivid imaginings of helping someone directly modulate the strength of that desire. Importantly, positive affective valence of the simulation is an important factor in determining whether one's motivation is enhanced (see Gaesser, DiBiase, & Kensinger 2017), which suggests that one expects the realization of that motivation to be positively valenced as well.

Regarding food cravings, there is evidence indicating that simulations of consuming food have a robust influence on the strength of food cravings. Simulations of eating and its consequences are

triggered by food-related cues (Papies 2013; Papies et al. 2017, 406). These simulations in turn strengthen or weaken the respective desires, depending on how the imaginings are elaborated and how the imagined scenes appear to the subject: if they appear appealing, the strength of the craving increases (Tiggeman & Kemps 2005; Kemps & Tiggemann, 2015; Kemps, Tiggemann, Woods, & Soekov, 2004; Haasova et al. 2016; Keesman et al 2016). The same seems to apply to drinking simulations in response to alcohol related cues (Carter & Tiffany 1999).

It seems, then, that the same dynamic of congruency/incongruency-based learning can be mirrored at the level of simulated experience: desires are strengthened or weakened on the basis of whether their satisfaction appears satisfactory or disappointing. If a person who has a food craving, for instance, focuses in their imagination on the aspects of the object of the craving that they predict would satisfy them, the strength of the craving is increased; if they focus on the aspect that they consider as frustrating or disappointing, the strength of craving is decreased. That a desire for *x* is strengthened in response to apparent *S*-satisfaction both in actual experience and in imagining of *D*-satisfaction enables SSR to have an immediate effect on one's desires. Furthermore, there is a neural basis for this: brains compute fictive reward error prediction signals both in response to experienced *and* imagined outcomes (e.g., Lohrenz et al. 2007; Chiu et al. 2008).

SSR thus provide us with DOR that satisfy the desiderata and are preferable to the alternatives that we considered in the previous section. In contrast with DSR, SSR can provide positive normative guidance as to what desires to have. Unlike DDR, the present proposal does not face an infinite regress. Also, it allows that DOR can regulate desires independently of second-order desires, suggesting a more fundamental level at which desires can be responsive to reasons. Finally, unlike the Good-based proposal, the present account also articulates the standard and DOR in a way that makes explicit how they can be Effective and Accessible. As it was said in the previous section,

this need not be seen as a rejection of the GR, but as a specification of what the relative goodness or desirability that is fit to guide desires amounts to.

5. Concerns

In this section, I address a variety of challenges that the proposed account faces. The first two concerns are targeted at the idea that S-satisfaction is the constitutive standard of desire. The rest challenge SSR in particular.

5.1 One-off desires

The explanation of how S-satisfaction can be a regulative standard is grounded in the way in which our desire-regulating mechanism functions. In particular, it appeals to the idea that considerations bearing on S-satisfaction strengthen the desire and increase its probability of being tokened in the future. The explanation therefore assumed that the occurrences of desire are repeatable, otherwise we couldn't talk about the same desire being tokened at different times. However, there are desires that can be satisfied only once, such as someone's desire to climb the Mount Everest for the first time or to publish their first academic paper. If these desires are not repeatable, in what sense can S-satisfaction as a standard apply to them?

In response, one should note that although desire regulation through the experience of actually satisfying the desire does not apply to one-off desires, because there is no desire to be strengthened after it is satisfied, they are still regulated by updating through imagination. As I've argued above, desire-regulation can happen both through actual D-satisfaction and simulated D-satisfaction. Insofar as the same mechanism that aims at S-satisfaction governs both one-off and repeatable

desires, one-off desires also have S-satisfaction as their standard. This implies that a conative system which is capable of only having one-off desires would not have S-satisfaction as the standard of its desires. After all, that kind of system would presumably have no use for a congruence/incongruence-based learning mechanism. This implication is acceptable, however, given that human beings are not built like that.

5.2 On reducing S-satisfaction to D-satisfaction

The second concern is about the status of S-satisfaction as a distinct standard, different from D-satisfaction. I have explicitly left S-satisfaction unanalyzed. It could be argued, however, that S-satisfaction actually reduces to D-satisfaction (Braun 2015, 159). One might claim that if I have a desire for x, then x obtains, and I find myself disappointed with it, then this means that I also had another desire, for y, and, because by getting x I did not get y, I am disappointed with x. For instance, if an ice cream which I wanted does not satisfy me, this could mean that by wanting ice cream, I wanted to eat something delicious, but this other desire was not satisfied. If that is the case, S-satisfaction and its lack are understood in terms of D-satisfaction. The reductionist about S-satisfaction could also try to argue that S-satisfaction should be understood in terms of a more global desire to increase one's well-being in which case the satisfaction of a particular desire might be disappointing if that global desire is frustrated.

