

Desire-Fulfillment Theory

Chris Heathwood

7,457 words

draft of November 27, 2014

1. Introduction and Historical Background

The desire-fulfillment theory of well-being — also known as desire satisfactionism, preferentism, or simply the desire theory — holds, in its simplest form, that what is good in itself for people and other subjects of welfare is our getting what we want, or the fulfillment of our desires, and what makes things go worse for us is our wanting something to be the case when it is not or does not become the case. Most or all desire theorists would agree that the stronger the desire, the more beneficial is its satisfaction and the worse its frustration. There is less consensus over whether how long the desire is held is directly relevant to the value of its fulfillment or frustration. On the question of how good an entire life would be for a person, there are two main ways a desire approach might go: it can simply sum the values of all the instances of desire satisfaction and frustration *within* that life; or it can look to the person's desires *about* that whole life and hold that the best life is the one the person most wants to lead. These views differ because a person may prefer to lead a life that contains less preference satisfaction. A desire is fulfilled, according to standard forms of the theory, just if the desired state of affairs occurs; the subject need not experience any feelings of fulfillment.

The desire-fulfillment theory is a form of subjectivism about well-being in the rough sense that, according to it, getting a good life has to do with one's attitudes towards what one gets in life rather than the nature of those things themselves. There are other forms of subjectivism — e.g., aim-achievement theories, value-realization theories, happiness theories, and some forms of hedonism — but the desire-fulfillment theory is the archetype. Objective theories of well-being — such as perfectionism and the objective-list theory — maintain, by contrast, that at least some things that are intrinsically good or bad for us do not essentially involve our pro- or con-attitudes. Desire-fulfillment also plays a central role in some hybrid theories of well-being, which combine subjective and objective elements.

The desire-fulfillment theory is nowadays undoubtedly one of the leading theories of well-being. Some philosophers regard it to be *the* leading theory, “the theory to beat,” “[t]he dominant account among economists and philosophers over the last century or so ...” (Haybron 2008: 3). If it is the dominant theory of the 20th and 21st centuries, it received much less attention before then. Some leading ancient and medieval philosophers brought up the view in order to reject it. In Plato's *Gorgias* (c. 380 B.C.E.), for instance, it is Socrates' foil Callicles who asserts that “he who would truly live ought to allow his desires to wax to the uttermost, and ... minister to them and to satisfy all his longings” (491e-492a). In *De Trinitate* (c. 416 C.E.), St. Augustine briefly discusses the idea “that all are blessed, whoever live as they will,” claiming that Cicero refuted it (XIII, 5). Augustine goes on to assert, however, that desire (or will) fulfillment is at least necessary for well-being. In *Summa Theologiae* (c. 1274 C.E.), Thomas Aquinas speaks favorably of “the definition of beatitude that some have posited—viz., that the

blessed man is he who has everything that he desires” (I-II.5.8), but he does not in the end endorse a true desire-fulfillment theory.

Some major figures of the early modern period were more sympathetic to a genuinely subjectivist view. Thomas Hobbes is often mentioned as an early adopter of a desire theory of well-being, due to this passage in *Leviathan* (1651):

whatsoever is the object of any man’s appetite or desire that is it which he for his part calleth *good*; and the object of his hate and aversion, *evil* For these words of good [and] evil ... are ever used with relation to the person that useth them, there being nothing simply or absolutely so (ch. 6)

In his *Ethics* (1677), Baruch Spinoza writes, “in no case do we ... desire anything, because we deem it to be good, but ... we deem a thing to be good, because we ... desire it (Spinoza 1677: Part III, Prop. IX). It is not clear that Spinoza is talking about well-being as opposed to just plain value, but because Hobbes suggests that he rejects the very notion of value *simpliciter*, there are stronger grounds for interpreting him as talking about well-being.

That is how Henry Sidgwick interprets Hobbes when, in *The Methods of Ethics* (1907), Sidgwick begins what may be the first in-depth discussion of the desire-fulfillment theory of well-being (I.IX.3).¹ In that discussion, Sidgwick comes to the nowadays orthodox view that the theory is more promising if (simplifying somewhat) we

identify [a person’s good] not with the actually desired, but rather with ...

what would be desired ... supposing the desirer to possess a perfect forecast,

¹ The theory also seems to be endorsed by Joseph Butler in his *Fifteen Sermons* when he writes that “the very idea of an interested pursuit, necessarily presupposes particular passions or appetites; since the very idea of interest, or happiness, consists in this, that an appetite; or affection, enjoys its object” (1726, preface §31).

emotional as well as intellectual, of the state of attainment or fruition.

