Noûs

NOÛS 53:3 (2019) 664–688
doi: 10.1111/nous.12232

Which Desires Are Relevant to Well-Being?

CHRIS HEATHWOOD
University of Colorado Boulder

Abstract

The desire-satisfaction theory of well-being says, in its simplest form, that a person's level of welfare is determined by the extent to which their desires are satisfied. A question faced by anyone attracted to such a view is, *Which desires?* This paper proposes a new answer to this question by characterizing a distinction among desires that isn't much discussed in the well-being literature. This is the distinction between what a person wants in a merely behavioral sense, in that the person is, for some reason or other, disposed to act so as to try to get it, and what a person wants in a more robust sense, the sense of being *genuinely attracted* to the thing. I try to make this distinction more clear, and I argue for its axiological relevance by putting it to work in solving four problem cases for desire satisfactionism. The theory defended holds that only desires in the latter, genuine-attraction sense are relevant to welfare.

The desire-satisfaction theory of well-being says, in its simplest form, that a person's level of welfare is determined by the extent to which their desires are satisfied. A question faced by anyone attracted to such a view—or any view in which desire plays a role[1]—is, *Which desires?* Is satisfying just any desire beneficial in itself, or might it be that only one's *idealized* desires are directly relevant to well-being, or perhaps only one's *global* desires, or one's *self-regarding* desires, or one's *second-order* desires, or one's *non-moral* desires, or one's *non-altruistic* desires, or one's *now-for-now* desires, or one's *autonomous* desires? For each of these kinds of desire, some philosophers have suggested restricting desire satisfactionism to count just that kind.[2]

When it comes to which desires are relevant to well-being, I tend towards the inclusive. I am inclined to reject each of the restrictions just mentioned, as well as any kind of idealization.[3] I want to count a person's *actual* rather than idealized desires, and I want to count *local* as well as global desires, *other-regarding* as well as self-regarding, *first-order* as well as second-order, *moral* as well as non-moral, *altruistic* as well as non-altruistic, *now-for-then* as well as now-for-now, and *non-autonomous* as well as autonomous. But there is a distinction among desires—one less discussed in the well-being literature—that I believe to be fundamentally axiologically relevant. This is the distinction between what a person wants in a behavioral sense, in that the person is, for some reason or other, disposed to act so as to try to get it, and what a person wants in a more robust sense, the sense of

© 2017 Wiley Periodicals, Inc.

being *genuinely attracted* to the thing, or of the thing's being *genuinely appealing* to the person. In this paper, I try to make this distinction more clear, and I argue for its axiological relevance by putting it to work in solving four problem cases for the desire theory of welfare.

I begin by describing the problem cases. They pose plausible counterexamples to a simple unrestricted desire theory: such a theory implies that a benefit is taking place when in fact that's not the case, or that there is more benefit than there intuitively is, or that the benefit is occurring at the wrong time, or some such. Along the way, I sometimes mention some of the ways that some other desire theorists have suggested we handle these cases. This is usually in the form of some kind of restriction to the desire theory. In most cases, I hint at why I don't think the mentioned suggestion will work, but, unfortunately, it is beyond the scope of the present paper to fully explain why I believe alternative solutions to be unsuccessful.

Next, I introduce a general sort of distinction among kinds of desire—or, more accurately, senses of the term 'desire'.[4] This is a distinction that appears more in other literature than it does in the well-being literature. Partly to forestall the suspicion that the distinction I am appealing to is an *ad hoc* invention brought in to solve the problem cases, I describe what others have said about it. I draw out what seem to be some consensus ideas concerning the distinction. I don't pretend to have given strict definitions of the two senses of 'desire', but I hope to have conveyed the distinction well enough.

Finally, I explain how drawing this distinction among senses of 'desire', and formulating desire satisfactionism in terms of one rather than the other of the senses, enables the theory to handle the problem cases. This is evidence that the best desire theory will be formulated in the way I am proposing. This theory puts forth a new answer to the question of which desires are relevant to well-being.

1. Four Problem Cases

a. Idealistic Desires

One kind of case we'll be concerned with is that of *moral* desires, or, more generally, desires pertaining to certain *ideals*. Following Robert Adams, who raises this as a problem for desire theories of welfare, we can call these 'idealistic desires' (not to be confused with ideal*ized* desires). According to Adams,

> Altruistic desires might lead you to sacrifice your own good for the good of another. This seems to imply that what you would prefer, on the whole, with full knowledge, is not necessarily what is best, on the whole, for you. . . . Something like [this] problem [also] arises in connection with desires that are not necessarily altruistic but may be called "idealistic." One may clearheadedly do what is worse for oneself out of regard for virtue, or for some other ideal. Love of truthfulness, or of human dignity, may lead a person to tell the truth, or to refuse to abase herself, at great cost to herself and for nobody else's benefit. (Adams 1999: 87–88)

I agree with Adams that this is an important problem for the desire theory. Here, first, is a "warm-up case," a simple illustration of how which outcome we prefer can come apart from which outcome would be best for us:

> *Trolley Variation*: A runaway trolley is heading towards five workers on the track, about to kill them. You can pull a lever, which would divert the trolley onto a spur of track, thereby saving these five. Unfortunately, there is a person on this spur of track who will die if you do this. Even more unfortunately, this person is you. After some quick reflection, you are impressed by the fact that each of the five on the other track is a person just like you, who matters just as much as you. Why should you get to live at the expense of *all five* of them dying? On the tail of these thoughts, you grit your teeth and, heroically, flip the switch.[5]

I'm not bringing this case up to discuss whether this act is obligatory or praiseworthy or rational.[6] I want us to consider *benefit*. And I say that surely you did not do what would most benefit you when you turned the trolley onto yourself. You would have fared better had you not pulled the lever.[7] However, this was the outcome you most preferred. You were free to bring about whichever outcome you wanted. You voluntarily brought about the outcome in which you die; so this was the outcome you wanted most. The desire theory may thus seem to imply, absurdly, that you acted in your best interest in this case.

I describe this as a warm-up case because I think that in fact the desire satisfactionist can reject the understanding of her view on which the outcome that is most in a person's interest is the one they now most prefer. She can instead hold that it is the outcome that contains, for them, the greatest balance of desire satisfaction over frustration. The outcome you most preferred in Trolley Variation is not like that, since you don't even exist in the outcome you most prefer, and it's hard to accrue desire satisfactions when you don't exist. We can call this reply to the warm-up case 'the appeal to future desires' (cf. Heathwood 2011: §3; Crisp 2017: §4.2).

To make real trouble for desire theories, we need a case in which the subject continues to exist and continues to prefer the intuitively worse option for herself. For then the appeal to future desires won't help. So consider a case that I'll call *Pig Slice.* The term 'pig slice' is a "sniglet," which is "any word that doesn't appear in the dictionary, but should" (Hall 1984). I no longer have any *Sniglets* books, but the Urban Dictionary defines 'pig slice' as "the last unclaimed piece of pizza that everyone is secretly dying for."[8]

> *Pig Slice*: You and your good friend are down to the last slice of pizza, and are both secretly dying for it. In the face of how much you want the pig slice, you manage a white lie: "I couldn't eat another thing; you take the last slice." Your craving for the pig slice lasts throughout the evening, but still you wouldn't change your choice if you could—you continue to prefer that your friend got the pig slice—since you regard it as the right and selfless thing to do.[9]

Surely yours was a selfless act. But the desire theory struggles to agree, since you did what you most wanted to do, and what you continued to most want that you did. It thus implies that in fact you acted most in your own interest, which seems like the intuitively wrong result. (We'll stipulate that there are no ill health effects

or any other bad side effects of eating the extra slice of pizza.) The appeal to future desires does not help here, since, unlike in Trolley Variation, your decision to forgo the pig slice does not deprive you of other desire satisfactions. In the future, you'll still form myriad desires that you will satisfy—e.g., to go for a walk, to see a movie, to read a book. Not so in Trolley Variation, which was why, desire theorists can say, your choice in that case was so harmful to you.