The first thing to say in response is that it is just implausible that the standard of desire can be understood in terms of D-satisfaction because congruence/incongruence-based desire-regulation is not governed by (second-order) desires. The strengthening and weakening of our desires through experiential and imaginative updating happens largely independently of whether we want to

strengthen or weaken them. Since the standard of desire derives from those updating processes, it is not derivable from D-satisfaction.

Second, if the apparent lack of S-satisfaction with the content of a desire can be traced back to the frustration of some other desire, this suggests that the former desire is instrumental with respect to the latter. If that is the case, the cases in question are not immediately relevant in the context of this paper which is concerned solely with intrinsic desires and their constitutive standard. Presumably, in the case of an intrinsic desire, an apparent lack of S-satisfaction is not explained by some other desire being frustrated.

Third, even if S-satisfaction can be reduced to D-satisfaction along the lines just envisaged, this does not imply that the present account is theoretically uninteresting. It is still informative to say that S-satisfaction is the constitutive standard of desire, even if S-satisfaction is eventually understood in terms of the satisfaction of some privileged set of desires. This result would still show that desires are constitutively embedded in a broader normative structure on the basis of which they can be evaluated.

5.3 Desire-satisfaction without S

SSR may seem awfully egocentric in that S-satisfaction requires the existence of a subject whose satisfaction it is. However, people can have desires upon whose satisfaction they wouldn't exist, desires for future generations, for instance. If we accept SSR as DOR, does it mean then that they have no DOR to have such desires? This is counterintuitive.

Biting the bullet in response this concern is not out of the question. Our desire-regulation mechanisms are arguably disposed to respond to feedback regarding our (potential) psychological

condition, not to conditions in which we do not exist. Our desires are at the service of our survival and considerations which do not promote survival are presumably merely extrinsic to desire and do not count as DOR. Also, not all considerations that make having a desire seem fitting have to be DOR. It can still be prudentially or morally reasonable to have desires whose satisfaction implies one's non-existence.

We do not have to fully concede to that objection, however. The proposed account is modally liberal. It allows that a consideration can be a DOR to want x when, if S were to exist and experience x , x would be S -satisfying. Understood in this way, the account can be extended to at least some desires upon whose satisfaction S doesn't exist anymore. If I were to exist and be aware of the well-being of future generations, for instance, this would arguably S -satisfy me, and that fact provides me with a DOR to want it. In virtue of desires being subject to updating through imagination, such considerations can also accessibly guide my desire-formation.

Where SSR do seem to lack application are desires whose satisfaction not only contingently implies my non-existence but also *requires* my not existing. For instance, I could have a desire that I be remembered after my death. In such a case, the possibility of imagining S -satisfaction if I were to exist seems to be precluded. After all, if I were to exist, the desire to be remembered after my death would not be satisfied. One could argue in response that subject could imagine from the perspective of a ghostlike entity or a Cartesian self of how other people reminisce about them. If they imagined being S -satisfied with this, then we could say that they have a DOR to want to be remembered. However, leaving aside the issue that the person would then still exist in that scenario in some elusive form, the problem with this idea is that we would have to accept the existence of facts about what S -satisfies ghostlike entities or Cartesian selves, and this is difficult to stomach. It is thus doubtful if SSR are applicable to such cases. But instead of taking this limitation to be a

reason to reject SSR, I think that we should treat it as its interesting consequence, in line with the idea that desires have a distinctive normative profile that can divide up different kinds of reasons in an unexpected way. It turns out that in the case of desires which are satisfied when one does not exist, it is easier to have DOR for other-regarding desires than for self-regarding desires.

5.4 Desire-satisfaction without awareness

Aside from desires whose satisfaction implies S's non-existence, there are also desires whose satisfaction implies that S is not aware of it. Take a desire that my friend do something without my awareness. I could have such a desire in a situation, for instance, wherein the awareness of that deed would cause me grave discomfort. Yet, since I care about my friend and know that the deed in question would please them, I still want them to do it, just without my awareness. Such desires are surely possible and yet pose a challenge to the present account of DOR: because their D-satisfaction implies that I am not aware of it, it also implies that I can't get any S-satisfaction from it either (I assume here that one cannot be S-satisfied with x if one is not aware of x). But in that case, it seems that there cannot be SSR for one to have such desires.

