

## THE PROBLEM OF DEFECTIVE DESIRES

Chris Heathwood

This paper defends the actualist desire-satisfaction theory of welfare against a popular line of objection—namely, that it cannot accommodate the fact that, sometimes, it is bad for a person to get what he wants. Ill-informed desires, irrational desires, base desires, poorly cultivated desires, pointless desires, artificially aroused desires, and the desire to be badly off, are alleged by objectors to be defective in this way. I attempt to show that each of these kinds of desire either is not genuinely defective or else is defective in a way fully compatible with the theory.

The desire-satisfaction theory of welfare says, roughly, that one's life goes well to the extent that one's desires are satisfied. On standard 'actualist' versions of the theory, it doesn't matter what you desire. So long as you are getting what you actually want—whatever it is—things are going well for you.<sup>1</sup> There is widespread agreement that these standard versions are incorrect, because we can desire things that are bad for us—in other words, because there are 'defective desires'. The aim of this paper is to defend the actualist desire-satisfaction theory against the problem of defective desires. I aim to show how the theory can accommodate the obvious fact that we can desire things that are bad for us. Admittedly, there are kinds of allegedly defective desire the theory cannot accommodate, but these desires, I argue, turn out not to be defective in the relevant way.

### I. The Problem of Defective Desires

Cases abound in which it appears to be bad for a person to get what he wants. We have *ill-informed desires*. Suppose I'm thirsty and therefore desire to drink from the river. Suppose also that the river is polluted and that drinking from it would make me sick [Carson 2000: 72–3]. Clearly, it would be bad for me to drink from the river, even though it is something I want to do. If only I had more information, I would no longer desire to drink from the river, and would consequently be better off.

We have *irrational desires*. Suppose I need to see a dentist for a procedure that will require drilling. And suppose this time lack of information isn't my

<sup>1</sup>The theory is called 'actualist' to distinguish it from a popular variant: an 'idealist' theory, which counts the satisfaction not of one's actual desires but of one's 'ideal desires'—typically, the desires one would have if one were fully informed and free from cognitive error.

problem: I know it is in my long-term interest to see the dentist. But still I want no part of it. Clearly, it would be bad for me not to go, even though it is something I want. If only I weren't weak-willed, I would come to desire to see the dentist, and would consequently be better off [Schwartz 1982: 196].

We have *base desires*. Suppose my strongest desires would be satisfied by a perpetual indulgence in bestiality [Moore 1993: 146–7]. Some venture to say that any theory that implies that the satisfaction of such desires is good is as false as it is paradoxical. Actualist desire theories are thought to have this implication.

We have *poorly cultivated desires*. Suppose I prefer Muzak over Mozart, mashed potatoes over masterpiece cuisine, and pushpin over poetry. It is thought that I would be better off if I desired the better things, and got them, than if I desired and got the mediocre things I now desire [Schwartz 1982: 196].

These last two objections are similar. The idea behind them might be (though need not be) that certain things are somehow inherently more worthy of desire than other things, and this desire-worthiness translates into enhanced welfare when the worthy things are gotten. The objects of base desires are at one end of the spectrum, such that getting them is positively bad; the objects of noble desires are at the other end, and getting them is super; somewhere in the middle are the mediocre things, and getting them isn't so great. That we put effort into cultivating a taste for certain things we think worthy, and that we encourage some desires over others in our children and other loved ones, suggests that the idea that the objects of desire can vary in worth is common.

Yet another variation of this same idea is found in the possibility of *pointless desires*. Suppose my only desire is 'to count blades of grass in various geometrically shaped areas such as park squares and well-trimmed lawns' [Rawls 1971: 432]. All I want is to count the blades of grass, all day, every day. It is thought that I would be better off if I desired to do something more worthwhile, and did that, than if I did the utterly pointless thing I now desire to do.

Continuing the litany, we have *artificially aroused desires*. Suppose as a result of a late-night indulgence in infomercials, I want nothing more than to own a Flowbee Vacuum Haircut System. Suppose the desire is artificially aroused: I have it only due to the in-your-face marketing campaign's effect on my insomniac mind. It is tempting to think that the satisfaction of such desires does not benefit their subject.<sup>2</sup>

Finally, we *desire to be badly off*. Suppose that, guilt-ridden by past crimes, I seek to punish myself by taking an arduous, boring, and insignificant job [Kraut 1994]. Suppose I land one and become pretty miserable. Clearly, it is bad for me to get the job, even though it is what I want. If I were less susceptible to guilt, I would have no desire for such a job, and would consequently be better off.<sup>3</sup>

<sup>2</sup>I thank Jason Raibley for encouraging me to discuss this class of allegedly defective desire. I cannot here discuss Parfit's drug-addiction case [1984: 497], which might be classified as another kind of case of artificial arousal.

<sup>3</sup>A class of desires I will not discuss here are what Robert Adams calls idealistic desires [1999: 88]. We desire idealistically when we sacrifice the maximization of our own welfare for something else we value, such as virtue, knowledge, dignity, or the welfare of others (see also Overvold [1980]). In Heathwood [2004], I attempt to show how actualist desire theories of welfare can accommodate these self-sacrificial desires.

The problem of defective desires, in one or another of the variants catalogued above, is regarded as a decisive objection to actualist desire-satisfaction theories of welfare by quite a number of writers on the topic. Robert Adams [1999: 87] considers it to be ‘the gravest disadvantage of the theory’. James Griffin [1986: 10] maintains that this ‘objection to the actual-desire account is overwhelming’. Fred Feldman [2004: 16] deems it ‘the central problem’ for actualism. Endorsement of the objection is given also by Elizabeth Anderson [1993: 129], Richard Brandt [1979: ch. 6], Thomas Carson [2000: 80–2, 88–92], Shelly Kagan [1998: 38–9], Richard Kraut [1994], and Thomas Schwartz [1982: 195–7].