One natural and common solution to these kinds of cases is simply to exclude idealistic desires from the theory—to claim that satisfying them is of no intrinsic benefit. Thus, in his seminal discussion of desire theories of welfare, Sidgwick sets down that he will consider "only what a man desires ... for himself—not benevolently for others" (Sidgwick 1907: 109). More recently, Richard Brandt mentions excluding altruistic desires, and Harriet Baber excludes preferences that we have "in response to feelings of moral obligation" (Brandt 1979: 247; Baber 2007: 107).[10] She acknowledges, however, that "Sorting out moral from nonmoral motivations may be problematic" (Baber 2007: 4n).[11]

For the desire theory to exclude idealistic desires, it indeed needs to be able to say what they are. One natural way to do this is by the *content* of the desire, another by its *genesis*. To illustrate the former, we might say that altruistic desires are those that are about the interests of others, or are simply about others. It is indeed a common strategy among desire theorists to restrict the theory to count only self-regarding desires, or desires that are about oneself or one's own life.[12] But this restriction excludes too much. If it is my heart's desire that the team I root for win, then surely it is a good thing for me when the team wins, even though this desire isn't about me or my life. Notice that a hedonistic theory would never exclude from one's own good the *pleasure* taken in the victory of the team one roots for. Nor should desire theories exclude the corresponding desire satisfaction.[13]

As for the suggestion of genesis, we might say that idealistic desires are those caused by or based in the person's ideals. So your preference not to have the pig slice might be caused by your belief that selflessness is a virtue. It is not plausible to exclude all such desires, however. It will be easier to explain why after I present my own positive proposal. I will suggest instead that, to solve the problem, we should look to the *way* in which the objects of idealistic desires are typically desired rather than their content or genesis.

b. Compulsive Desires

Another problem case for desire theories of well-being involves compulsive desires that intuitively seem to be of no benefit to the compulsive person. One such case is due to Derek Parfit:

> Knowing that you accept a Summative [desire-fulfillment] theory, I tell you that I am about to make your life go better. I shall inject you with an addictive drug. From now on, you will wake each morning with an extremely strong desire to have another injection of this drug. Having this desire will be in itself neither pleasant nor painful, but if the desire is not fulfilled within an hour it would then become extremely painful. This is no cause for concern, since I shall give you ample supplies of this drug. Every morning, you will be able at once to fulfil

> this desire. The injection, and its after-effects, would also be neither pleasant nor painful. You will spend the rest of your days as you do now. (Parfit 1984: 496)

Parfit then writes, "if I make you an addict, I would be increasing the sum-total of your desire-fulfilment," and so if the desire-fulfillment theory of welfare is true, "I would be benefitting you." "This conclusion," he claims, "is not plausible" (Parfit 1984: 496).

Another well-known case of compulsion is Warren Quinn's *Radio Man*. "Suppose," Quinn writes, "I am in a strange functional state that disposes me to turn on radios that I see to be turned off" (Quinn 1993: 32). Quinn says that he "cannot see how this bizarre functional state," which may be hard to deny is a desire, "in itself gives me even a *prima facie* reason to turn on radios, even those I can see to be available for cost-free on-turning" (Quinn 1993: 32). Quinn's target is a Humean or desire-based theory of normative reasons rather than a desire-based theory of well-being. But if Quinn is right that this functional state gives him no reason to turn on radios, then, since welfare is reason-providing, turning on radios would give him no benefit either.

Parfit's suggested solution is that the desire theorist appeal to a person's *global* desires rather than the sum of their day-to-day desires to explain the benefit to them of some outcome (Parfit 1984: 495–96). "A preference is global," Parfit explains, "if it is about some part of one's life considered as a whole, or is about one's whole life" (496). Because a person "might globally prefer one of two possible lives even though it involved a smaller total sum of local desire fulfilment," the global version of the desire-fulfillment theory may not imply that one is benefitted if Parfit makes one in this way addicted (135).

Note that the earlier appeal to future desires that I mentioned as a way of avoiding the objection from Trolley Variation seems to walk one right into Parfit's objection. The appeal to future desires is essentially an appeal to the summative desire-fulfillment theory. But also note that Parfit's appeal to global desires seems to walk one right back into the problem posed by Trolley Variation. We can thus see these two cases—Trolley Variation and Parfitian Addiction—as a Scylla and Charybdis for desire theories: avoid one, and the other threatens. I believe that the general solution that I will propose handles both cases.

c. Prudential Desires

The appearance of this category may be surprising. Will the suggestion be that prudential desires—that is, desires for or concerning one's own self-interest—are problematic for a desire theory of prudential value? Wouldn't these be just the desires with which such a theory is most at home?

I think we'll see that these cases can be similar to cases of idealistic desires. Here is such a case:

> *The Smoker*: A smoker wants to quit, and finally resolves to do so. For the next month, he has intense cravings for cigarettes almost constantly. But he knows that cigarettes cause cancer. So it's also true that all day, each day, he wants himself not to have a cigarette. In the end, he is able to resist his desires for

> cigarettes; his desire not to smoke wins out. Eventually, he loses his desire to smoke, and for that's he is grateful.

So what's the problem? Since not to smoke is what he most wanted that month, the desire theory would seem to imply that what he did—not smoke—was most in his interest. And this is intuitively the right answer. It *is* beneficial not to smoke.

The problem has to do with the timing of the benefit. The man, we can suppose, would himself say that this was a very difficult month, one that he should never want to repeat. This suggests that it was a bad month for him. But it was a month during which he, at each moment, got the thing he most wanted, which was not to smoke. So a desire theory might seem to imply, incorrectly, that this was a good month for him, rather than a difficult one to be endured for the sake of a later good.

The point about timing can also be made with a variant of the case. Suppose that at the end of his difficult month of quitting, the man dies suddenly of a brain aneurysm. Suppose that this is unrelated to his smoking or to his quitting: it was going to happen either way at the end of this month. Given that he had only a month to live, and given that smoking, if it would end up being harmful to him at all, would be harmful to him only decades later, we should say that the man was *worse off* by doing what he most wanted to do that month, which was not to smoke. This is ostensibly in conflict with the desire theory of well-being. The man did what, at each moment, he most wanted to do, but seems to have gotten a bad month.

One familiar desire-theoretic solution to this case is to idealize for full information. If, in the second, variant case, the man knew he was going to die either way, perhaps he would have had no desire not to smoke, and so he indeed would have been worse off, on the idealized desire theory, if he had somehow been unable to smoke. But ideal desire theories bring with them new problems, problems which, in my view, make them untenable. Unfortunately, I can't defend this claim here.[14] In any case, I want to pursue another solution, the same one as for the earlier cases. I am going to suggest that we exclude the desire to quit, not because it might be based on false information, but because of something more intrinsic to it.

Another possible solution is to appeal to the intrinsic-desire/instrumental-desire distinction. According to this solution, only the satisfaction of intrinsic desire benefits us, and the smoker's desire not to smoke is merely instrumental.[15] I myself have been tempted by this solution, but I have started to doubt whether the restriction to intrinsic desires is viable. Though I'm unable to give a full defense of this doubt in this paper, here is a glimpse at the sort of consideration that worries me. Suppose a father very strongly wants to see A's on his son's report card. The report card arrives and indeed the son has earned straight A's. Plausibly, this is a good thing for the father and it is in the spirit of the desire-satisfaction theory to agree. But, for all that, such a desire on the father's part might be merely instrumental, even if it would take the father some thinking to figure out just why it's important to him to see A's on his son's report card. This suggests that satisfying at least some instrumental desires may be intrinsically good for people.

But certainly there are also cases in which the satisfaction of a merely instrumental desire is intuitively of no benefit to the desirer, and that of course is what

motivates the restriction to intrinsic desires.[16] I believe that the distinction that I will lay out and defend the relevance of *correlates* with the intrinsic desire/instrumental desire distinction, and that when they come apart, my distinction gets things right. Thus, I believe my distinction can play the role that the restriction to intrinsic desire has been hoped to play.

Since I cannot here fully defend these claims either, I emphasize that the restriction to intrinsic desires would not help in any of the other three problem cases under consideration. Idealistic desires, compulsive desires, and desires concerning unlikely possibilities (to be introduced presently) can all be intrinsic.

d. Desires Concerning Unlikely Possibilities

Fred Feldman presents an interesting argument against Wayne Davis' (1981) desire-based theory of happiness. But the case Feldman describes makes trouble for desire-based theories of welfare as well:

> Suppose Lois is emotionally neutral—neither happy nor unhappy. Suppose she is taking some children through a museum where they see a dinosaur exhibit. Lois is looking at the skeleton of an apparently ferocious dinosaur in the museum. She hears some other visitors talking. One remarks on how horrible it would be to be eaten by one of those things. Lois thinks about how horrible it would be to be eaten by a dinosaur. Of course she wants not to be eaten by a dinosaur. At the same time, she knows that dinosaurs are extinct and have not eaten anyone in hundreds of years, so she is quite confident that she will not be eaten by a dinosaur. Davis's theory implies that this should constitute an increase in her level of happiness, but it doesn't. Her neutral emotional state persists. She gains no joy from the realization that she is not going to be eaten by a dinosaur. She never felt that she was in danger of being eaten by a dinosaur, and so reflection on the fact that it is not going to happen does not bring any relief. (Feldman 2010: 66)

Though Feldman is interested in drawing conclusions about the nature of happiness, I want to consider the case as in support of the claim that Lois does not *benefit* when she comes to want not to be eaten by dinosaur. A standard desire-based theory would have to say that Lois does benefit when her desire not to be eaten by this dinosaur is satisfied.