(110-11)

Though Sidgwick does not ultimately endorse a view of this sort, the doctrine that he formulates later inspires John Rawls' view in *A Theory of Justice* (1971) that

A person's good is determined by what is for him the most rational long-term plan of life ... the plan that would be decided upon as the outcome of careful reflection in which the agent reviewed, in the light of all the relevant facts, what it would be like to carry out these plans and thereby ascertained the course of action that would best realize his more fundamental desires. (92-93, 417)

When the desire-fulfillment theory of welfare finally took root in the early- to mid-20th century, it did so perhaps most deeply among economists. Early welfare economists, such as A.C. Pigou, accept the classical utilitarian doctrine that "the elements of welfare are states of consciousness" (1920: I.5; II.1) But, recognizing the need for something scientifically measurable, Pigou proposes that these welfare states "be brought into relation with a money measure" (II.1). And he saw that this could be done only indirectly: it must be "mediated through desires and aversions." Later welfare economists drop the underlying view that ultimate value lay wholly in the states of mind, and come to understand preference satisfaction itself as constituting rather than merely indicating well-being. John Harsanyi, for example, states his adherence to "the important philosophical principle of *preference autonomy*," "the principle that in deciding what is good and what is bad for a given individual, the ultimate criterion can only be his own wants and his own preferences" (1977: 645).

At the same time, philosophers, too, came to endorse preference-based accounts in larger numbers. In *The Varieties of Goodness* (1963), for example, the Finnish philosopher G.H. von Wright explains the notion of “a positive constituent of our good (welfare)” in terms of what “we should rather have than continue to be without” (107). In addition to von Wright and Rawls, other prominent, early advocates among philosophers include the political theorist Brian Barry (1965) and moral philosophers Richard Brandt (1966), Peter Singer (1979), and R.M. Hare (1981).

The desire-fulfillment theory’s rise to prominence is also partly attributable to its role in decision theory. Although early statements of the principle of expected utility are explicitly evaluative, the principle later came to be understood simply in terms of desires and aversions. In “Truth and Probability” (1926), for instance, F.P. Ramsey stipulates that he will “call the things a person ultimately desires ‘goods’,” and “emphasize[s] that in this essay good and bad are ... to be understood ... simply as denoting that to which a given person feels desire and aversion” (173–4).

And “Today,” some writers believe, “the desire-satisfaction theory is probably the dominant view of welfare among economists, social-scientists, and philosophers, both utilitarian and non-utilitarian” (Shaw 1999: 53).

2. Arguments for the Desire-Fulfillment Theory

The fundamental principles of axiology might be the most basic normative truths, and so we might not expect to find many direct arguments for them. Still, there is at least one interesting such argument for the subjectivist approach to well-being, one that provides at least

indirect support for the desire-fulfillment theory. The argument appeals to *internalism about well-being*, which Peter Railton (1986: 9) puts as follows:

... what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware.

Since desiring is a paradigm way of finding something compelling or attractive, this principle suggests a link between welfare and desire. Why think the principle is true? I suspect that, to many people, as it does to Railton, it simply seems right. It is indeed hard to believe that we can benefit someone by giving her things with which she is utterly unimpressed and in which she will remain forever uninterested. Other philosophers have offered arguments for internalism.²

Another kind of argument for the desire-fulfillment theory is based on the idea that it fits well with a naturalistic metaethic, and hence a naturalistic worldview more generally. This may be related to the theory's popularity among economists. One naturalistic approach in metaethics holds that normative or evaluative properties are to be identified with those natural properties that elicit certain responses, or are the object of certain attitudes, in certain observers. Such an approach might hold that the property of being beneficial for some subject, *S*, *just is* the property of being an object of a desire of *S*. This metaethical thesis implies a version of the desire theory of welfare. It is sometimes thought that pluralistic or objective theories of welfare are more difficult to square with naturalism.³

² E.g., Rosati 1996 and Velleman 1998. For criticism, see Sarch 2011. For some discussion of how best to formulate internalism, see Heathwood 2014: §2.

³ Though see Hooker 1991.

Another way to argue for a desire-fulfillment theory of well-being is from a desire-based, or internalist, theory of reasons for action. The latter asserts, roughly, that the only thing a person has reason to do is satisfy her desires. It may seem to be a datum for any theory of reasons that it accommodate that a person always has some reason to do what is in her own interests. The way for a reasons internalist to accommodate this datum is to endorse a desire-fulfillment theory of well-being.⁴

Yet another line of reasoning in support of the desire-fulfillment theory begins with the intuitive idea that getting what you want is at least *a* good thing for us, and then subjects the strengthened, unified hypothesis that it is the *only* good thing to scrutiny, attempting to falsify it; the argument then claims that the unified hypothesis survives the scrutiny, and we are thus justified in accepting it. Hedonism can be argued for on similar grounds. But desire theorists may claim that the desire-fulfillment hypothesis is more plausible than the hedonistic hypothesis in two ways. First, one of the most popular arguments against hedonism — the experience machine objection — does not apply (at least not in the same straightforward way) to the desire theory.⁵ Second, if we consider someone who is familiar with pleasure and doesn't want it as much as she wants other things, there is some plausibility to the claim that it is better for her to get the other things. This intuition favors the desire theory over hedonism. Hedonistic theories that make use of a desire theory of pleasure — the view, roughly, that for an experience to be pleasurable is for the person experiencing it to want to be experiencing it — may avoid this argument, but may also collapse into a desire theory (Heathwood 2006).