## II. The Actual Desire-Satisfaction Theory

What is the theory thought to be so refuted by the problem of defective desires? First, assume that desires are propositional attitudes. *Desire satisfactions*, then, are states of affairs in which a subject desires that some proposition be true and in fact the proposition is true. *Desire frustrations* are states of affairs in which a subject desires that some proposition be true but the proposition is not true.

All parties to the debate seem to agree that any desire-satisfaction theory must count only *basic* (or intrinsic) desires [Brandt 1979: 111; Parfit 1984: 117]. If I desire to turn on my CD player only because I desire to hear the Pixies, I am made better off only if the latter, basic desire is satisfied. My desire to turn on my CD player is merely *instrumental* (or extrinsic): I have it only because I desire something else, and I think the thing instrumentally desired will lead to that something else. The theory is therefore interested only in ‘basic desire satisfaction’—that is, the satisfaction of one’s basic desires. Intuitively, a basic desire is a desire for something for its own sake, not merely for something else that it will lead to, or otherwise realize.

The simplest possible actualist desire-satisfaction theory of welfare would contain the following theses:

- (i) Every basic desire satisfaction is intrinsically good for its subject; every basic desire frustration is intrinsically bad for its subject.
- (ii) The intrinsic value for its subject of a basic desire satisfaction = the intensity of the desire satisfied; the intrinsic value for its subject of a basic desire frustration = – (the intensity of the desire frustrated).
- (iii) The intrinsic value of a life (or segment of a life) for the one who lives it (in other words, the total amount of welfare in the life (or life-segment)) = the sum of the intrinsic values of all the basic desire satisfactions and frustrations contained therein.

We can take this theory to be a paradigmatic version of an actualist desire-satisfaction theory, and the target theory of the problem of defective desires.

The theory is *actualist* rather than *idealist*, since it counts towards welfare the satisfaction of one’s actual desires, not the desires one would ideally have. The

theory is *summative* rather than *global*: it implies that the amount of welfare in a life is obtained by summing over all the satisfactions in the life, not just the satisfactions of one’s ‘global desires’ (the desires about one’s whole life, or about a largish part of one’s life).<sup>4</sup> The theory is a *satisfaction* version rather than an *object* version of the desire theory, since it is desire satisfactions rather than the objects of desire that are intrinsically good and bad for a subject [Rabinowicz and Österberg 1996; Bykvist 1998]. Importantly, I mean the theory above to require *concurrency*: in order for a state of affairs to count as a genuine instance of desire satisfaction, the state of affairs desired must obtain *at the same time* that it is desired to obtain. If I desire fame today but get it tomorrow, when I no longer want it, my desire for fame was not satisfied. A desire of mine is satisfied only if get the thing while I still desire it, and continue to have the desire while I’m getting it. The theory above therefore does not take the *duration* of a desire to be as prudentially significant as some have taken it to be. It doesn’t matter, welfare-wise, how long I desire something before I get it. All that matters are the following: that I still desire it while I’m getting it; how long the concurrent desiring and getting last; and how intense the desire is while it is being satisfied.<sup>5</sup>

## III. The Form of The Argument

We can abstract away from the particulars of the types of argument outlined in Section I and be left with the general form of the argument from defective desires:

- (1) There are defective desires: desires whose satisfaction does not make their subject better off.
- (2) But if the actualist desire-satisfaction theory is true, then there are no defective desires.
- (3) Therefore, the actualist desire-satisfaction theory is not true.

If these premises are true and univocal, actualist forms of the desire-satisfaction theory are done for. Call this the ‘Main Argument’.

Though the Main Argument captures the structure of all the arguments suggested above, the arguments subdivide reasonably well into three categories: (i) arguments from ‘all-things-considered defective’ desires, (ii) arguments from ‘intrinsically defective’ desires, and (iii) a third category to be named later.

<sup>4</sup>On global desires and on the distinction between global and summative theories see Parfit [1986: 496–9] and Carson [2000: 73–4]. Though I use Parfit’s terminology here, I think the label ‘summative’ is misleading. Even Parfit’s global theory is summative in the sense that the welfare value of a whole life, according to the theory, is equal to the sum of the values of the smaller bits of welfare in the life (in this case, the satisfaction or frustration of global desires).

<sup>5</sup>What if the intensity of the desire changes over the time that it is being concurrently satisfied? We can avoid this by defining desires as things that occur at instants (or at very brief intervals of time). At each brief interval at which some concurrent desire satisfaction occurs, there is just one intensity. We can say that the intensity of a desire has ‘changed’ when a person has a desire of some intensity for some proposition at some brief interval and then, at the next brief interval, has a desire of a different intensity for the same proposition.

#### IV. Arguments from All-Things-Considered Defective Desires

Some things are intrinsically bad for people. It is the job of a theory of welfare to identify these things. According to hedonists, episodes of pain are intrinsically bad for people. According to actual desire-satisfaction theorists, desire frustrations are intrinsically bad for people. In each case, the allegedly bad items are bad in themselves, rather than for what they lead to. The items would still be bad even if they led to nothing else bad. We can put this more precisely, and theory-neutrally, as follows: a state of affairs *p* is *intrinsically bad* for someone *S* iff given two lives exactly alike except with respect to *p*, the *p*-life is worse for *S* than the not-*p*-life.

Other things are bad for people in a different way. They are bad not in themselves but for what they lead to, or what they prevent. Many desire satisfactions are like this—e.g., many ill-informed desires and irrational desires of the sorts described above (and likewise for some of the other kinds of desire catalogued above, but more on that later). Recall my ill-informed desire to quench my thirst by drinking from the polluted river. There is nothing intrinsically or necessarily bad about having my thirst quenched. The desire is defective merely accidentally: its satisfaction just happened to be bad overall, in the long run, due to the fact that the river happened to be polluted.