As with the previous case, some people might be inclined to say, What's the problem? It *is* a good thing for her that this desire will be satisfied. When I say that the desire theory faces a problem here, I am not saying that the problem is that the theory implies that it would be better for Lois not to be eaten by the dinosaur when in fact this would not be better for her. Of course it would be better for her not to be eaten by a dinosaur. The true claim that it would be better for her not to be so eaten involves a comparison between the outcome in which she is eaten and the outcome in which she is not. To get the objection that I am interested in going, we need to make a comparison between two possibilities in neither of which is she eaten by a dinosaur:

O1: a quieter day at the museum, in which she never overhears any visitors talking about being eaten by a dinosaur, and so never considers the question of being

eaten by a dinosaur, and so never comes to have any desires concerning that possibility;

O2: the day as described by Feldman in which Lois overhears the other visitors say that they wouldn't want to be eaten by a dinosaur, and in which Lois therefore thinks to herself, "I don't want that either."

O1 and O2 are otherwise exactly similar outcomes for Lois. The putative problem for the desire theory of welfare is this: O2 contains an extra desire satisfaction for Lois, but, intuitively, Lois fares no better in O2 than in O1.

It's worth noting that it's easy to explain on a desire view why it would be worse for Lois to be eaten by a dinosaur even if she has never considered the possibility and so has never formed any desires about it (in other words, why O1 is a better outcome for her than a fantastical third outcome, O3, in which the dinosaur comes to life and eats her). For if suddenly she were to start to be eaten by a dinosaur, she would very intensely want it not to be happening. Moreover, and much more significantly, since being eaten by a dinosaur would kill her, she would also lose out on an enormous amount of desire satisfaction in the future that she would have received had she not been eaten by the dinosaur.[17] Thus, we don't need it to be true that her desire satisfaction in the original example is itself a benefit to explain why it would be bad for her on a desire theory to be eaten by a dinosaur.

To strengthen the case against some versions of preferentism, consider that Lois' desire in Feldman's example, though utterly lacking in emotion, will be *very strong* on at least one way that desire theories might determine the strength of a desire. Of all the things that Lois comes to desire during her day at the museum, her not being eaten by a dinosaur will rank at or near the top, in the following sense: she'd pick satisfying that one ahead of all or almost all others. She'd pay more money to satisfy it than all or almost all others. So it would be not just a benefit but a *tremendous* benefit, on the suggested theory. But surely her overhearing these visitors and so forming this strong (in this sense) desire doesn't give her a *tremendous* benefit. Surely O2 is not *hugely* better for her than O1.

The desire theory of welfare, then, faces difficulties in cases of (a) idealistic desires, (b) compulsive desires, (c) prudential desires, and (d) desires concerning unlikely possibilities. I will now argue that all four problems can be solved in a satisfying way if we distinguish two senses of 'desire', and formulate the desire-fulfillment theory using just one of these senses.

2. Two Senses of 'Desire'

a. Characterizing the Two Senses

There is a certain natural and intuitive distinction among kinds of desire or senses of the term 'desire'. Different philosophers have been on to it in different contexts. I'll begin by reporting on the distinctions that some philosophers have drawn.

In his book *Pleasure and Desire*, Justin Gosling discusses some "familiar beginners' arguments on free will" (1969: 86). He lays out the following case, which is a useful place to start:

> I am sitting in my room after a morning's work, and badly want to get out. There is nothing I should like so much as fresh air and the smell of country flowers. I am just rising from my chair when there is a knock at the door. A note has come from an acquaintance who, I know, relies on my company and sympathy to keep him from melancholy and self-pity: . . . he is in the blackest of depressions; would I go round? I gaze longingly at the day outside, and with a sinking heart go to spend the afternoon trying to bring this man back to an even keel. When asked about my choice I shall of course say that what I really wanted to do was go out for a walk, and that I did not want to visit my friend at all. (Gosling 1969: 86)

Though this sort of case is commonplace, a reflective reader might balk at Gosling's last claim, the claim that he didn't want to visit his friend at all. Such a reader might reason as follows:

> *A Reaction to Gosling*: "The case involved a genuine choice. Gosling was free to do as he pleased. He could have gone out for a walk if that was what he most wanted to do, but he didn't; so it must be that that wasn't in fact what he most wanted to do. He must have more strongly wanted to visit his friend, or do his duty. So not only is it false that Gosling did not want to visit his friend at all, his desire to visit his friend was in fact even stronger than his desire to go for a walk."

Both parties—Gosling and the imagined reader—say things that seem reasonable, and true. Considered on its own, it seems that Gosling could be saying something true (and consistent with the rest of his story) when he says that he had no desire to visit his friend. And considered on *its* own, the imagined reader likewise seems to say something true when she says that Gosling must have wanted to visit his friend more strongly than he wanted to go for a walk. But these two reasonable thoughts are evidently mutually contradictory. A natural response is to posit an ambiguity. What the imaginary reader above says is true on one sense of the term 'desire' or 'want', but there is another, perfectly good sense of these terms in which Gosling's last remark—that he did not want to visit his friend at all—is true.

For an even quicker argument for ambiguity, due to Derek Parfit (2011: 43), consider a remark that one friend might make to another as they are trying to decide what to do on a day that they plan to spend together: "I don't want us to do what *I* want us to do; I want us to do what *you* want us to do." Since the first clause of this remark isn't in fact self-contradictory, 'want' must be ambiguous.

Language aside, there does indeed seem to be an attitude that Gosling bears towards going for a walk that he doesn't bear to any degree—let alone to a stronger degree—towards visiting his friend or doing his duty. Here are some terms that Gosling uses to characterize what is lacking in the latter case: "I had no enthusiasm for my duty whatever"; "That had no appeal for me, no interest or excitement"; "I could not view the prospect with pleasure" (89–90). Going for the walk, on the

other hand, did appeal very much to him. That is the attitude that he bore towards visiting his friend that he didn't bear towards his duty.

I am inclined to join many other philosophers in thinking that there are these two senses of 'desire'. There is a sense that is connected to voluntary action, intention, choice, and will. If we voluntarily and intentionally do something, it follows by that very fact that we wanted to do it, and indeed that we most wanted to do it (as compared with our other options). And there is a distinct sense of 'desire' that is connected with notions like enthusiasm, appeal, interest, excitement, and attraction.

This distinction is found in many places. One is G. F. Schueler's book *Desire*:

> On the one side is what we might be called *the philosophers' sense*, in which, as G. E. M. Anscombe (1963, 68) says, "the primitive sign of wanting is trying to get," that is, the sense in which desires are so to speak automatically tied to actions because the term "desire" is understood so broadly as to apply to whatever moves someone to act. On the other side is *the more ordinary sense*, in which one can do things one has no desire to do, that is, the sense in which one can reflect on one's own desires, try to figure out what one wants, compare one's own desires with the desires of others or the requirements of morals, the law, etiquette or prudence, and in the end, perhaps, even decide that some desires one has, even very strong ones, shouldn't be acted on at all. (Schueler 1995: 1, italics added)[18]

Another is Wayne Davis' paper "The Two Senses of Desire." Davis claims that the term 'desire' expresses two different propositional attitudes, one he calls 'volitive desire' and the other 'appetitive desire'. About appetitive desires (Schueler's "ordinary sense" and the sense Gosling uses), Davis says, its objects are appealing and are "viewed with pleasure" (Davis 1986: 66). Volitive desires (Schueler's "philosophers' sense"), Davis says, are a better indicator of action, are based on reasons, are influenced by value judgments, and are entailed by intentions. Davis notes that appetitive desire influences volitive desire, but not vice versa (1986: 63).

Melinda Vadas begins her paper "Affective and Non-Affective Desire" by asking, "Is it possible to act without desire?" (1984: 273). She answers affirmatively, and says that those who dismiss this possibility do so because they fail to distinguish between two senses of the term 'desire': the affective sense and the non-affective sense (274). Here is how Vadas characterizes affective desire:

> 'Desire' (as a noun) may refer to an affect, that is, a feeling, emotion, or mood, such as a desire to eat pizza, have children, or run in a marathon. . . . Desire, in the affective sense, is both a *present* affect – I now relate affectively to, say, the thought of my having a pizza – as well as a projection of affect; that is, I in a sense 'predict' that I will feel a certain way in the future." (Vadas 1984: 276)

"But," she goes on, "there is another sense of 'desire,' a non-affective sense. In this sense I can be said to desire whatever goals I intentionally pursue" (Vadas 1984: 277, italics removed). If I do something voluntarily, *ipso facto* I wanted to do it in the non-affective sense. But, she writes, "nothing is phenomenologically plainer

than that we do not always act upon or even in accordance with our affective desires" (278).[19]

More or less this distinction is also found in T. F. Daveney 1961 (inclinational vs. intentional wanting), Thomas Nagel 1970 (unmotivated vs. motivated desire), Philippa Foot 1972, David Lewis 1988 (warm vs. cold desires), L.W. Sumner 1996 (attitudinal vs. behavioral wanting), Derek Parfit 2011 (narrow vs. wide wanting), Stephen Campbell 2013 (desire in the attitudinal sense vs. desire in the motivational sense), and Tamar Schapiro 2014 (the substantive sense vs. the placeholder sense). There is also David Hume's distinction between "calm and violent passions" (Hume 1739: III.IV).[20]

The distinction that these philosophers and others all appear to be on to is the one that I want to claim is axiologically relevant. Before attempting to characterize it further, I'll say that I'm not sure what terms to use for it. The terms 'affective' and 'non-affective' desire invite the thought that an affective desire is simply a non-affective desire with a certain phenomenological feel tacked on. But that wouldn't be right. For what is different about the one kind of state isn't that it is just the other kind of state plus a certain phenomenology. Rather, it is a difference in the intentional aspect of the state, in the way of relating to the object of the desire. I don't deny that there is a phenomenological difference between the two states; they plainly do differ with respect to what it is like to be in them. I just want to insist that the affective part isn't some inessential feature of the relation to the object of the desire, but part of the very stuff of the relation. Vadas, who uses the "affective" terminology, seems to agree; when we have an affective desire, she says, we are "relat[ing] affectively" to the object of the desire (Vadas 1984: 276).