⁴ However, see Lin forthcoming a.

⁵ This objection is derived from the thought experiment in Nozick 1974: 42–5.

Whether this last argument for the desire approach succeeds depends on the extent to which the approach has the resources to deflect the many lines of objection that have been advanced against it. Before we turn to those, we'll conclude this section with a final line of thought in support of the desire-fulfillment theory. It concerns the ability of the desire approach to accommodate the goods posited by competing theories. The main competing theories are hedonistic and objective theories. Concerning hedonism, either a desire-based account of the nature of pleasure is true, or it isn't. If it is true, then the desire theory of well-being can accommodate the data the pleasure is good and pain bad for their subjects.⁶ If it isn't true, then so much the worse for that alleged data; for if pleasure is instead a certain distinctive kind of feeling or feeling tone, one a subject may have no interest in, then it's not at all clear that it is a good thing for such a subject to experience this (to him) neutral feeling (cf. Sobel 2005: 444-6). As for putative objective goods, such as knowledge or friendship, note that such goods are desired by virtually everyone. The desire theorist can thus explain why they might seem to be universal, objective goods. But when we imagine a strange person who truly has no interest in them, the desire-theoretic commitment that they are of no benefit to that person is at least as plausible as the objectivist insistence that they are (cf. the doctrine of internalism about well-being, discussed earlier). Desire fulfillment may be the common denominator on the scene in cases of apparent objective and hedonic goods, the factor that indeed explains the value in these cases.

⁶ Though see Lin forthcoming b.

3. Arguments Against the Desire-Fulfillment Theory

a. Mere instrumental fulfillments

We begin with a maximally unadorned theory, according to which whenever someone wants something to be the case, and it is or becomes the case, this is a benefit to the person. But suppose the person wants the thing to be the case only as a means to something else. For example, suppose she wants it to snow in the mountains so that the skiing will be good for her upcoming trip there, and only for that reason. Suppose it does snow in the mountains, but that she had to cancel her trip. Intuitively, the fulfillment of her desire that it snow was not in the end of any benefit to her.

The obvious solution is for the desire-fulfillment theorist to restrict the theory to count as intrinsically good for us only the fulfillment of desires for things their own sakes, or what are sometimes called *intrinsic desires*.⁷ This restriction is usually accepted uncritically; however, it isn't obviously unproblematic. Suppose a father 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-fulfillment theory to agree. But, for all that, such a desire on the father's part might be merely instrumental.

b. Ill-Informed Desires

There is a cherry pie before me and I am dying for a slice. Unbeknownst to me, I have recently developed a severe allergy to cherries and so it would in fact not be in my interests to satisfy my desire to eat the slice. This appears to conflict with the unadorned desire-fulfillment

⁷ See, e.g., Sidgwick 1907: 109 and von Wright 1963: 103–4; on a related solution, there simply are no such things as instrumental desires (Murphy 1999).

theory, according to which all desire fulfillments benefit a person. The restriction to intrinsic desires, while it will exclude some ill-informed desires (e.g., those based on false beliefs about what means might bring about a desired end), appears not to help here, since my desire to eat the pie is intrinsic.

About such cases, it might often be true that if the person knew all the facts, he would not have the problematic desire. This inspires the standard solution to the problem of ill-informed desires: idealization. The *informed desire theory* holds that what is good in itself for us is our getting what we *would* want if we knew and vividly appreciated all of the non-evaluative facts (Sidgwick 1907: §3; Rawls 1971: 417). If I knew how eating the pie would affect me, probably I would no longer want to eat it.