Let's say that if something is bad in this way, then it is *all-things-considered* bad for the subject. That is, a state of affairs *p* is all-things-considered bad for someone *S* iff the life *S* would lead if *p* were to obtain is worse for *S* than the life *S* would lead if *p* were not to obtain. That is, iff *S* would have been better off had *p* not obtained.

Given this simple distinction, it is not difficult to see how actualist desire-satisfaction theories have the resources to deem ill-informed and irrational desires defective. Despite the fact that the theories declare all desire-satisfactions to be intrinsically good for their subject, the theories can nevertheless also declare some desire-satisfactions to be bad — *all-things-considered bad*.<sup>6</sup> Since intrinsic badness and all-things-considered badness can come apart, the theory is consistent with—indeed, it entails—the claim that many desire-satisfactions are all-things-considered bad. To make this more clear, let's take a closer look at the examples.

##### A. The Argument from Ill-Informed Desires

Assume my desire to quench my thirst is a basic desire. Assume the object of this desire is the proposition *that I quench my thirst within the next hour*. We all agree that this ill-informed desire is defective because, were I to satisfy it, I would be worse off than I would be if I did not satisfy it. The future I would have if I did satisfy this desire would start out pretty good—my thirst would be quenched. But it would quickly turn bad as sickness set in. The future I would have if I chose not to satisfy the desire to quench my thirst

would start out not so great—I would be thirsty. But that's about as bad as it would get (at least as far as this case goes), since I would avoid illness. All told, the life I would lead were I not to quench my thirst then and there—call it 'L1'—is better than the life I would lead—call it 'L2'—were I to quench my thirst then and there.

These are the facts. And actualist forms of the desire-satisfaction theory fully accommodate them. The actualist theory entails that L1 is better than L2 because, due to the sickness suffered in L2 (and despite the extra desire satisfaction (the thirst-quenching within the hour) in L2), L1 contains a more favourable balance of satisfaction over frustration. It avoids those desire frustrations that accompany being ill, whose negative value outweighs the positive value of satisfying the desire to quench within the hour. Far from being a counterexample to actualist theories, this case of an ill-informed defective desire actually provides confirmation to actualist desire-satisfaction theories of welfare.

So we say that my desire to quench my thirst in this example is all-things-considered defective. If the concept of all-things-considered defectiveness is the one being used in the Main Argument above, then premise (2) is false: according to the actualist desire-satisfaction theory, there are indeed defective desires.

The other concept was that of *intrinsic* defectiveness, the idea of something being bad in itself for a subject, whether or not it happens to lead to anything else bad or prevent anything good. So long as we attend only to basic desires (as any desire-satisfaction theory requires), it is not plausible to say that merely ill-informed desires are intrinsically defective. If I intrinsically desire to quench my thirst right now, and the desire is satisfied, this is intrinsically good for me. Saying this is consistent with saying that it is not in my best interest to quench my thirst, that I ought not to quench my thirst, and that it would be all-things-considered bad for me to quench my thirst. Once we see what is meant—and what is *not* meant—by saying that the satisfaction of the desire is intrinsically good for me, there is no compelling reason to deny it.

And there are compelling reasons not to deny it. Consider what it is to say that my desire to quench my thirst is intrinsically defective. To do this, we need to compare initial segments of two possible lives. First imagine the initial segment of L1, the life in which I do not quench my thirst. This segment—call it 'S1'—ends, let's say, a few minutes after I choose not to quench my thirst. These last few minutes of S1 are spent with a parched throat and a longing for liquid. Now imagine the corresponding initial segment of L2, the life in which I quench my thirst. This segment—call it 'S2'—is just like S1 except that I choose to satisfy my desire to be quenched. The last few minutes of S2 are spent with a wet whistle and little longing. Importantly, S2 ends before any pollutant-induced illness sets in.

Now I ask you to compare S1 and S2. I think you will agree that the total level of welfare is higher in S2 than in S1. For S2 has everything S1 has and more: the satisfaction of one extra, intense, basic desire, the existence of some added enjoyment, the avoidance of an annoying thirst. Actualist desire-satisfaction theories therefore correctly classify S2 as better than S1.

<sup>6</sup>This idea is hinted at by Carson [2000: 72–3] and Schwartz [1982: 197–8]. I hope to offer a more developed account of this reply.

If the Argument from Ill-Informed Desires is interpreted as about intrinsic defectiveness, then, yes, according to actualist desire-satisfaction theories, the ill-informed desire is not defective. But this is the correct result. Premise (1) of this interpretation of the Main Argument is false.

### B. The Argument from Irrational Desires

An analogous treatment is available for the version of the Main Argument based on irrational desires. My desire to avoid the pain of drilling at the dentist *is* defective according to actualist desire-satisfaction theories. It is all-things-considered defective. This is because, were the desire to be satisfied, a less favourable balance of desire satisfaction over frustration would result than would result if the desire were not satisfied. This is due to the desire frustration that would ensue from avoiding the dentist (such as that accompanying toothaches, the inability to eat the foods one likes, etc.). Again, it is not plausible to say that the desire to avoid the pain of drilling is intrinsically defective. If I were to die immediately after the time of the dentist appointment no matter what I did, it would clearly be better for me, in terms of welfare, to avoid the dentist and spend my last minutes doing something I really wanted to do.

### C. The Argument from Artificially Aroused Desires

Oscar Wilde quipped that the only thing worse than not getting what one wants is getting what one wants.<sup>7</sup> Sidgwick lamented a related point, noting that the objects of our strongest desires often come to us as ‘Dead Sea apples’, no longer wanted once they are gotten, ‘mere dust and ashes in the eating’ [1907: 110]. Sidgwick thought that moving to an ideal desire theory, one requiring full information about the objects of desire, could solve the problem, since we could learn in advance whether the objects of our desires would come to us flat [1907: 110–11].