Davis' term 'appetitive desire' seems too narrow. It suggests that we are talking only about "urges," or a sense that applies only to desires for things like water, sex, or a cigarette. Those are certainly included, but the mental state that I am talking about is also one that is paradigmatically had towards things like getting a job offer, having your candidate win the election, and making a free throw.

The term 'behavioral desire' seems fine for the other sense of 'desire' and I will use that.

For the earlier sense, I won't settle on one term. The unwieldy phrases 'desire in the genuine-attraction sense' and 'desire in the finding-appealing sense' do a decent job of capturing the notion, I think.[21] I'll sometimes use 'true desire', 'real desire', and 'the ordinary sense of "desire"'. I'll also sometimes use 'affective desire'.

Whatever terms we use, the following summarizes some key aspects of the distinction as I intend it:

"Desire in the Genuine-Attraction Sense"

- If a person has a genuine-attraction desire for some event to occur (or to have occurred or to be occurring), the person finds the occurrence of the event attractive or appealing, is enthusiastic about it (at least to some extent), and tends to view it with pleasure or gusto;
- When it comes to the genuine-attraction sense of 'desire', a person can voluntarily do an action that they had no desire to do, and they can refuse to do what they most desire to do;[22]

- For desire in the genuine-attraction sense, *strength* of desire is the strength of the genuine attraction to the event's occurring, or the degree to which the event's occurrence genuinely appeals to the desirer, or the degree to which they are enthusiastic about it.

"Behavioral Desire"

- Behavioral desire may simply be a "functional state," or a state defined by what it does; in this case: an intentional state that disposes the person in it to try to act in the ways that (according to the person's beliefs) would make its content true.[23]
- When it comes to the behavioral sense of 'desire', a person cannot voluntarily do an action that they had no desire to do; voluntarily doing an act entails having wanted to do it.
- For desire in the behavioral sense, *strength* of desire is determined by hypothetical choices.

So, for example, if a person faces five options, the one they would or do choose is by definition the one they most want, in the behavioral sense of 'want'. The one they would otherwise choose if that option weren't available is the one for which they have the next strongest desire, and so on. This understanding of strength of desire encourages the idea that only an ordinal ranking of strength of desire is possible.

In the *Treatise*, Hume suggests an account of our distinction that is different from my sketch in at least one significant respect. However, I don't think it matters for our purposes. Hume writes,

> There is not in philosophy a subject of more nice speculation than this of the different causes and effects of the calm and violent passions. It is evident passions influence not the will in proportion to their violence, or the disorder they occasion in the temper. (1739: III.IV)

Hume's calm and violent passions seem to correspond, respectively, to our behavioral and genuine-attraction desires. For Hume talks about the *influence that a passion can have on the will*—this is the behavioral sense—and contrasts this with the *violence* of the passion, or *the disorder it occasions in the temper*—which evokes our genuine-attraction sense. But Hume's remarks suggest that one and the same state might have *both* a degree of influence on the will and a degree of violence.[24] My characterization of the distinction, as well as the characterizations of most of the philosophers mentioned above, suggest, alternatively, that in such a case there would be two different states: a behavioral desire, which has a degree of influence on the will, and a true desire, which has a degree of violence.

But I doubt anything hangs on which of these two descriptions we choose. I think I could say everything I want to say about the problem cases for desire theories of welfare under the assumption that there is just one kind of thing, desire, that has magnitudes along two dimensions: influence on the will and violence. Provided that we allow magnitudes of zero along each dimension, the two ways of describing things seem straightforwardly intertranslatable.[25,26]

b. The Two Senses of 'Desire' and the Theory of Well-Being

Some philosophers of welfare assume that if desire is relevant to welfare, it is the behavioral sense that is relevant. In discussing a case of a person who voluntarily chooses to punish himself, Richard Kraut writes, for example,

> In any ordinary sense of "want," he doesn't want to punish himself, but the desire theory cannot take refuge in this point, since it uses a much broader notion of desire, according to which what we voluntarily seek is what we desire. And in this sense, our self-punisher does want above all to punish himself. (Kraut 1994: 41)

Some versions of the desire theory do use this sense of 'desire'. On Harriet Baber's preferentism, for example, "[p]reference is inextricably linked to choice" (Baber 2007: 106). This is also true of the desire theories of welfare favored by economists and decision theorists.[27] This is in part because economists have been especially interested to have an account of welfare that makes welfare measurable, and it is thought that at least on the behavioral sense of 'desire', desire is measurable by empirical methods, because it is revealed through our choices. Desires in the genuine-attraction sense are more difficult to measure, since our choices can float free of these desires. It should be said, however, that the fact that a theory of welfare makes welfare easier to measure is no indication that it is true.

A desire satisfactionism formulated in terms of genuine-attraction desire is, however, compatible with other, better motivations for the desire approach. The most important such motivation, in my view, is internalism about well-being, or the resonance constraint (Railton 1986: 9; Rosati 1996; Noggle 1999: 303). On Railton's seminal statement of the doctrine,

> what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive .... It would be an intolerably alienated conception of someone's good to imagine that it might fail in any such way to engage him.

Since the terms Railton uses—"attractive," "engage"—are about genuine attraction rather than mere behavioral disposition, his statement actually supports formulating the theory in terms of genuine-attraction desire *over* formulating it in terms of behavioral desire. Formulating the theory in genuine-attraction terms is consistent with other motivations as well, such as that the desire approach fits well with a naturalistic metaethic, is supported by reasons internalism, and enjoys the advantages of monism while avoiding the experience-machine objection.[28]

Not all discussants assume that the behavioral sense is the one that desire theories of welfare should use. Most writers simply make no comment on the matter. However, one philosopher, Wayne Sumner, shares the view that desire theories should make use of the genuine-attraction sense of 'desire'. Sumner writes, "[i]t is only in the attitudinal sense [his term for the genuine-attraction sense] that preference can be plausibly connected with welfare" (1996: 120). Sumner is no desire satisfactionist, however; he makes this remark only in passing, and gives no argument for it.[29]

I wish to argue for it. My aim in what follows is to show that if, for the purposes of the desire theory of welfare, we understand 'desire' in the ordinary, attitudinal, true, affective, inclinational, warm, appetitive, violent sense rather than in the merely behavioral, intentional, volitional, non-affective, cold, calm, wide, philosophers' sense, we can provide plausible solutions to the four problem cases described earlier. That the desire theory can solve a host of problems if it is formulated in a certain way is good evidence that it should be formulated in that way.

3. Application to the Problem Cases

a. Idealistic Desires

In *Pig Slice*, despite how much you want it, you let your friend have the last slice of pizza. Moreover, you never come to regret your decision: you continue, for as long as it ever crosses your mind, to prefer your friend's having had the slice to your having had it. The problem for preferentism is that it nevertheless seems that the correct way to describe this case is as one in which you acted against your own interests, for the sake of your friend's interests, despite your having gotten the outcome you most wanted and would continue to most want.

But if we distinguish between what you most wanted in the behavioral sense and what you most wanted in the genuine-attraction sense, and adopt a desire theory of well-being that makes use of the latter sense, we avoid this objection. For although you chose to give your friend the pig slice, this outcome held no appeal for you. What really appealed to you, or what you were genuinely attracted to, was having the slice yourself. When you give away the slice, you miss out on the valuable satisfaction of this true desire. Although you do fulfill a merely behavioral desire, one grounded in your ideals, this is of no benefit to you, according to the new theory. The theory thus agrees with the intuitive verdict that you chose against your own interest in this case.