An alternative response to the objection from ill-informed desires requires no modification to the theory (Heathwood 2005). The objection claims that the unmodified theory implies that it is in my interests to satisfy my desire to eat the allergenic pie. But consider two things we might have in mind when we say that it is in my interests to satisfy some desire. We might mean that it is in my interests *all things considered* — that is, taking all the effects of satisfying the desire into account. Or we might mean merely that it is good *in itself* for me — intrinsically good for me — to satisfy the desire. The objection assumes, plausibly, that it is not in my interests *all things considered* to satisfy my desire for the pie. But the original unidealized desire theory can accommodate this, for if I satisfy my desire to eat the food, this will cause many of my other desires — desires not to feel sick, desires to go on a hike, etc. — to be frustrated on into the future. The original theory is committed only the claim that it is *good in itself* for me to satisfy my desire to eat the food. But, ignoring the effects — which is what one does when

evaluating a claim of intrinsic value — it intuitively *is* good for me to get to eat this piece of pie I very much want to eat. One advantage of this solution is that it is not hostage to the empirical conjecture that if I were to become idealized, I would lose all desire for the pie. Another is that it avoids problems that idealization may introduce.

c. Unwanted Fulfillments of Ideal Desires

Idealizing theories are indeed subject to objections that non-idealizing theories don't face.

James Griffin writes,

It is doubtless true that if I fully appreciated the nature of all possible objects of desire, I should change much of what I wanted. But if I do not go through that daunting improvement, yet the objects of my potentially perfected desires are given to me, I might well not be glad to have them; the education, after all, may be necessary for my getting anything out of them. That is true, for instance, of acquired tastes; you would do me no favour by giving me caviar now, unless it is part of some well-conceived training for my palate. (1986: 11)

Suppose we do give Griffin caviar now. The informed desire theory implies that we have indeed done him a favor, since, although he in fact has no interest in caviar, we have satisfied a desire that (we can suppose) he would have had if he were fully and vividly informed about the taste of caviar. Giving caviar to Griffin's *idealized* self might very well benefit *that* person, but theories of welfare are also supposed to tell us what things are good for schleps like you and me. Perhaps the underlying problem here is that an idealized desire theory of the sort under consideration

seems to abandon internalism about well-being, a basic intuition that motivates the desire theory in the first place.

The standard response to this problem is not to abandon idealization but to move to the *ideal advisor theory* (Railton 1986: 16; Rosati 1996). On one way to understand this proposal, what is good for a person to get is not what she would want herself to get were she idealized, but what she would want her actual, unidealized self to get were she idealized. Though Griffin's ideal self wants caviar *for himself*, perhaps he would not want his roe-averse actual self to get it.

But the ideal advisor version of idealization brings with it new problems. One is that it's at least possible that one's ideal advisor finds one's ignorance, inexperience, and poor taste pathetic, and consequently feels only disdain for one, and wishes one ill. Griffin's ideal advisor might think, "If I'm ever that ignorant and uncultivated, then shoot me," or, less fanatically, "... then give me caviar anyway." One might attempt to emend the ideal advisor theory by having it appeal to one's *benevolent* and informed desires. We could stipulate that "The ideal advisor's sole aim is to advance the well-being of the advisee" (Arneson 1999: 127). But such an account appears viciously circular. It seems essentially to be telling us that what is good for a person to get is what someone who wants what is good for this person wants this person to get.⁸

d. Base desires, malicious desires, pointless desires

Those who think that enjoyment is in general a good thing sometimes doubt that all enjoyment is good, for some instances of it are base and others malicious. But desires can be similarly base or malicious. There are also desires that seem simply unworthy even if not base or

⁸ For further problems with idealizing theories, see Sobel 1994 and Rosati 1995.

malicious, as in Rawls' imagined case of a talented intellect whose aim in life is "to count blades of grass in various geometrically shaped areas such as park squares and well-trimmed lawns" (Rawls 1971: 432).

For desire theorists who have already signed on for idealization, it is tempting to call on it whenever problems arise. Thus an ideal desire theorist might hope that no one who was fully and vividly informed about all of their possibilities would want to spend their time breaking crockery while drunk, torturing kittens, or counting blades of grass. But it is hard to see why full and vivid information must in all cases extinguish such desires. Some suspect that idealizers who would make such claims are unconsciously assuming that the idealization process includes ridding one of desires for things it's simply not good to get. But such an appeal would evidently require there to be desire-independent welfare goods, and thus require abandoning the desire theory.

Another response is simply to "bite the bullet" and insist that the subjects are no worse off for desiring in the unconventional ways in which they do. This reply is bolstered when we are reminded that some such desires are still criticizable morally and aesthetically even if not prudentially.

e. Remote desires

"Since my desires can range over spatially and temporally remote states of affairs," L.W. Sumner writes,

it follows that the satisfaction of many of them will occur at times or places too distant from me to have any discernible effect on me. In such cases it is

difficult to see how having my desire satisfied could possibly make my life go better. (1996: 125)

A concrete case due to Derek Parfit has become stock in the literature:

Suppose I meet a stranger who has what is believed to be a fatal disease. My sympathy is aroused, and I strongly want this stranger to be cured. We never meet again. Later, unknown to me, this stranger is cured. On the Unrestricted Desire-Fulfillment Theory, this event is good for me, and makes my life go better. This is not plausible. (1984: 494)

A special case of the problem concerns the fact that our desires can be fulfilled after we are dead.