But actualist desire-satisfaction theories can accommodate the Dead-Sea-apple phenomenon. They can do so by building concurrence into the theory, as I have done above. Genuine desire satisfaction is had only when the desire remains once its object is gotten. The concurrence requirement ensures that the getting of Dead Sea apples doesn’t improve welfare, since the very reason the thing is a Dead Sea apple is that the desire for it has vanished.

I claim that, typically, artificially aroused desires are special cases of the Dead-Sea-apple phenomenon. If my desire for the infomercial gizmo is merely an effect of clever marketing, chances are that once I get the device, I’ll no longer want it. A desire-satisfaction theory with concurrence will not

<sup>7</sup>Or at least these words are commonly attributed to him. What he actually said (in print, anyway) comes to the same thing: ‘In this world there are only two tragedies. One is not getting what one wants, and the other is getting it. The last is much the worst; the last is a real tragedy!’ [*Lady Windermere’s Fan*, 1892: Third Act].

entail that the satisfaction of this artificially aroused desire is good for me. Actual desire-satisfaction theories therefore are compatible with the idea that satisfying artificially aroused desires is often not good for us. Such desires are often defective because their objects are often Dead Sea apples.

But what if the device arrives in the mail and I am glad to see it; and I remain satisfied with the product? Actualist desire theories will imply that I’m made better off because of it. I say, anti-paternalistically, that is the correct result. A desire’s arising through some process we find objectionable is not on its own enough to make the desire itself defective.

It is important to add that there is another way in which artificially aroused desires can be deemed defective by actualist desire-satisfaction theories: such desires may be all-things-considered defective. Suppose I’m happy to see the infomercial item when it arrives, but that, in order to afford it, I broke my budget. In such cases—and no doubt cases are common in which marketing seduces one into buying what one can ill-afford—the satisfaction of the desire would typically be all-things-considered defective. The life I would lead were the desire not satisfied is better than the life I would lead were it satisfied. This is consistent with actualist desire-satisfaction theories.

Maybe there is yet something else behind the Argument from Artificially Aroused Desires. If I acquire a desire by means of some artificial process (such as overzealous marketing, or brainwashing), it is tempting to think that satisfying the desire is not good for me because I don’t ‘really want’ the thing, because the desire conflicts with my ‘true self’, because the desire is in some way ‘inauthentic’. If this is the idea behind the Argument from Artificially Aroused Desires, then the argument will not undermine the actual desire-satisfaction theory. For the only sense that can be made of the idea of a desire being ‘inauthentic’, or in conflict with one’s ‘true self’, is that the desire conflicts with many other desires held by the person. And if a desire conflicts with many other desires held by a person (in that satisfying it means frustrating others), the desire is (or could be) all-things-considered defective: the life the subject would lead were the desire not satisfied is worse than the life he would lead were it satisfied.

### D. The Argument from Poorly Cultivated Desires

The Arguments from Base Desires, Pointless Desires, and Poorly Cultivated Desires are most naturally interpreted as based on the claim that the desires in question are *intrinsically* defective. I treat arguments from intrinsically defective desires below in Section V. However, particularly regarding the Argument from Poorly Cultivated Desires, the claim might be understood as that such desires are merely all-things-considered defective.

But any such argument will fail, for the same sorts of reasons already discussed. Let me briefly illustrate. One reason it might be true that a person has ‘poorly cultivated desires’ is that the person has a set of desires that will lead to an overall lower balance in his life of desire satisfaction over frustration than he would have if he had some other, readily attainable set of

desires. Suppose I have frequent, and frequently satisfied, desires for Muzak and mashed potatoes. Suppose my twin has frequent, and frequently satisfied, desires for Mozart and masterpiece cuisine. How can an actualist desire-satisfaction theory imply that my twin's set of desires is better?

In several ways. My set of desires might be worse because I'm bound to get bored with Muzak and mashed potatoes. Eventually, I'll no longer want either of them, but, not having cultivated any more interesting desires, I'm able to get no satisfaction. Having these mediocre desires therefore makes me worse off, but not because satisfying them is no good; rather, it is because the mediocre things are less able to keep me desiring them. Another reason why my twin might be better off with his desires is that, once desires for such things as wonderful as Mozart and masterpiece cuisine are cultivated, the desires for them are more intense than are the desires for mediocre things. Since the desires are going to be satisfied anyway, one gets more out of them, in terms of welfare, the more intense they are. My mediocre desires never light my fire so much, and I'm made worse off for it.<sup>8</sup>

These are perfectly sensible ways in which one can say that one's set of desires is poorly cultivated without having to say that any desire is intrinsically defective. But maybe promoters of the Argument from Poorly Cultivated Desires mean to be saying something stronger: that the satisfaction of an intense desire for Muzak is less good *in itself* for its subject than the satisfaction of an equally intense desire for Mozart. To arguments of this sort I now turn.

## V. Arguments from Intrinsically Defective Desires

Sometimes critics of actualist desire theories have in mind an objection that is different in kind from those so far discussed, despite the fact that we can classify it under the same name. Some critics think that some desires are 'intrinsically defective'. The claim is that, for some desires, a person can be made worse off by having them satisfied, not because the satisfaction leads to a lower net balance of satisfactions, but simply because the satisfaction is bad (or not very good) in itself. The case of base desires is perhaps the most common example.