How should we treat those cases of D-satisfaction without awareness? The present account has resources to provide such desires, or at least some of them, with DOR. In order to see this, note that we do not need to assume that S on the left-hand side of SSR is the same as S on the right-hand side. Perhaps also the S-satisfaction of *other* subjects can ground reasons for *our* desires? Somewhat speculatively, it can be argued that there are empathetic desires regarding other people in which case a DOR for one person to have such a desire is a consideration that explains why x would be S-satisfactory for some other person.

As it stands, this response needs to be fleshed out more because if S on the right-hand side can refer to whoever, the account over-generates the DOR that one has. It can't be the case that one has DOR for all desires whose satisfaction is S-satisfying for *some* other person, whoever that other person is. In order to constrain the set of relevant subjects of DOR-relevant S-satisfaction, we should consider how one person's desire-regulation can become robustly sensitive to what appears to S-satisfy some other person. Let's call the first person S_1 and the other S_2 . S_1 's desires can become sensitive to S_2 's subjective satisfaction when the two are sufficiently exposed to each other, so that the modulation of S_1 's desire strength becomes coupled with S_2 's expressive behavior. This can presumably happen in the contexts of close relationships of attachment, between spouses or other family members, for instance. And if this happens, then we can apply to S_1 's desires the same reasoning that we applied in intrasubjective cases: insofar as S_1 's desire regulation is geared toward S_2 's subjective satisfaction, considerations that bear on the latter count as DOR for S_1 . Note that this explanation would also allow us to attribute DOR to a person's desire to be remembered after death, insofar as those who remember are sufficiently close to that person.

As already admitted, this sketch of an explanation of how other people's S-satisfaction can be relevant for our DOR is speculative. But here the aim was just to indicate how the present account can make sense of DOR for desires whose satisfaction implies a lack of awareness on the part of the subject. Interestingly, the suggested explanation is only applicable to cases in which there is still some other subject who is capable of S-satisfaction. It doesn't explain how one could have DOR for a desire whose satisfaction implies that no subject is aware of it. Someone's nihilistic desire for the end of the world might be an example of such a desire. However, that there can be no DOR for the latter is intuitively an attractive limitation.

5.5 Irrelevance of SSR

One might also raise a concern that the proposed account provides DOR with insufficient weight for them to ever really matter for the evaluation of a person's desires. Arguably, there are kinds of considerations that count for or against having a desire that are always weightier than SSR, such as moral or prudential reasons. In addition, there are also hedonic reasons for having a desire which can compete with SSR and often outweigh them. This is basically a practical objection to SSR: even if there are conative reasons that bear on whether satisfaction of desire would be subjectively satisfactory, these reasons are less important than some extrinsic considerations, insofar as the latter exist. If this objection goes through, then the account of DOR that I've proposed might be correct but is for all intents and purposes uninteresting and normatively insignificant. Admittedly, this objection can be raised with equal force against other accounts of DOR as well, with a possible exception of GR.

I am not entirely sure how to respond to this concern, largely because it trades on intuitions about normative weight which can be untrustworthy. That being said, even if it is true that certain non-SSR considerations always have more weight than SSR, the latter are still relevant in at least two kinds of situations. First, presumably there are cases in which none of the non-SSR considerations that are weightier than SSR matter for what to want. For instance, one could be in a circumstance in which moral and prudential reasons are simply not relevant for a particular desire in question. If that is the case, SSR to have that desire can be decisive. Second, SSR can also play a decisive role in contexts where weightier non-SSR considerations do not settle the question of whether to have a desire for x or not. This might happen, for instance, when non-SSR reasons for and non-SSR reasons against are equally strong. SSR can play a tie-breaking role in such situations. The

practical relevance of SSR is therefore not cancelled by the putative fact that there are stronger reasons out there.

Finally, a case can be made that there is something lacking in a person who has non-DOR reasons but no SSR for some of their desires and that SSR therefore always have relevance for the evaluation of one's conative perspective. If a desire does not have SSR, then it fails to function in ways that are conducive to S-satisfaction, and there is a sense in which one's conation is somewhat poorly organized in such a situation. This becomes noticeable, for instance, when we think of cases in which one's desire has a hedonic reason but no SSR. Take, for instance, insatiable desires whose satisfaction offers pleasure but no S-satisfaction. Although such desires have hedonic reasons, they do not have DOR, and the latter fact explains why there seems to be something problematic about having such desires, although the latter are reasonable in the sense of having a hedonic reason. The relevance of SSR for the evaluation of persons' desires is thus not screened off by other, non-DOR reasons.