Mark Overvold is a desire theorist who admits that “it is hard to see how anything which happens after one no longer exists can contribute to one’s self-interest” (1980: 108), and proceeds to develop a theory that delivers the desired result in the sorts of cases we are considering. On Overvold’s proposal, a desire had by some person is relevant to her welfare just in case it is a desire for a state of affairs that can obtain at some time only if she exists at that time (1980: 10n). On this *self-regarding desire theory*, since *the stranger’s being cured* can obtain at some time without Parfit existing at that time, the fulfillment of Parfit’s desire for it is of no benefit to Parfit. Overvold’s theory also rules out posthumous harm and benefit.

Overvold’s restriction to self-regarding desires may exclude too much, however. A persuasive example is the desire that the team one roots for wins. It is very important to some people that their team win, and they hope for it as intently as they hope for anything about themselves. It does not seem plausible to claim that the fulfillment of such a desire is of no benefit to the desirer simply because it did not involve a desire that was about her. This

objection also makes trouble for an alternative solution: that it is the fulfillment of our *aims* rather than our desires that benefits us.

On a third kind of solution, the remoteness that is anathema to welfare is remoteness from what we are aware of, or what we experience (Heathwood 2006: §2). The reason Parfit isn't benefitted when the stranger is cured is that the stranger is cured unbeknownst to Parfit. Note that it does seem more plausible that Parfit receives a benefit in a variant of the case in which Parfit learns that the stranger has been cured. Unlike the previous solutions, this solution allows that the fulfillment of desires that aren't about me, such as my team's winning, can nevertheless benefit me.

This solution does, however, imply that nothing that fails to enter or otherwise affect my awareness or experience can benefit me. If my spouse has an affair — something I am strongly averse to — some thinkers want to say that I am harmed by this even if I never find out about it and it never affects anything else that I have desires about. If I am harmed and, more generally, what you don't know *can* hurt you, then this theory of experienced desire fulfillment fails, and we are left without a solution to the problem of remote desires. Some philosophers bite the bullet up front and insist that things do go better for Parfit when, unknown to him, the stranger is cured (Lukas 2010).

f. Unwanted Desires

“Knowing that you accept a Summative theory” — the kind of desire theory that determines the value of your life by summing the values of the desire fulfillments and frustrations within it — Derek Parfit tells you,

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. ... 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. (1984: 496)

Parfit believes that few people would take him up on his offer, yet a summative desire-fulfillment theory implies that we would be better off if we did. Although we might often wish that we were not addicted to this drug, the disvalue of these desire frustrations would (we can suppose) be outweighed by the value of the repeated daily fulfillments.

It is sometimes thought that “complication[s] ... created by the fact that sometimes we have desires—those created by addictions, for example—that we wish we were without ... can easily be handled in familiar ways by giving special weight to second-order desires” (Kraut 1994: 40). On this proposal, only fulfillments of those desires that one wants to have contribute to one’s well-being. This solution may help with the addiction case, but it would seem to exclude too much. Unreflective people — people who live in the moment and never pause to consider their desires or take up any attitudes towards them — don’t all have worthless lives. Likewise for those mentally disabled people and animals who are incapable of higher-order mental states.

Parfit himself believes that his case shows that “Global versions” of the desire theory are superior. Since these theories “appeal only to someone’s desires about some part of his life, considered as a whole, or about his whole life,” they “ignore your particular desires each morning for a fresh injection” (1984: 497). It was a global desire theory that was discussed or

endorsed in the earlier passages by Sidgwick and Rawls. But, again, what of those of us who don't have global desires, or can't have them?⁹ Another objection to the move to global desires calls into question the presumption that when global and local desires conflict, global desire should always be considered authoritative (de Lazari-Radek and Singer 2014: 221).

There is a familiar distinction among desires, between what a person "truly desires" or finds truly appealing, and what a person wants in the thinner, merely behavioral sense that he is simply disposed to try to get it.¹⁰ This distinction isn't discussed much in the welfare literature, though one exception is Sumner, who, though not a desire theorist, maintains that "[i]t is only in the [former, "true appeal"] sense that preference can be plausibly connected with welfare" (1996: 120). Perhaps a theory restricted to this narrower sense of desire can answer Parfit's objection, since, as he describes them, the daily desires for the drug seem merely behavioral; taking the drug holds no genuine appeal for the addict.¹¹

g. Idealistic Desires, Self-Sacrificial Desires

Robert Adams points out that

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. ...