I claimed that the prospect of your friend getting the pig slice held no appeal for you, and only your getting it did. That is unrealistic; most real-life cases are more complicated. We are sympathetic beings, and we are often genuinely attracted, to some degree, to the prospect of another person getting something that they are genuinely attracted to. But our sympathies are limited, and so even if the thought of your friend getting the pig slice holds some appeal for you, chances are that the thought of yourself getting it holds more appeal (and also that this appeal will continue more robustly on in the future, whereas the appeal of your friend having gotten the slice will diminish more quickly as time passes). Thus even in more realistic cases, we act against our own interest, on the imagined desire theory, if we give away the pig slice, since we would satisfy a more intense genuine-appeal desire if we ourselves indulge.[30]

Another real-world complication is that if our emotional responses are integrated with our values, we are likely to feel a little guilty or ashamed if we take the pig slice for ourselves. This will constitutively involve being repulsed, even if only somewhat, by what one has done; that is, it will involve having a genuine desire not to have done it. This will diminish how much it would be in your interest to take the pig

slice for yourself, according to the desire theory in question, since doing so would frustrate this true desire. This sort of desire can arise before the decision as well, if we have even more integrity. If one feels strongly enough that it wouldn't be right for one to take the pig slice, the prospect of taking it will lose much of its luster.

The case could be such that all of these factors combine to make it the case that the balance of desire satisfaction over frustration for the agent (even in the intended genuine-attraction sense of 'desire') tilts against choosing the pig slice for himself. The desire theory in question would then imply that it is not in fact in the person's interest to get it. But in such a case, this, I submit, is the right result. If you would have been rather into seeing your friend get the slice, and would have felt rather guilty had you taken it, and, due to your values, found the prospect of yourself getting it not all that attractive anyway, it very well can be in your own best interest to do the "selfless" thing and let your friend have the slice.[31] Just as a hedonist would never claim that idealistic pleasures don't benefit us, a desire satisfactionist should not claim that satisfying idealistic desires can never benefit us.

That this is so creates trouble for alternative desire-theoretic solutions to the problem of idealistic desires. It shows that it is not plausible simply to exclude idealistic desires on the grounds of their being based in our ideals and values, or their being about others' interests. For if we *truly care*, in the right way, about these values—they aren't just abstract ideals that we coolly or grudgingly subscribe to—or if we truly care, in the right way, about these people, it can be worse *for us* if the values are not realized or the others not benefitted. The proposals of some other philosophers for dealing with cases like these—such as the proposals of Sidgwick, Brandt, and Baber, which exclude such desires—cannot accommodate this fact.

The restriction to genuine-attraction desires thus appears to solve cases of idealistic desires as they are presented in manufactured thought experiments, and also generates the right results in the various more complex, real-world variations.

b. Compulsive Desires

Parfit's drug addict wakes each morning with a strong desire for the drug, a desire that is itself neither pleasant nor painful, and that he easily fulfills each morning by taking the drug. Parfit claims that a standard desire-fulfillment theory of welfare implies, implausibly, that Parfit has benefitted this person by turning him into such an addict. This is true if the theory makes use of the behavioral notion of desire. But what if it instead makes use of genuine-attraction desires?

Parfit doesn't say whether consuming the drug holds genuine appeal for the addict. Some addicts, we know, simply find themselves compelled to act in accordance with their addiction, though they look to the prospect with no gusto or enthusiasm. From Parfit's description of the case, I'm inclined to think that his addict's daily desires are like this. That is, the addict desires to take the drug merely in the behavioral sense. The addict here would be just like Quinn's Radio Man, who is in a state that simply tends to cause him to turn on radios; he does so merely habitually, with no emotion or enthusiasm. If that's right, then a desire theory formulated in terms of genuine-attraction desire will avoid the implausible result that the drug addict or Radio Man is better off for having his condition.

But what if either taking the drug or the effects of taking the drug do hold genuine appeal for the addict each morning? The addict is genuinely attracted to the way the drug will make them feel, and then gets to satisfy this desire, all the while remaining attracted to the way they are feeling as a result of taking the drug. Assuming, unrealistically, but as Parfit does, that there really are no bad side effects to this addiction, the desire theory I am imagining would then indeed imply that having this addiction is a benefit to the drug addict. But I believe that this is the right result, even if the addict wishes not to have the addiction.

For note that Parfit's unusual case has none of the usual hallmarks of addiction. The user doesn't need continually and dangerously to increase the dosage in order to satisfy their desires; and there are no undesirable side-effects of the addiction, such as illness, or opportunity costs, or worry about obtaining more of the drug. In fact, as this term is used in medicine, the case wouldn't even qualify as one of addiction. Here is a standard characterization of addiction:

> Addiction is a condition that results when a person ingests a substance (e.g., alcohol, cocaine, nicotine) or engages in an activity (e.g., gambling, sex, shopping) that can be pleasurable but the continuation of which becomes compulsive and interferes with ordinary responsibilities and concerns, such as work, relationships, or health. People who have developed an addiction may not be aware that their behavior is out of control and causing problems for themselves and others. (*Psychology Today*)

In Parfit's example, the use of the drug does not interfere with ordinary responsibilities and concerns and is not causing problems for the user or others. It's thus unfair, and likely biasing, to describe the case as one of addiction. The essentials of the case make it more akin to introducing someone to a new, easily-obtainable food, or to a new hobby, than to making them into a drug addict. And we indeed do benefit others when we turn them on to a new food or a new hobby that has no bad side effects and that they come to love.[32]

c. Prudential Desires

Consider next the smoker. The desire theory that conceives of desire merely behaviorally implies that the smoker had a good month, since he was at each moment getting what he wanted most—not to be smoking. But the smoker himself would surely tell us that it was a rough month, one he'd rather not have to repeat.

But if we instead focus only on the man's genuine-attraction desires, we see that his case is rather like the case of idealistic desire. (Perhaps it literally *is* a case of idealistic desire, where the relevant ideal is prudence, one of the traditional cardinal virtues.) The man has a very strong desire to smoke. He wants this badly, with gusto. By contrast, his not smoking holds no real appeal. Just like with your belief in *Pig Slice* in the value of selflessness, and how this gets you to choose against what you really want, likewise the smoker has the evaluative belief that smoking is bad for him, and this gets him to choose against what really appeals to him. Just going with what he really wants, the smoker is not faring well this month. Those desires are being frustrated continually. His strong behavioral desire is being

satisfied, true, but this doesn't matter on the new theory. Thus, the desire theory that appeals to what one really wants generates the correct result that the smoker had a bad month.[33]

d. Desires Concerning Unlikely Possibilities

Finally, recall the desire concerning the unlikely event that Lois is eaten by a dinosaur. She doesn't want this to happen. And, in the behavioral sense, this desire is quite strong. She would choose for this not to happen above most other possibilities she can think of. But what are her affective desires concerning this? Feldman's description of the case implies that they are weak or non-existent. Lois can't muster any enthusiasm against this possibility because it's just too remote. She knows it's not going to happen, and so she can't be bothered by it. Thus, the theory that is restricted to count only desires in the genuine-attraction sense can solve this problem case as well. The view agrees that it is no benefit to Lois to be made not to want, in the behavioral sense, to be eaten by a dinosaur. By contrast, desire theories that appeal to behavioral desires and determine their strength in terms of hypothetical choices arguably must say, implausibly, that Lois receives a tremendous benefit in being caused to form a behavioral desire about being eaten by a dinosaur.[34]

In "Of the Causes of the Violent Passions" from the *Treatise*, after introducing the distinction between "calm and violent passions," Hume notes (as we quoted earlier) that "passions influence not the will in proportion to their violence, or the disorder they occasion in the temper" (1739: III.IV). Hume goes on to point out that "The same good, when near, will cause a violent passion, which, when remote, produces only a calm one." What kind of nearness and remoteness does Hume have in mind? Surely at least *temporal* nearness and remoteness, as we noted in the discussion of the smoker case. But Hume might also have had in mind nearness and remoteness in the sense of *possibility* or *likelihood*. Applying this idea to our fourth case, the evil of the dinosaur attack is so remote—is such a remote possibility—that our passion against it has little to no violence. This thought is echoed later on in the section when Hume says, of the passions, that "uncertainty ... encreases them." Because we are certain that no dinosaur attack is coming, our passions concerning the possibility are calm, not at all violent. This is why fulfilling them is of no intrinsic benefit, on the theory I am suggesting. If we become less certain that no attack is coming—if we think that we really might be in danger—our passion would increase in violence, and its being satisfied would indeed be of intrinsic benefit to us.

4. Might My View Also Exclude Too Much?

Against desire-fulfillment theories that exclude certain sorts of desires from the utility calculus—such as idealistic desires or other-regarding desires—I complain that these theories exclude too much. For there are cases in which the satisfaction of one of these desires is intuitively a good thing for the person. Might the same complaint be made against my exclusion of merely behavioral desires?

One way to put this objection is to ask us to consider the coolest and calmest of rational beings: a being who desires things only in the behavioral sense. Suppose

this being is playing a game of chess, something it excels at, and its aim is to win. Suppose the being wins. Haven't things thereby gone better for the being than they would have if the being had lost? If the answer is Yes, this suggests that satisfying merely behavioral desires can be good in itself for a subject, something that my proposed theory denies.