6. Conclusion

In this paper, I claimed that subjective satisfaction governs our mechanisms of desire-regulation in virtue of the latter being responsive to considerations that explain why the satisfaction of the desire would be subjectively satisfying. These considerations are desire's own reasons.

That subjective satisfaction is what our desires are oriented toward might make desires appear to be myopic creatures, ultimately just aiming at the fulfilment of agents' self-centered interests and perhaps also interests of those who are in their close circles. It also suggests that people can have DOR for rather reprehensible desires, insofar as their satisfaction would satisfy them. However,

the standard of subjective satisfaction is the most we can reasonably expect from non-ideal agents like us. Also, as humble as it is, the standard still upholds certain forms of evaluation and criticism, targeted at desires in particular. The present account also leaves open the possibility that some kinds of subjective satisfaction are better or more fitting than others. For instance, subjective satisfaction that fades quickly is presumably of lesser merit than satisfaction that is more stable or even grows over time.

References

Adler, J. E. (2006) *Belief's Own Ethics*. Cambridge, MA: MIT Press.

Arpaly, N., & Schroeder, T. (2014) *In Praise of Desire*. Oxford: Oxford University Press.

Braun, D. (2015) 'Desiring, desires, and desire ascriptions'. *Philosophical Studies*, 172, 141–62.

Carruthers, P. (2018) 'Valence and value'. *Philosophy and Phenomenological Research*, 97, 658–80.

Carter, B. L., and S. T. Tiffany. (1999) 'Meta-analysis of cue-reactivity in addiction research'. *Addiction*. 94, 327–40.

Chiu, P.H., Lohrenz, T.M. and Montague, P.R. (2008) 'Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task'. *Nature Neuroscience*, 11, 514–20.

Côté-Bouchard, C. (2016) 'Can the aim of belief ground epistemic normativity?'. *Philosophical Studies*, 173, 3181–98.

D'Arms, J. and Jacobson, D. (2000) 'Sentiment and value'. *Ethics*, 100: 722–48.

- D'Arms, J. and Jacobson, D. (2014) 'Wrong kinds of reason and the opacity of normative force'. *Oxford studies in metaethics*, 9, 215–42.
- Damasio, A. (2019) *The Strange Order of Things: Life, Feeling, and the Making of Cultures*. Vintage.
- Damasio, A. and Damasio, H. (2016) 'Exploring the concept of homeostasis and considering its implications for economics'. *Journal of Economic Behavior & Organization*, 126, 125–29
- De Sousa, R. (1998) 'Desire and serendipity'. *Midwest Studies in Philosophy*, 22, 120–34.
- Enoch, D. (2006) 'Agency, shmagency: Why normativity won't come from what is constitutive of action'. *The Philosophical Review*, 115, 169–98.
- Gaesser, B., Shimura, Y. and Cikara, M. (2020) 'Episodic simulation reduces intergroup bias in prosocial intentions and behavior'. *Journal of Personality and Social Psychology*, 118, 683–705.
- Gaesser, B., DiBiase, H. D., & Kensinger, E. A. (2017) 'A role for affect in the link between episodic simulation and pro-sociality'. *Memory*, 25, 1052–62.
- Gaesser, B., Horn, M. and Young, L. (2015) 'When can imagining the self increase willingness to help others? Investigating whether the self-referential nature of episodic simulation fosters pro-sociality'. *Social Cognition*, 33, 562–84.
- Gaesser, B., Keeler, K., & Young, L. (2018) 'Moral imagination: Facilitating prosocial decision-making through scene imagery and theory of mind'. *Cognition*, 171, 180–93.
- Gaesser, B., & Schacter, D. L. (2014) 'Episodic simulation and episodic memory can increase intentions to help others'. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 4415–20.

- Goldie, P. (2000) *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.
- Goldman, A.H. (2017) 'What desires are, and are not'. *Philosophical Studies*, 174, 333–52.
- Haasova, S., Elekes, B., Missbach, B. and Florack, A. (2016) 'Effects of imagined consumption and simulated eating movements on food intake: thoughts about food are not always of advantage'. *Frontiers in Psychology*, 7, 1–8.
- Hawkins, J. (2007) 'Desiring the bad under the guise of the good'. *The Philosophical Quarterly*, 58, 244–64.
- Haybron, D. 2007. 'Life satisfaction, ethical reflection, and the science of happiness'. *Journal of Happiness Studies*, 8, 99–138.
- Hazlett, A. (2019) 'The guise of the good and the problem of partiality'. *Canadian Journal of Philosophy*, 49, 851–72.
- Keesman, M., Aarts, H., Vermeent, S., Häfner, M. and Papies, E.K. (2016) 'Consumption simulations induce salivation to food cues'. *PloS one*, 11(11): e0165449.
- Kelly, T. (2002) 'The Rationality of Belief and Some Other Propositional Attitudes'. *Philosophical Studies*, 110, 163–96.
- Kemps, E., & Tiggemann, M. (2015) 'A role for mental imagery in the experience and reduction of food cravings'. *Frontiers in Psychiatry*, 5, 193, 1–4.
- Kemps, E., Tiggemann, M., Woods, D. and Soekov, B. (2004) 'Reduction of food cravings through concurrent visuospatial processing'. *International Journal of Eating Disorders*, 36, 31–40.