### A. The Argument from Base Desires

In a colourful passage, G. E. Moore presents an argument from base pleasures in an attempt to refute a hedonist theory of value [1993: 146–7]. But the point can easily be modified to apply to actualist desire-satisfaction theories of welfare. Modified Moore has us imagine a scenario in which people have no

<sup>8</sup>Note that we're not *always* better off cultivating desires for the refined. It's a contingent matter whether some individual would be better off cultivating such desires. If, say, one will never be able to afford fancy dinners and tickets to the symphony, and so will be stuck with Muzak and mashed potatoes no matter what one desires, one will probably be better off keeping the mediocre desires. I thank Maurice Goldsmith for this point.

desires for the refined—no desires for the 'contemplation of beauty' or for 'personal affections'—but in which instead the strongest possible desires are satisfied 'by a perpetual indulgence in bestiality' [1993: 147]. The upshot is that the advocate of an actual desire-satisfaction theory is forced to maintain, absurdly, that 'such a state of things would be heaven indeed, and that all human endeavours should be devoted to its realization' [1993: 147].

Another provocative case (due to Feldman [2002]) provides the basis for a similar argument. Suppose a terrorist hates children and wants nothing more than to see them suffer; suppose he devotes his life to an ongoing terror campaign, and succeeds. None of us is inclined to think the terrorist's life is worth emulating, but an actual desire view must judge it to be a very good one.

An actualist reply begins by examining the cases more closely. In the Moorean example, a series of claims is made about the desire-satisfaction theory. It is claimed that (a) the theory entails that the scenario of perpetual indulgence so described would be very good, and in fact much better than the current state of the world. It is also claimed that (b) according to the desire-satisfaction theory, we should see to it that such a scenario is actualized. Finally, it is suggested, though not explicitly stated, that (c) the people in such a scenario would be very well off according to the desire-satisfaction theory.

It is important to distinguish these claims because, in my view, only one of them—(c)—is true. The desire-satisfaction theory, as a theory of *welfare*, itself implies nothing about the value of complex scenarios containing many welfare subjects. Theories of welfare do have implications for how well things are going for each subject in Moore's indulgence scenario. But it is a step to go from the claim that everyone in some scenario is doing well to the claim that the scenario itself is good. The step must be bridged by a principle about how to compute the value of 'situations', or possible worlds. But such principles move beyond the scope of theories of welfare and into the jurisdiction of full-blown axiologies. Since the desire-satisfaction theory of welfare contains no such principles, claim (a) is false. An argument based on it therefore won't get off the ground.<sup>9</sup>

Likewise, the desire-satisfaction theory of welfare does not imply that we ought to devote all of our efforts to the actualization of the indulgence scenario. Even if we stipulate that the indulgence scenario is the best scenario available to us, a theory of welfare still does not imply that we ought to actualize it. If theories of welfare have any implications about behaviour, they imply only what we *prudentially ought* to do. But since sometimes I prudentially ought to do what I morally ought not to do, and sometimes I morally ought to do what I prudentially ought not to do, theories of welfare on their own imply nothing about what we morally ought to do.<sup>10</sup> So claim (b) above is false, and any argument against an actualist desire-satisfaction theory based on it will be unsound.

<sup>9</sup>Let me emphasize that this kind of rebuttal may be ineffective against the argument of the real G. E. Moore, whose target appears to have been a hedonist theory of *value* rather than a hedonist theory of *welfare*.

<sup>10</sup>Here I am interpreting the word 'should' in the Moorean quotation 'all human endeavours should be devoted to its realization' to be the *moral* 'should'. But even if Moore had some other sense of 'should' in mind, such as the *just-plain* 'should', the rebuttal still stands, since the just-plain 'should' and the prudential 'should' likewise come apart.

Theories of welfare do indeed have implications about how well off a perpetual indulger in bestiality would be. Since it is possible for a person to have strong desires in that direction, and have them satisfied, and satisfied without concomitant frustrations, actualist desire-satisfaction theories of welfare do indeed imply that things can be going very well for a perpetual indulger in bestiality. But I maintain that the actualist desire-satisfaction theories deliver the correct verdict here.

To see why, we need to get very clear on what we're imagining vis-à-vis the thought experiment, and on what we are—and are *not*—saying in claiming that the perpetual indulger has a high level of welfare. For the Argument from Base Desires to work, it would need to establish that the desire (in humans, at least) for sex with animals is *intrinsically* defective. If the claim is merely that such a desire is all-things-considered defective (because, say, the indulger would get bored or lonely, or would come to feel extreme shame or guilt, or would come down with some barnyard diseases), then the actual desire-satisfaction theory will agree. To be sure our judgement is not clouded by such considerations, we should imagine a perpetual indulger who would suffer no such maladies. We need also to imagine a person who has genuine desires to have sex with animals, not a person who finds it disgusting but is forced to do it. So it's probably no good to imagine yourself in the barnyard, with all of your current desires. The subject of our Moorean thought experiment must want and get sex with animals, and without large amounts of ensuing desire frustration.

What are we *not* saying if we agree with the desire-satisfaction theory that the perpetual indulger does have a high level of welfare? We are not saying that it is good *simpliciter* that he is doing what he is doing. The desire theory of welfare says nothing about what is good *simpliciter*, only about what is good *for a subject*. It is coherent to say that every indulger in Moore's 'heaven' is well off even though it is a very sad state of affairs.

Nor are we giving our approval of the indulger's behaviour. It may turn out that what the indulger is doing is morally wrong, and if it is, this is consistent with the desire theory of welfare. Nor are we saying that the life of the perpetual indulger ranks high on the other scales on which we rank lives, such as the scales that measure virtue, dignity, or achievement. In short, it is perfectly consistent for an actual desire-satisfaction theorist to issue the following judgement about the perpetual indulger: 'What a pity! Sure, he's well off there in the barnyard, happy doing his thing, getting just what he wants, but his life is pathetic: he will achieve nothing; what he does is degrading; and his moral character is woefully underdeveloped. I would not wish this life upon anyone'.