Initially it may seem right to agree that things do go better for this being if it wins, but I believe that in fact winning the game is of no benefit to this being—and indeed that nothing is.[35] To begin, it's very important to make sure that we are imagining the case properly. Some readers might be imagining a creature like the famous character Spock from *Star Trek*. Doing so certainly helps to elicit the intuition that things go better for the being if it wins the game of chess. For surely Spock is a being for whom things go better and worse, and if so, then winning this game of chess seems like a good candidate for being a good thing for Spock. But to imagine the case this way is illegitimate, for Spock isn't in fact a being who lacks true desires. Although there is a tendency to think of Spock as an emotionless, detached being, this isn't true. Spock even cries (Diehl 1968). Spock desires things not merely in virtue of the fact that he makes voluntary choices; things genuinely appeal to him.

So what we need to imagine instead is a being who has goals, and can bring them about, but is in no way "invested" in them. The achievement of their goals doesn't genuinely appeal to this being, or excite them to any degree. The being is simply disposed, like a machine, to coolly and detachedly bring them about. When the being wins its game of chess, it does so joylessly, with no enthusiasm or real interest. Nor is the being genuinely averse to losing the game of chess. The being will try to win, of course, but won't mind if it loses. It wins chess games the way Deep Blue, IBM's chess-playing computer, does, or the way Quinn's Radio Man turns on radios.

Once we see that this is how this being would have to be, it is no longer intuitive to say that the being benefits when it satisfies its merely behavioral desire to win the game. Just as winning holds no appeal for the being, losing doesn't bother it one bit. It seems to get nothing out of it either way. Thus I believe that on closer examination, this case—similar as it is to the cases of Quinn's Radio Man and Parfit's drug addict—in fact provides confirmation for the restriction to desires in the genuine-attraction sense rather than evidence against it.

Let me conclude this section by mentioning a certain plausible possibility, though doing so raises vexed issues that we lack the space to explore in any depth. It may be that beings who (say) play and excel at chess the way Deep Blue does, but to whom nothing is genuinely appealing, *can* be benefitted or harmed *in an analogical or metaphorical sense*. This might be the same sense in which plants and cars can be said to be doing well or badly, or in which certain things can be said to be good or bad for them (as when we say that sunshine is good for plants and that failing to change the oil is bad for cars). That the sorts of beings that we are discussing in this section can do better or worse in this analogical or metaphorical sense may help to explain why some people would have the false impression that these beings are subject to genuine benefits and harms. Needless to say, the present theory is

not a theory of the analogical or metaphorical notion. Nor, in my view, is the analogical or metaphorical notion a notion with moral or normative importance.

* * * * *

There are many objections, issues, problems, and challenges concerning the desire-satisfaction theory of welfare. I took note of four such problems in this paper. I tried to articulate a distinction between two senses of 'desire': desire in the behavioral sense and desire in the genuine-attraction sense. It is a distinction that isn't discussed much in the well-being literature, but it is one that I believe to be axiologically crucial: only the latter sense of 'desire' is relevant to well-being. I have argued for this by showing how a desire-fulfillment theory of welfare formulated in terms of that sense of 'desire' can solve the four problems in a satisfying way.[36]

Notes

[1] Such as a hybrid theory on which desire satisfaction is an essential component of the basic welfare goods (as in Parfit 1984: 501), an objective-list theory with desire satisfaction on the list (as in Arneson 1999: 117 or Keller 2009: 659), a subjectivist pluralist view in which desire satisfaction is among the goods (as in Lin 2016), or a value-realization theory on which desires play a role in the account of valuing (as in Raibley 2010).

[2] For idealized theories, see Sidgwick 1907: 109–111; Rawls 1971: 92–93, 417; and Brandt 1979: 268, 113. For global desire theories see Parfit 1984: 495–498 and Carson 2000: 94. For self-regarding desire theories, see Dworkin 1977: 234; Overvold 1980: 10n; and the "Success Theory" throughout Parfit 1984. For second-order desire theories, see Railton 1986: 17; Hyde 2011; and perhaps Raibley 2010. For restrictions to non-moral or non-altruistic desires see Sidgwick 1907: 109; Brandt 1979: 247; and Baber 2007: 107. For a restriction to now-for-now desires, see Hare 1981: 101–106. For a restriction to autonomous desires, see Haji 2009.

[3] I explain why idealization isn't needed for many problems for which it is believed to be needed in Heathwood 2005. For a different kind of criticism of idealization, see Enoch 2005.

[4] See Matthews 1972.

[5] Cf. York 2015.

[6] Though I feel confident that it is both praiseworthy and not irrational. The latter may be relevant to our topic in that some desire theorists may wish to get the result that you don't benefit in bringing about this outcome by excluding your desire for it on the grounds that the desire is irrational. This strategy does not seem promising.

[7] If you think it matters, we can stipulate that you wouldn't have been wracked with guilt for years or otherwise been traumatized by your choice not to sacrifice yourself; you would have accepted your decision as reasonable in the circumstances and gone on to live a good life on any account.

[8] Cf. "Man Betrays His Heart" (2016).

[9] Allan Gibbard (1987: 139) presents a similar case:

> Consider a piece of cake to be divided between Desdemona and Iago. Desdemona is altruistic: given the choice, she would divide the cake equally. Iago is selfish, and given the choice he would take the entire cake for himself. They are of similar size and appetite, they eat cake with equal signs of gusto, and they will each undergo similar kinds of inconvenience in order to eat a cake that would otherwise go to waste.

[10] More exactly, Baber denies the existence of such preferences; she writes, "a plausible account of preference should not ascribe preferences to us on the basis of choices we make in response to feelings of moral obligation" (Baber 2007: 107). See also Broome 1978.

[11] In this connection, she cites Velleman 2002. John Harsanyi's preference-utilitarian theory "exclude[s] all clearly antisocial preferences, such as sadism, envy, resentment, and malice" from the

social-utility function. It is not clear to me, however, whether he means to do this by excluding it from individual welfare, or only at the level of social utility (Harsanyi 1977: 647).

[12] As in see Dworkin 1977: 234; Overvold 1980: 10n; and the "Success Theory" throughout Parfit 1984.

[13] The restriction to self-regarding desires is often made in response to what I call the problem of remote desires, a canonical illustration of which is Parfit's case of the stranger on a train (Parfit 1984: 494). My rejecting this standard solution (the restriction to self-regarding desires) to the problem of idealistic desires might raise the question of how I would solve the problem of remote desires. I present my solution, which does not rely on the distinction among desires that is the topic of the present paper, briefly in Heathwood 2006 and more fully in Heathwood (unpublished).

[14] But see Sobel 1994 and Rosati 1995 for important problems for idealized desire theories. See also Griffin 1986: 11 and Enoch 2005. Idealization is also not required to solve many or perhaps any of the problems it is commonly brought in to solve—non-idealized theories have the resources to solve them; see Murphy 1999, Heathwood 2005, and Lin unpublished.

[15] The restriction to intrinsic desires is orthodox. See, e.g., Sidgwick 1907: 109 and von Wright 1963: 103–4.

[16] See Heathwood 2016: 138–139.

[17] Summative desire theories that appeal to future desires allow for this at least. This more significant factor is why it would still be bad for Lois to be eaten by a dinosaur even if she never formed any desires about it at all, even while it was happening, as might happen if, say, she were eaten by a dinosaur while totally unconscious.

[18] This passage suggests that Schueler takes the first sense of 'desire' to be something of a philosopher's invention. Don Locke says something similar in spirit: "once 'desire' is interpreted in the broader, more formal way, they cease to be real existences" (1982, p. 243). For my part, I feel that the speech above by our imagined reader of Gosling comes naturally enough to non-philosophers that Schueler's and Locke's view is probably unfair. This aspect of their view is no part of how I understand the distinction.

[19] Ruth Chang (2004) also uses the term 'affective desire' and distinguishes affective desires from other sorts of desire.

[20] Not surprisingly, some philosophers implicitly reject ambiguity by claiming that one of our two senses corresponds to the true nature of desire while the other sense corresponds to an inessential but perhaps common feature of desire. Michael Smith (1994: §§4.5-4.6), for example, rejects what he calls the phenomenological conception of desire in favor of the dispositional conception. On the other side, Galen Strawson writes, "the link to the notion of affect dispositions is internal to and fundamentally constitutive of the notion of desire in a way that the link to the notion of behavioral dispositions is not" (Strawson 2010: 282).

[21] Though they still do not always do so unambiguously. For example, if a person is in a functional state that disposes him to turn on radios, there is some sense of 'attracted' in which it follows trivially that he was attracted to doing this (perhaps it is the same sense in which electrons are attracted to protons). But I hope it's clear enough that there is another sense—a more phenomenological or affective rather than behavioral sense—in which this fellow was in no way attracted to turning on radios.

[22] Is it also possible for a person to have literally no motivation to do or no disposition to choose what they most desire to do in the genuine-attraction sense? In other words, can there be truly non-motivating genuine-attraction desires? I'm not sure, but I can remain neutral here. The standard view seems to be that genuine attraction entails motivation. For many writers describe the behavioral or motivational sense of 'desire' as a *broader* or *wider* sense, which suggests the relevant entailment (see the passages by Schueler, Kraut, and Locke quoted earlier, and the fact that Parfit uses the labels 'narrow' and 'wide' to characterize the distinction). For a thought experiment suggesting that there can be non-motivating genuine-attraction desires, see Galen Strawson's "Weather Watchers" (2010: ch. 9).