⁹ An *idealized* global theory, which asks which whole lives such people would want if they *were* to have global desires, is an option worth considering.

¹⁰ See, e.g., Davis 1986 and Schueler 1995: 1

¹¹ Cf. the view of psychologist Kent Berridge, who argues that "'wanting' can be activated without 'liking'" and that this phenomenon "has special relevance for understanding the causes of addiction." The suggestion in the main text is that the desire theorist count only those desires that are involved in Berridge's liking, and that Berridge's wanting involves merely behavioral desire.

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. (1999: 87-88)

A related case is that of self-sacrificial desires, though the objection here is a little different. According to the argument from self-sacrifice, desire theories fail because they imply, absurdly, that self-sacrifice is impossible (Overvold 1980). For an act to count as an act of self-sacrifice, it would seem that it must be (i) voluntary, (ii) informed, and (iii) not in the agent’s best interest. But, the argument claims, if (i) and (ii) are satisfied, (iii) cannot be, given standard desire-fulfillment theories of welfare. For if an act is voluntary, it is the one the agent most wants to do; if it is also informed, then, on either simple desire fulfillment theories or full-information variants, it is thereby in the agent’s best interest, and so condition (iii) cannot be satisfied.

One natural solution to the problems created by idealistic desires is simply to exclude them from the theory by fiat. Mill holds a view along these lines for determining the value of a pleasure, excluding preferences that are based on a “feeling of moral obligation” (1863: 12). In his discussion of the desire theory, Sidgwick sets down that he will consider “only what a man desires . . . for himself—not benevolently for others” (1907: 109).

Such proposals face problems similar to those faced by theories that restrict to self-regarding desires. Plausibly, devoted parents are sometimes benefitted when their intrinsic

desires concerning their children's welfare are satisfied; presumably some such desires are altruistic. Conversely, desires based on moral considerations should, intuitively, also sometimes count. People can become quite invested in justice, for example; if the just outcome is their heart's desire, it doesn't seem right to rule out all possibility of benefit.

Perhaps we need not exclude idealistic or self-sacrificial desires to solve the problems they raise. It has been argued that even the simplest, fully unrestricted sort of desire theory can accommodate self-sacrifice, so long as it is of the sort described above as 'summative' (Heathwood 2011). Even if an agent brings about the outcome she most prefers, that outcome can still contain within it less desire satisfaction for her than some alternative outcome available to her, making the act not in her best interest, even if voluntary and informed. Another solution, combinable with the one just mentioned, counts only the narrower sense of 'desire' mentioned above, the sense of finding the object of the desire truly appealing. These solutions are more flexible than those that simply exclude idealistic desires: they allow us to say that in cases of grudging obedience to the ideal, no benefit accrues, whereas in cases of enthusiastic embrace of the value, benefit does accrue.

h. Changing Desires

According to Richard Brandt,

The fundamental difficulty for the desire-satisfaction theory is that desires change over time: Some occurrence I now want to have happen may be something I did not want to have happen in the past, and will wish had not happened, if it does happen, in the future. (Brandt 1982: 179).

So suppose I want, for years, to go skydiving on my 40th birthday. But as the day approaches, my interests change, and I become strongly averse to doing this.

Plausibly, when my 40th birthday comes, it is in my interest to satisfy my present desire *not* to go skydiving at the expense of frustrating my past desires to go skydiving (at least if we assume that I won't later have persistent desires in the future to have done it). And perhaps this remains true no matter how long held and strong the past desires to go skydiving were. This suggests that to determine what benefits a person, we can ignore her past desires completely.

However, sometimes we do act so as to satisfy the merely past desires of people we care about. For example, we heed the wishes of the dead concerning how to treat their remains. Do we do this for their *benefit*? It's not obvious that we do, but if we do, that implies that we believe that it is in their interests to have this merely past desire satisfied. One kind of theory ignores only those past desires that are "conditional on their own persistence," or that we want satisfied only if we still have the desire when the time comes to satisfy it. Presumably, the desire in the skydiving case is conditional on its own persistence whereas our desires about how to treat our remains after we die are not conditional on their own persistence.¹² Another possible solution holds that fulfilling a past desire does result in a benefit, but a benefit that occurs, as it were, retroactively, when the desire was held (Dorsey 2013). Perhaps in the skydiving case we care only about present and future benefit, while in the death case we care about past benefit.