Notice that we have pointed to another way in which a desire can be all-things-considered defective. A desire can be all-things-considered defective not because it adversely affects one's net level of welfare, but because it adversely affects one's character, or one's dignity, or what one achieves in life, each of which corresponds to a scale of evaluation of lives that we care about. Actually, it is not obvious that this sort of defectiveness is merely accidental, at least not for all of these alternative scales on which a life can be measured. For example, perhaps there is a necessary connection between

bestiality and low dignity. If so, then we do have a case of an intrinsically defective desire, but not one whose defects contaminate welfare. We therefore need a distinction among types of intrinsically defective desire—there are those that are welfare-defective, virtue-defective, dignity-defective, and achievement-defective. (There may be others for any additional scales on which a life can be ranked.) The lesson is that the Moorean argument must find an intrinsically welfare-defective desire, not merely an intrinsically defective desire or an all-things-considered welfare-defective desire.

We should acknowledge one final factor that might be judgement-clouding: that we find sex with animals disgusting. But that we are viscerally repulsed by some state of affairs does not mean we have good reason to condemn it. Indeed, I think it means we must be very careful not to jump to prejudicial conclusions without giving the matter some thought in a cool moment.

I acknowledge that the disfavour commonly felt towards a life like the perpetual indulger's may survive reflection. But the reason, I submit, that we react strongly against a life of perpetual indulgence in bestiality is that we care about more than just welfare. We don't merely want ourselves (or those we love) to be well off; we also want (or want them) to do good things, to be good people, to achieve worthwhile goals. A life of perpetual indulgence—in bestiality or whatever—lacks the other elements that we care about.

We have teased out what is and what is not implied by the desire-satisfaction theory. The theory implies that the desire for a perpetual indulgence in bestiality is not intrinsically welfare-defective. That is, it implies that the satisfaction of this desire, on its own, augments its subject's welfare. It is consistent with the desire theory that the satisfaction of this desire is bad *simpliciter*, is not in the subject's long term best interest, and diminishes the subject's virtue and dignity. Once all this is pointed out, the Argument from Base Desires loses its thunder. The perpetual indulger is pleased to be with the animals, is getting what he wants, is happy about it all. Given this and everything else that we have said, the claim that the satisfaction of this desire contributes positively to his welfare does not seem implausible.

Let me return briefly to the case of the terrorist, because it has elements that may not be present in the Moorean thought experiment. For instance, just about everyone agrees that the terrorist is a terrible person and that what he is doing is morally reprehensible. A parallel judgement about the perpetual indulger is far more controversial.

The desire-satisfaction theory of welfare implies that things are going well for the terrorist as he delights in the children's suffering. Again, I submit that this is the correct verdict. But let's remind ourselves what this verdict is not. By granting that the terrorist is well-off, the desire-satisfaction theory is not saying that his being well-off is good; it is not saying that we ought to let the terrorist do what he is doing; it is not saying that the terrorist's life ranks high on the other scales. The theory is just saying that his level of welfare is high as he continues to get what he wants. But isn't this the correct result? I think one reason we are so outraged at the situation is that the terrorist *is* well off despite that fact that he such a horrible person. We find the injustice of it all maddening. Any theory that delivered the result that the terrorist

was not doing well could not explain this injustice [Lemos 1994: 43–4; Goldstein 1989: 269, 271].

I conclude that if the Main Argument is interpreted as being about intrinsically defective desires (that is, intrinsically welfare-defective desires), then premise (1) is false: there is no such animal.

### B. The Argument from Pointless Desires

A desire might strike us as intrinsically defective not because its object is base or depraved, but simply because its object is pointless. The most compelling examples of the claim involve entire lives: an afternoon spent tossing a ball against a wall is quaint; a lifetime spent doing so is pathetic. But actualist desire-satisfaction theories will approve of lives spent doing the most pointless things, so long as the person continues to want to be doing the pointless things.

The argument rests on the idea that some states of affairs are intrinsically pointless. What could this mean? It cannot mean that the states of affairs fail to achieve their point (the purpose for which they were brought about), that they are inadequate means to a worthy end. This is ‘instrumental pointlessness’, and desire theories restricted to basic desires ignore the satisfaction of instrumentally pointless desires, since they ignore instrumental desires altogether. Nor does ‘intrinsically pointless’ seem to mean intrinsically bad, as the objects of base desires might be. What the terrorist wants is bad. But what Rawls’s grass counter wants doesn’t seem *bad*; it just seems . . . pointless. Might, then, it mean intrinsically morally neutral? I don’t think so. Many activities are neither morally good nor morally bad without being what most would want to call ‘pointless’. Consider playing the piano, collecting stamps, or climbing rocks.

We can’t say that piano playing, stamp collecting, and rock climbing escape pointlessness because their point is the fun of doing so. Counting the blades of grass and throwing the ball against the wall can have a point in that sense, too. The case in question is of a person who has a *basic* desire to count grass.

Consider yet another hypothesis. These activities are unworthy not because they lack moral goodness (or have moral badness) but because they lack *excellence*. And, it might be claimed, human persons, as beings with dignity and worth, ought to pursue the excellent over the pointless. Although I have some sympathy with this way of thinking, if it is what lies behind the Argument from Pointless Desires, I think the argument fails. For excellence is a scale on which a life can be measured that is distinct from the welfare scale. The grass counter’s life lacks engagement in the excellent, and is all the worse for it, but this is not to say that its well-being is damaged. Indeed, it is likely that we could harm the guy (i.e., diminish his well-being) precisely by refusing to let him pursue his heart’s desire, despite the lack of excellence of the activity.