[23] For a statement of such a conception of desire, see, e.g., Stalnaker 1984: "To desire that *P* is to be disposed to act in ways that would tend to bring it about that *P* in a world in which one's beliefs, whatever they are, were true" (15).

[24] Cf. Railton's thought that "desire is, underneath, a compound rather than simple state, with two distinctive aspects: a degree of *positive affective attraction* and a degree of *focused appetitive striving*" (Railton 2012: 31).

[25] One might wonder how the two notions of desire that I am calling our attention to relate to the main theories of the nature of desire, and whether this has the potential to create problems. The main psychological commitment of the theory of well-being that I will be proposing is that there be psychological states of finding something appealing, of being genuinely attracted to something, of being enthusiastic about something, or some similar notion. So long as some theory of desire is a (non-eliminativist) theory of this sort of state, my theory of well-being should be compatible with it. Now, other theories of desire have it that desires are very different sorts of animal than either of our two senses of desire would have it. On Timothy Schroeder's (2004) view, for example, "a desire can exist in a creature that cannot, by its nature, move or feel" (Schroeder 2015). On the assumption that being genuinely attracted to something requires the ability to feel, Schroeder's view implies that the states that I think are essential to well-being are not in fact desires. Though I think these states are properly called 'desires', I can remain neutral on this as far as the theory of welfare is concerned; if Schroeder is right, then I am a "genuine attraction theorist" rather than a desire theorist of welfare. It's no part of Schroeder's theory of desire that there is no such thing as genuine attraction.

[26] One might also wonder about the connection between the genuine-attraction form of desire and *pleasure*. My project in this paper does not require that I take a stand on the matter, but my view is that there is an intimate—indeed metaphysical—connection between these two phenomena (see Heathwood 2006: §4; Heathwood 2007; and Heathwood 2017). To be sure, desiring something in the genuine-attraction sense does not entail taking pleasure in it; one takes pleasure in something only if one thinks it is true or thinks it obtains, and we can of course desire what we know doesn't and won't obtain.

Incidentally, this connection to pleasure can in fact help to identify the concept of genuine-attraction desire: it is that state such that when one bears it to some object (e.g., a state, event, or proposition) while simultaneously believing that that object obtains, one thereby takes pleasure in the object. To be clear, this is not necessarily to analyze this sense of 'desire' in terms of pleasure; in my view, the desire notion is in fact the more basic one here, and pleasure is analyzable in terms of it (for a brief argument for this, see Heathwood 2009: 93).

In any case, nothing in this paper requires these views. For all I say here, a felt-quality rather than an attitudinal theory of the nature of pleasure could be true. And for all I say here, it could be that genuine-attraction desires are analyzable in terms of pleasure; so it could be both that desire satisfactionists should formulate their view in terms of genuine-attraction desire and that this form of desire is definable in terms of pleasure.

[27] In *Analytical Welfare Economics*, D. M. Winch writes, for example, "[A]n individual shall be considered better off if he is in a *chosen* position. This assumption relates the Paretian value judgement directly to the utility function. Since we define utility as that which the individual *attempts* to maximize, it follows that he will choose more rather than less utility. An increase in his utility can then be regarded as synonymous with his being better off" (Winch 1971, 33, both italics added; qtd. in Sumner 1996: 117n).

[28] For brief overviews of these arguments, see Heathwood 2016: 137–138.

[29] In a more recent paper, Peter Railton appreciates that conceiving of desire simply as a disposition to act may make it ill-suited to play a role in theories of normative phenomena, such as theories of reasons for action or theories of "an individual's objective interests," i.e., welfare (Railton 2012: 23). Railton proceeds to lay out an alternative "prospective model of desire" (35). He goes on to explore the implications of this theory for all manner of philosophical issues (39–44), but not for the theory of welfare.

Ruth Chang (2004) argues for the distinctive normative significance of affective desire, claiming that affective desires are the sole desiderative source of normative reasons for action.

[30] It is also worth mentioning that, in a case like this, we would also, if we were to eat the slice ourselves, satisfy numerous purely sensuous desires involving the gustatory experience of pizza eating. Such *de re* desires are formed in reaction to, rather than in advance of, experiencing the specific tastes of the particular episode. These genuine-attraction desire satisfactions, for which there are no corresponding

satisfactions in the outcome in which we forgo the slice of pizza, are also important in helping to explain why we act against our interest if we forgo the slice of pizza.

[31] It is an interesting question how to characterize the sense in which the act just imagined should not count as selfish despite being most in one's interest.

[32] To better understand how my theory handles certain sorts of cases, it is useful to consider another drug example. (I owe this objection to an anonymous referee.) Suppose a person has a genuine-attraction desire to take a certain drug, but doing so unexpectedly gives them no good feelings. If we add that the person also thinks that it would be wrong to indulge in the drug, it might seem implausible to say that anything has gone well for the person in this scenario. Does my theory imply otherwise?

It depends on how we fill in the details of the case. The most natural way to specify the case would be to say that what the person wants in taking the drug is to feel a certain way or to get a certain kind of experience. When a person has a desire to take some drug—e.g., to drink some liquor, to smoke a cigarette, to take a painkiller—they almost always want it for how they think it will make them feel. So let's suppose that the person in this case finds themself with a desire for this drug because they think that it will make them feel mellow and relaxed and they just want to feel mellow and relaxed. They give in to temptation and take the drug. However, to their surprise, it doesn't make them feel mellow and relaxed. In fact, the drug causes no changes in their experiences at all.

If this is how the case is specified, then I agree that nothing has gone well for the person in this situation. But my theory also agrees. For if this is the way the case is specified, what the person is genuinely attracted to are the feelings they think they'll get if they take the drug. They don't find the actual act of taking the drug appealing. Consequently, no satisfaction of a genuine-attraction desire has occurred. And so the theory agrees that no benefit occurs.

But what if the person is in fact genuinely attracted to the actual act of taking the drug? This isn't so unusual. It is a feature of our psychology that we tend to form genuine attractions to things that are closely associated with things that we are genuinely attracted to. A heroin addict may start to find the ritual of preparing and injecting the drug appealing in itself. Thus we have the second way to specify the case. The person in this version of the case is genuinely attracted to *the act of taking* the drug. Let's say that the person is a recovering heroin addict and has come to find the practice of preparing the drug, tying on the tourniquet, inserting the needle, and feeling the liquid coursing through their veins appealing in itself. In this case, the subject wants to take the drug both for what's involved in taking it and for how they think taking the drug will make them feel.

They take the drug, and their genuine-attraction desires for preparing the drug, tying on the tourniquet, inserting the needle, and feeling the liquid coursing through their veins are all satisfied. My theory thus implies that this is a direct benefit to them. But this is plausible. The person's genuine-attraction desire for feeling the effects of the drug, though, is unfulfilled, because, for some reason, there are no such effects. Things are even worse, though, because presumably this person, who is conflicted about all of this, finds the fact of his having given in to temptation and taken the immoral drug quite unappealing. This consists in more desire frustration for them. My theory thus will imply that the whole affair is a mixed result for them, and probably of negative overall value; they get a small temporary benefit associated with the act of taking the drug, but things go badly with respect to having unfulfilled desires to experience the high of the drug and unfulfilled desires not to have succumbed to temptation. These all seem like reasonable results.

[33] A bad month *in itself*, I of course mean. The month hopefully is an instrumentally very good month for him, in that it prevents his getting cancer decades later.

In case it's not already obvious, the way that my suggested desire theory would explain the misfortune of the smoker's getting cancer later, even if he cares not a whit about it now, is via the appeal to future desires: in particular, by appeal to (i) the desire frustrations that accompany getting cancer (most significantly those connected to suffering but also those connected to desires to do things that cancer prevents one from doing), and (ii) the desire satisfactions that cancer prevents, especially when it is fatal.

[34] However, won't Lois find the prospect of being eaten by a dinosaur unappealing, even in her emotionless state? Yes, but this just an illustration of the unfortunate fact that even the terms that I am using to characterize genuine-attraction desire don't all do so unambiguously. We can use the notions of appeal and attraction in a thinner sense as well, to refer to behavioral desire (see endnote 21 above). Language aside, what I need for my solution to work is for there to be two importantly different ways

that Lois might relate to the prospect of being eaten by a dinosaur: (i) calmly disposed to choose against it, which is what she is in the example, and (ii) affectively averse to it, which is what she would be if she thought it a live possibility.

[35] To be sure, I don't need this stronger claim that nothing could be good for such a being. For my topic is the question, "Which desires are relevant to well-being, assuming that some desires are?" This question arises even if things other than desire satisfaction are intrinsically good for people, such as excellence or achievement or knowledge. Our imagined rational being can have these things even if they can't have valuable desire satisfactions.