If, however, fulfilling merely past desires is never a benefit at any time, this suggests the view that the desire theory count only desires for what goes on at the time of the desire. As R.M. Hare, a proponent of this view, puts it, the theory "admits only now-for-now and then-for-then

¹² Cf. Parfit 1984: 151 and Bradley and McDaniel 2008: §10.

preferences,” to the exclusion of any now-for-then or then-for-now preferences (1981: 101–3).¹³ But might this exclude too much? Suppose that I do in fact strongly regret, for years, not having gone skydiving on my 40th birthday. If so, perhaps it was in my interests to force myself to go skydiving, despite my strong aversion to it at the time, for the sake of satisfying the “then-for-now” desires I would come to have. If that’s right, this suggests a surprising asymmetry: the desire theory of well-being should ignore future-directed desires but count present- and past-directed desires. There is a possible explanation for such an asymmetry. When we have a future-directed desire, we can’t now experience its satisfaction. But with present- and past-directed desires, we often are aware that they are satisfied. If the asymmetrical view is most plausible, this may provide an indirect argument for including an awareness requirement into the theory, as discussed earlier.

4. Conclusion

There are other objections to the desire approach worthy of our attention. When someone can’t get what he really wants, he may adapt his preferences to his predicament. If he succeeds in doing this, he is now getting everything he wants. This seems like an unfortunate situation, but the desire theory may be unable to accommodate this intuition.¹⁴

The theory may even lead to paradox. Suppose, out of self-loathing, I want only to be badly off. Either I am badly off or I am not. If I am badly off, then my only desire is fulfilled, and so, on the desire theory, I am not badly off. If, on the other hand, I am not badly off, then my

¹³ Cf. the “concurrence requirement” in Heathwood 2005.

¹⁴ On adaptive preferences, see Nussbaum 2000: ch. 2; Baber 2007; and Bruckner 2009.

only desire is frustrated, and I am badly off. In short, the desire-fulfillment theory implies the contradictory thought that, in some cases, a person is badly off if and only if he is not badly off.¹⁵

There is a Euthyphro objection: at the times when we are thinking just about ourselves and our interests, don't we want the things we want because they are good for us? But the desire theory suggests the opposite, that these things are good for us because we want them. There is an objection from Buddhism: doesn't Buddhism teach that the way to well-being is the extinction of all desire? There are objections from manipulated or non-autonomous desires: if subliminal advertising brainwashes us into wanting some silly gadget, does it really benefit us to get it?

Despite all of these potential objections facing the desire-fulfillment theory of well-being, the theory remains a leading one. As the above discussion illustrates, desire theorists have resources at their disposal when attempting to undermine these objections. And no doubt further resources will be discovered and developed by the theory's proponents, and novel objections will be advanced by its opponents, as the philosophical study of well-being continues to grow.

REFERENCES

Aquinas (1274) *Summa Theologiae*, trans. A. J. Freddoso (2014), URL = <<http://www3.nd.edu/~afreddos/summa-translation/TOC.htm>>.

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

Augustine (416) *De Trinitate*, many editions.

¹⁵ On the paradox for desire theories of well-being, see Heathwood 2005: §VI; Bradley 2007; and Skow 2009.

- Baber, H. E. (2007) "Adaptive Preference," *Social Theory and Practice* 33: 105–126.
- Barry, B. (1965) *Political Argument*, London: Routledge & Kegan Paul.
- Berridge, K. (1999) "Pleasure, Pain, Desire and Dread: Hidden Core Processes of Emotion," in D. Kahneman et al (eds.) *Well-Being*, New York: Russell Sage Foundation, 525–557.
- Bradley B. and McDaniel K. (2008) "Desires," *Mind* 117: 267–302.
- Bradley B. (2009) "A Paradox for Some Theories of Welfare," *Philosophical Studies* 133: 45–53.
- Brandt, R. B. (1966) "The Concept of Welfare," in S. R. Krupp (ed.) *The Structure of Economic Science*, Englewood Cliffs, NJ: Prentice-Hall, 257–76.
- Brandt, R. B. (1982) "Two Concepts of Utility," in H. B. Miller and W. H. Williams (eds.) *The Limits of Utilitarianism*, Minneapolis, MN: University of Minnesota Press, 169–85.
- Bruckner, D. (2009) "In Defense of Adaptive Preferences," *Philosophical Studies* 142: 307–324.
- Butler, J. (1726) *Fifteen Sermons Preached at the Rolls Chapel*, London: J. and J. Knapton.
- Davis, W. (1986) "The Two Senses of Desire," in J. Marks (ed.) *The Ways of Desire*, Chicago, IL: Precedent Publishing, 63–82.
- de Lazari-Radek, K. and Singer, P. (2014) *The Point of View of the Universe*, Oxford: Oxford University Press.
- Dorsey, D. (2013) "Desire-Satisfaction and Welfare as Temporal," *Ethical Theory and Moral Practice* 16: 151-171.
- Elster, J. (1982) "Sour Grapes — Utilitarianism and the Genesis of Wants," in A. Sen. and B. Williams (eds.) *Utilitarianism and Beyond*, Cambridge: Cambridge University Press.
- Hare, R. M. (1981) *Moral Thinking*, Oxford: Clarendon Press.