So my reply to the Argument from Pointless Desires is similar to my reply to the Argument from Base Desires. Each argument posits intrinsically

welfare-defective desires. I deny that there is any such thing. No way of making sense of the notion of pointless desire yields a kind of desire that is intrinsically welfare-defective. I explain our feeling that the lives described in the arguments are lacking by pointing out that the lives rank low on other scales of evaluation that we care about, scales different from the welfare scale.

At this point one might protest: ‘Then what is welfare?’ Am I employing an idiosyncratic concept of welfare, one made so narrow that no objection can touch it? Is my concept of welfare the same one others have been talking about, the one that is relevant to moral theory?<sup>11</sup> I definitely mean it to be, and I believe it is. I do distinguish a life rich in personal welfare from lives rich in virtue, excellence, dignity, achievement, aesthetic value, etc. But this is standard practice; many writers on well-being do precisely the same thing when characterizing their subject matter for their readers [Griffin 1986: 23; Sumner 1996: 20–5; Crisp 2001; Feldman 2004: 8–12]. And the concept isn’t so narrow or unconnected to other concepts as to be irrelevant to moral philosophy, or to be such that theories about it are unfalsifiable.

Welfare as it is understood here is relevant in the sense that it still provides *reasons for action*: that some state of affairs would be good for some agent gives that agent a reason to bring about that state of affairs. Nothing I have said in responding to the problem of defective desires contravenes this undeniable thesis. I accept (and believe it would be a mistake to deny) that Moore’s perpetual indulger, Feldman’s terrorist, and Rawls’s grass counter have reason to do what they do. Now, what I have said may be in tension with the idea that welfare, by itself, provides ‘agent-neutral’ reasons—i.e., that the fact that some state of affairs would be good for some subject provides anyone capable of bringing it about with a reason to bring it about. But this seems to me to be exactly the right result. In my view, value *simpliciter* (or value for the world)—rather than mere value for a person—is what provides agent-neutral reasons. That is, that some state of affairs would be good provides anyone capable of bringing it about with a reason to bring it about. Of course, whether a state of affairs is good will depend in part on the level of well-being of the people in it. But, I take it, it is no simple function. It will depend on other things as well, such as, for example, on the extent to which the people deserve to be enjoying whatever welfare they are enjoying, and perhaps even on the extent to which the welfare being enjoyed has a non-base and non-pointless source.

The actualist theory defended here is definitely not unfalsifiable. In fact, I myself think it is probably false! But this is due to a different objection.<sup>12</sup> My point here has not been to endorse once and for all the particular actualism presented here. My point rather is to show that the problem of defective desires is, after all, no problem for even very simple forms of actualism (and therefore that it is no problem for actualist desire theories generally, including more sophisticated versions designed to avoid the better objections).

<sup>11</sup>I thank anonymous referees for the *Australasian Journal of Philosophy* for pressing this complaint.

<sup>12</sup>What I call the objection from remote desires, the most well-known incarnation of which is Parfit’s case of the stranger on the train [1984: 494]. Griffin [1986: 16–17] and Kagan [1998: 37] also present versions of the objection from remote desires.

### C. The Argument from Poorly Cultivated Desires

Recall me and my twin. I like Muzak; he prefers Mozart. Let's focus on one hour-long stretch of each of our lives. During that hour I sit in the waiting room of a dentist's office relishing the elevator music. My twin is beside me with Mozart's Piano Concerto No. 23 playing through his headphones. Suppose we have equally strong desires to be hearing our respective music types. Suppose all else is equal. Actual desire-satisfaction theories imply that I am just as well off listening to Muzak as my twin is listening to Mozart.

Once an argument from intrinsically defective desire is presented so starkly, doesn't it lose its force? What reason could one give for the claim that I am worse off than my brother simply because I'm listening to worse music? Even granting the point that Mozart is more worthy of desire than Muzak, why should this have any effect on welfare? My brother and I have equally strong and lasting desires for our respective music. We enjoy hearing our respective choices equally. We're equally content sitting there doing it. Why say I'm not as well off? It may be plausible to say that I'm lacking in aesthetic virtue, or even that I degrade myself by listening to such schlock, but this not to say I'm *harmed* by it, welfare-wise.

It is plausible to say that my desire for Muzak may be all-things-considered defective, as we saw in Section IV.D. But we also saw there that the desire-satisfaction theory of welfare accommodates the fact that poorly cultivated desires can be all-things-considered defective. On either interpretation, the Argument from Poorly Cultivated Desires misses the mark.

## VI. The Argument from Desires To Be Badly Off

Recall the case: guilt-ridden by past crimes, I desire to be badly off, and so find an arduous, boring, and insignificant job, which makes me miserable. Surely, if ever there is a desire whose satisfaction is not good for its subject, the desire to be badly off is it.

Actualist desire-satisfaction theorists can reply as follows. We admit that the satisfaction of a (basic) desire to be badly off is itself intrinsically good for its subject. But we also point out that such desires are defective, according to desire-satisfaction theories, though in a way we haven't yet described. Perhaps the best way to describe the way in which the desire to be badly off is defective is to say that it is *necessarily all-things-considered defective*—all-things-considered defective because its defectiveness is due to the frustration it will lead to, but necessarily defective because it will lead to the frustrations of necessity.

The example will help to clarify. Here is how the actual desire-satisfaction theorist should describe what goes on in the case. I have a job that is arduous, boring, and insignificant. My daily life is thus jam-packed with desire frustration: to feel bored is, necessarily, to want (intrinsically) to be doing something other than what one is doing. For a person to find a task arduous is also, necessarily, for him to have certain desires frustrated. Does it make sense to say that someone found some stretch of time boring, or

arduous, but that he was getting everything he wanted at every moment of that stretch of time?