- Lohrenz, T., McCabe, K., Camerer, C.F. and Montague, P.R. (2007) 'Neural signature of fictive/counterfactual learning signals in a sequential investment task'. *Proceedings of the National Academy of Sciences*, 104, 9493–8.
- Lycan, W. G. (2012) 'Desire considered as a propositional attitude'. *Philosophical Perspectives*, 26, 201–15.
- Moss, J. (2012) *Aristotle on the Apparent Good: Perception, Phantasia, Thought, and Desire*. Oxford: Oxford University Press.
- Nagel, T. (1970). *The Possibility of Altruism*. Oxford: Oxford University Press.
- Oddie, G. (2005) *Value, Reality, and Desire*. New York: Oxford University Press.
- Papies, E. K. (2013) 'Tempting food words activate eating simulations'. *Frontiers in Psychology*, 4, 838, 1–12.
- Papies, E.K., Best, M., Gelibter, E. and Barsalou, L.W. (2017) 'The role of simulations in consumer experiences and behavior: Insights from the grounded cognition theory of desire'. *Journal of the Association for Consumer Research*, 2, 402–18.
- Papineau, D. (1998) 'Teleosemantics and indeterminacy'. *Australasian Journal of Philosophy*, 76, 1–14.
- Papineau, D. (1993) *Philosophical Naturalism*. Oxford: Blackwell.
- Prinz, J. (2010) 'For valence'. *Emotion Review*, 2, 5–13.
- Rabinowicz, W, and Rønnow-Rasmussen, T. (2004) 'The strike of the Demon: On fitting pro-attitudes and value'. *Ethics*, 104, 391– 423.

- Rabinowicz, W. and Rønnow-Rasmussen, T. (2006) 'Buck-passing and the right kind of reasons', *The Philosophical Quarterly*, 56, 114–20.
- Railton, P. (2017) 'Learning as an inherent dynamic of belief and desire'. In F. Lauria & J. Deonna (eds.), *The Nature of Desire* (New York: Oxford University Press), pp. 249–76.
- Realo, A., Johannson, J. and Schmidt, M. (2017) 'Subjective well-being and self-reported health in osteoarthritis patients before and after arthroplasty'. *Journal of Happiness Studies*, 18, 1191–206.
- Renner, F., Murphy, F.C., Ji, J.L., Manly, T. and Holmes, E.A. (2019) 'Mental imagery as a "motivational amplifier" to promote activities'. *Behaviour Research and Therapy*, 114, 51–9.
- Rosati, C. S. (2009) 'Relational good and the multiplicity problem'. *Philosophical Issues*, 19, 205–34.
- Schroeder, M. (2012) 'The ubiquity of state-given reasons'. *Ethics*, 122(2), 457–88.
- Schroeder, T. (2004) *Three Faces of Desire*. Oxford: Oxford University Press.
- Shah, N. (2006) 'A new argument for evidentialism'. *The Philosophical Quarterly*, 56, 481–98.
- Sharadin, N. (2015) 'Reasons wrong and right'. *Pacific Philosophical Quarterly*, 97, 371–99.
- Stampe, D. W. (1987) 'The authority of desire'. *Philosophical Review*, 96, 335–81.
- Sumner, L. W. (1996) *Welfare, Happiness, and Ethics*. Clarendon Press.
- Sylvan, K. (2015) 'What apparent reasons appear to be'. *Philosophical Studies*, 172, 587–606.
- Tenenbaum, S. (2007) *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge: Cambridge University Press.

Tiggeman, M., and Kemps, E. (2005) 'The phenomenology of food cravings: The role of mental imagery'. *Appetite*, 45, 305–13.

Wilson, T. D. and Gilbert, D. T. (2005) 'Affective forecasting: Knowing what to want'. *Current Directions in Psychological Science*, 14, 131–34.