[36] For helping me to make this paper much better than it would have been, I would like to thank Andrew Alwood, Guy Bar Sadeh, Ben Bramble, Donald Bruckner, Stephen Campbell, Dale Dorsey, Molly Gardner, June Gruber, Elizabeth Harman, Jennifer Hawkins, Eden Lin, Jason Raibley, Connie Rosati, David Sobel, Jonathan Spelman, Peter Zuk, and audiences at the 2013 Princeton Workshop on Well-Being, the Center for Values and Social Policy at the University of Colorado Boulder, the 2014 Kansas Workshop on Well-Being, the Colorado Affective Sciences Laboratories at the University of Colorado Boulder, Virginia Commonwealth University, and Keio University. I am especially grateful to Eden Lin for valuable feedback on multiple drafts. Work on this essay was supported by a fellowship at the University Center for Human Values at Princeton University; I thank the Center for that support.

Bibliography

Adams, Robert M. (1999) *Finite and Infinite Goods* (Oxford University Press).

Anscombe, G. E. M. (1963) *Intention*, 2nd ed. (Cornell University Press).

Arneson, Richard J. (1999) "Human Flourishing Versus Desire Satisfaction," *Social Philosophy and Policy* 16: 113–142.

Baber, H. E. (2007) "Adaptive Preference," *Social Theory & Practice* 33: 105–126.

Brandt, Richard B. (1979) *A Theory of the Good and the Right* (Oxford University Press).

Broome, John (1978) "Choice and Value in Economics" *Oxford Economic Papers* 30: 313–333.

Campbell, Stephen M. (2013) "An Analysis of Prudential Value," *Utilitas* 25: 334–354.

Carson, Thomas L. (2000) *Value and the Good Life* (University of Notre Dame Press).

Chang, Ruth (2004) "Can Desires Provide Reasons for Action?" in R. J. Wallace, P. Petit, S. Scheffler, and M. Smith (eds.) *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Oxford University Press), 56–90.

Crisp, Roger (2017) "Well-Being," in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), <https://plato.stanford.edu/archives/fall2017/entries/well-being/>.

Daveney, T. F. (1961) "Wanting," *The Philosophical Quarterly* 11: 135–144.

Davis, Wayne (1981) "A Theory of Happiness," *American Philosophical Quarterly* 18: 111–20.

———. (1986) "The Two Senses of Desire," in J. Marks (ed.) *The Ways of Desire* (Precedent Publishing), 63–82.

Diehl, D. (1968) "Girls All Want to Touch the Ears," *The New York Times*, 25 Aug.: D17.

Dworkin, Ronald (1977) *Taking Rights Seriously* (Harvard University Press).

Enoch, David (2005) "Why Idealize?" *Ethics* 115: 759–787.

Feldman, Fred (2010) *What Is This Thing Called Happiness?* (Oxford University Press).

Foot, Philippa (1972) "Reasons for Action and Desires," *Proceedings of the Aristotelian Society, Supplementary Volumes* 46, 203–10.

Gibbard, Allan F. (1987) "Ordinal Utilitarianism," in G. R. Feiwel (ed.) *Arrow and the Foundations of the Theory of Economic Policy* (New York University Press), 135–153.

Gosling, J. C. B. (1969) *Pleasure and Desire* (Oxford University Press).

Griffin, James (1986) *Well-Being* (Clarendon Press).

Hall, Rich (1984) *Sniglets* (Macmillan).

Haji, Ishtiyaque (2009) *Freedom and Value* (Springer).

Hare, R. M. (1981) *Moral Thinking* (Clarendon Press).

Harsanyi, John (1977) "Morality and the Theory of Rational Behavior," *Social Research* 44: 623–56.

Heathwood, Chris (2005) "The Problem of Defective Desires," *Australasian Journal of Philosophy* 83: 487–504.

———. (2006) "Desire Satisfactionism and Hedonism," *Philosophical Studies* 128: 539–563.
———. (2007) "The Reduction of Sensory Pleasure to Desire," *Philosophical Studies* 133: 23–44.
———. (2009) "Desire-Based Theories of Reasons, Pleasure, and Welfare," *Oxford Studies in Metaethics* 6 (2011): 79–106.
———. (2011) "Preferentism and Self-Sacrifice," *Pacific Philosophical Quarterly* 92: 18–38.
———. (2016) "Desire-Fulfillment Theory," in G. Fletcher (ed.) *The Routledge Handbook of the Philosophy of Well-Being* (Routledge), 135–147.
———. (2017) "Unconscious Pleasures and Attitudinal Theories of Pleasure," *Utilitas* 30 (2): 219–227.
———. (unpublished) "The Problem of Remote Desires."
Hume, David (1739) *A Treatise of Human Nature*.
Hyde, Jason (2011) *Caring Satisfactionism: A New Theory of Personal Welfare*, Ph.D. dissertation, University of Colorado Boulder.
Keller, Simon (2009) "Welfare as Success," *Noûs* 43: 656–683.
Kraut, Richard (1994) "Desire and the Human Good," *Proceedings and Addresses of the American Philosophical Association* 68: 39–54.
Lewis, David (1988) "Desire as Belief," *Mind* 97: 323–32.
Lin, Eden (2016) "The Subjective List Theory of Well-Being," *The Australasian Journal of Philosophy* 94: 99–114.
———. (unpublished) "Why Subjectivists about Welfare Needn't Idealize."
Locke, Don (1982) "Beliefs, Desires and Reasons for Action," *American Philosophical Quarterly* 19: 241–9.
"Man Betrays His Heart By Telling Friend He Can Have Last Dumpling," *The Onion* (2016, Dec. 8), http://www.theonion.com/article/man-betrays-his-heart-telling-friend-he-can-have-l-54845.
Matthews, Gareth B. (1972) "Senses and Kinds," *The Journal of Philosophy* 69: 149–157.
Mill J. S. (1863) *Utilitarianism*.
Murphy, Mark C. (1999) "The Simple Desire-Fulfillment Theory," *Noûs* 33: 247–272.
Nagel, Thomas (1970) *The Possibility of Altruism* (Princeton University Press).
Noggle, Robert (1999) "Integrity, the Self, and Desire-Based Accounts of the Good," *Philosophical Studies* 96: 301–328.
Overvold, Mark Carl (1980) "Self-Interest and the Concept of Self-Sacrifice," *Canadian Journal of Philosophy* 10: 105–18.
Parfit, Derek (1984) *Reasons and Persons* (Oxford University Press).
———. (2011) *On What Matters*, Volume I (Oxford University Press).
Quinn, Warren (1993) "Putting Rationality in its Place," in R. G. Frey and C. W. Morris (eds.), *Value, Welfare, and Morality* (Cambridge University Press), 26–50.
Raibley, Jason (2010) "Well-Being and the Priority of Values," *Social Theory and Practice* 36: 593–620.
Railton, Peter (1986) "Facts and Values," *Philosophical Topics* 14: 5–31.
———. (2012) "That Obscure Object, Desire," *Proceedings and Addresses of the American Philosophical Association* 86: 22–46.
Rawls, John (1971) *A Theory of Justice* (Harvard University Press).
Rosati, Connie (1995) "Persons, Perspectives, and Full Information Accounts of the Good," *Ethics* 105: 296–325.
———. (1996) "Internalism and the Good for a Person," *Ethics* 106: 297–326.
Schapiro, Tamar (2014) "What Are Theories of Desire Theories Of?" *Analytic Philosophy* 55: 131–50.
Schroeder, Timothy (2004) *Three Faces of Desire* (Oxford University Press).
———. (2015) "Desire," in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*, URL = <http://plato.stanford.edu/archives/sum2015/entries/desire/>.
Schueler, G. F. (1995) *Desire* (The MIT Press).
Sidgwick, Henry (1907) *The Methods of Ethics* (Macmillan and Company).
Smith, Michael (1994) *The Moral Problem* (Blackwell).
Sobel, David (1994) "Full Information Accounts of Well-Being," *Ethics* 104: 784–810.
Stalnaker, Robert (1984) *Inquiry* (MIT Press).
Strawson, Galen (2010) *Mental Reality*, second edition (MIT Press).
Sumner, L. W. (1996) *Welfare, Happiness, and Ethics* (Oxford University Press).

Vadas, Melinda (1984) "Affective and Non-Affective Desire," *Philosophy and Phenomenological Research* 45: 273–279.

Velleman, J. David (2002) "Motivation by Ideal," *Philosophical Explorations* 5: 89–104.

Von Wright, Georg Henrik (1963) *The Varieties of Goodness* (The Humanities Press).

"What Is Addiction?" *Psychology Today* <http://psychologytoday.com/basics/addiction>, accessed September 2017.

Winch, David M. (1973) *Analytical Welfare Economics* (Penguin).

York, Kaya (2015) "Lesser-Known Trolley Problem Variations," *Timothy McSweeney's Internet Tendency* (McSweeney's Publishing), 20 Apr. <http://www.mcsweeneys.net/articles/lesser-known-trolley-problem-variations>.