- Harsanyi, J. (1977) "Morality and the Theory of Rational Behavior," *Social Research* 44: 623–56.
- Haybron, D. (2008) *The Pursuit of Unhappiness*, Oxford: Oxford University Press.
- Heathwood, C. (2005) "The Problem of Defective Desires," *Australasian Journal of Philosophy* 83: 487–504.
- Heathwood, C. (2006) "Desire Satisfactionism and Hedonism," *Philosophical Studies* 128: 539–563.
- Heathwood, C. (2011) "Preferentism and Self-Sacrifice," *Pacific Philosophical Quarterly* 92: 18–38.
- Heathwood, C. (2014) "Subjective Theories of Well-Being," in B. Eggleston and D. Miller (eds.) *The Cambridge Companion to Utilitarianism*, Cambridge University Press, 199–219.
- Hobbes, T. (1651) *Leviathan*, E. Curley (tr.), Indianapolis, IN: Hackett Publishing, 1994.
- Hooker, B. (1991) "Theories of Welfare, Theories of Good Reasons for Action, and Ontological Naturalism," *Philosophical Papers* 20: 25–36.
- Kraut, R. (1994) "Desire and the Human Good," *Proceedings and Addresses of the American Philosophical Association* 68: 39–54.
- Lin E. (forthcoming a) "Prudence, Morality, and the Humean Theory of Reasons," *Philosophical Quarterly*.
- Lin E. (forthcoming b) "Pluralism about Well-Being," *Philosophical Perspectives*.
- Lukas, M. (2010) "Desire Satisfactionism and the Problem of Irrelevant Desires," *Journal of Ethics and Social Philosophy* 5.
- Mill, J. S. (1863) *Utilitarianism*, London: Parker, Son, and Bourn.

- Murphy, M. (1999) "The Simple Desire-Fulfillment Theory," *Noûs*, 33: 247–272.
- Nozick, R. (1974) *Anarchy, State, and Utopia*, New York: Basic Books.
- Nussbaum, M. (2000) *Women and Human Development*, Cambridge: Cambridge University Press.
- Overvold, M. C. (1980) "Self-Interest and the Concept of Self-Sacrifice," *Canadian Journal of Philosophy* 10: 105–118.
- Parfit, Derek (1984) *Reasons and Persons*, Oxford: Oxford University Press.
- Plato 380 B.C.E.: *Gorgias*, B. Jowett. (tr.).
- Pigou, A. (1920). *The Economics of Welfare*, London: MacMillan & Co, 1932.
- Railton, P. (1986) "Facts and Values," *Philosophical Topics* 14: 5–31.
- Ramsey, F. P. (1926) "Truth and Probability," in R. B. Braithwaite (ed.) *Foundations of Mathematics and Other Essays*, London: Kegan, Paul, Trench, Trubner, & Co., 1931, 156–198.
- Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Rosati, C. (1995) "Persons, Perspectives, and Full Information Accounts of the Good," *Ethics* 105: 296–325.
- Rosati, C. (1996) "Internalism and the Good for a Person," *Ethics* 106: 297–326.
- Sarch, A. (2011) "Internalism About a Person's Good: Don't Believe It," *Philosophical Studies* 154: 161–184.
- Schueler, G. F. (1995) *Desire*, Cambridge, MA: The MIT Press.
- Shaw, W. (1999). *Contemporary Ethics*, Malden, MA: Blackwell Publishers.
- Sidgwick, H. (1907) *The Methods of Ethics*, seventh edition, London: MacMillan and Company.

- Singer, P. (1979) *Practical Ethics*, Cambridge: Cambridge University Press.
- Skow, B. (2009) “Preferentism and the Paradox of Desire,” *Journal of Ethics and Social Philosophy* 3.
- Sobel, D. (1994) “Full Information Accounts of Well-Being,” *Ethics* 104: 784–810.
- Sobel, D. (2005) “Pain for Objectivists: The Case of Matters of Mere Taste,” *Ethical Theory and Moral Practice* 8: 437—57.
- Spinoza, B. (1677) *Ethics*, R. Elwes (tr.).
- Sumner, L. W. (1996) *Welfare, Happiness, and Ethics*, Oxford: Oxford University Press.
- Velleman, J. D. (1998) “Is Motivation Internal to Value?” in C. Fehige and U. Wessels (eds.) *Preferences*, Berlin: de Gruyter, 88–102.
- Von Wright, G. H. (1963) *The Varieties of Goodness*, New York: The Humanities Press.