So each day at this job, I'm miserable. My life is filled with desire frustration. But that's not the whole story, for I do have at least one desire satisfaction on my plate: my desire that I be badly off. That's one point in my favour against all the points against me. The satisfaction of this desire to be badly off must, of necessity, count for less, in terms of welfare, than all the daily frustrations I rack up. If it were otherwise, then I wouldn't be badly off, and the desire to be badly off would no longer be satisfied.

The sense in which my desire to be badly off is 'necessarily all-things-considered defective' is this. Necessarily, any time the desire is satisfied, I will be badly off. This makes it necessarily defective, but it is a form of defectiveness that is accommodated by actual desire-satisfaction theories, as described above. Though the defectiveness is necessary, it is merely all-things-considered (rather than intrinsic); the satisfaction of the desire to be badly off is itself intrinsically good for its subject, because any desire satisfaction is intrinsically good for its subject.

### A. The Paradoxes of Desire

I said that the satisfaction of the above desire to be badly off must, of necessity, count for less, in terms of welfare, than all the daily frustrations I rack up. Otherwise, the desire to be badly off wouldn't be satisfied. But what if the desire to be badly off is stronger than the sum of the frustrations, so that I fail to be badly off?

Let's make the case more explicit. Suppose at  $t_1$  I suffer two desire frustrations, each of intensity 3. It would seem that, given the actual desire-satisfaction theory, my net welfare level at  $t_1$  is  $-6$ . But suppose I also desire at  $t_1$ , to intensity 10, that my welfare level at  $t_1$  be negative. If my welfare level is negative, then my intensity 10 desire is satisfied, which tilts the scales back to the positive side, to  $+4$ . But if that happens, then my intensity 10 desire that my net welfare at  $t_1$  be negative is *not* satisfied, tilting the scales back to the negative. We have a paradox. My desire that my welfare be negative is satisfied if and only if it is not satisfied.

There is another, simpler paradox in the vicinity. Imagine a person who has exactly one desire: that he be badly off. Our desire-satisfaction theory implies that this desire is satisfied if and only if it is not satisfied.<sup>13</sup>

But not just desire-satisfaction theorists are mired in paradox. Analogous paradoxes get off the ground without assuming a desire theory of welfare. Imagine a person who desires, to intensity 10, that his net balance of desire satisfaction over frustration at some time be negative. Suppose he gets 6 units of desire frustration at that time. It would seem his balance is  $-6$ . But if it is, then his intensity 10 desire is satisfied, and so his net balance is  $+4$ .

<sup>13</sup>The paradoxes for desire theories of welfare were brought to my attention by Fred Feldman and by Kris McDaniel. Feldman raises the paradox in print in Feldman [2004: 17].



But then his intensity 10 desire is *not* satisfied. In short, this desire is satisfied if and only if it is not satisfied.

A simpler paradox also exists. Imagine a person who has exactly one desire: that none of his desires is satisfied. This desire is satisfied if and only if it is not satisfied. These paradoxes of desire are everyone's paradoxes.

My response to the paradoxes for desire-satisfaction theories is that they are inherited from these more basic paradoxes. The paradoxes for desire-satisfaction theories are just variants of the paradoxes of desire. One possible answer to the paradox (at least to the simpler version) is to reject the possibility of having just one desire. If the concept of desire is analytically tied up with counterfactuals involving belief and action, then it is probably impossible to have just one desire. An alternative response (to the complex as well as the simpler version) is that although it seems for all the world that a person could desire that none of his desires be satisfied, the paradox itself shows that this is impossible (cf. some commonly accepted responses to Russell's paradox and to the liar paradox). It is beyond my scope here to attempt to solve the paradoxes of desire. I will say only that however the more basic paradoxes of desire are solved so will the paradoxes for desire-satisfaction theories be solved.

The paradoxes of desire notwithstanding, the actualist desire-satisfaction theory of welfare stands up well to what is commonly thought to be a devastating objection: the problem of defective desires.<sup>14</sup>

University of Colorado at Boulder

Received: June 2004

## References

- Adams, Robert Merrihew 1999. *Finite and Infinite Goods: A Framework for Ethics*, New York: Oxford University Press.
- Anderson, Elizabeth 1993. *Value in Ethics and Economics*, Cambridge MA: Harvard University Press.
- Brandt, Richard B. 1979. *A Theory of the Good and the Right*, Oxford: Clarendon Press.
- Bykvist, Krister 1998. *Changing Preferences: A Study in Preferentialism*, Ph.D. Dissertation, Uppsala University.
- Carson, Thomas L. 2000. *Value and the Good Life*, Notre Dame: University of Notre Dame Press.
- Crisp, Roger 2003. Well-Being, in *The Stanford Encyclopedia of Philosophy (Summer 2003 Edition)*, ed. E. N. Zalta. URL = <<http://plato.stanford.edu/archives/sum2003/entries/well-being/>>.
- Feldman, Fred 2002. The Good Life: A Defense of Attitudinal Hedonism, *Philosophy and Phenomenological Research* 65: 604–27.
- Feldman, Fred 2004. *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*, Oxford: Clarendon Press, 2004.
- Goldstein, Irwin 1989. Pleasure and Pain: Unconditional, Intrinsic Values, *Philosophy and Phenomenological Research* 50: 255–76.
- Griffin, James 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*, Oxford: Clarendon Press.
- Heathwood, Chris 2004. Desire-Based Theories of Welfare and the Possibility of Self-Sacrifice, presented at the Central Division Meeting of the American Philosophical Association, Chicago IL, April 25, 2004.
- Kagan, Shelly 1998. *Normative Ethics*, Boulder: Westview Press.
- Kraut, Richard 1994. Desire and the Human Good, *Proceedings and Addresses of the American Philosophical Association* 68: 39–54.
- Lemos, Noah M. 1994. *Intrinsic Value: Concept and Warrant*, New York: Cambridge University Press.

- Moore, G. E. 1993 (1903). *Principia Ethica*, rev. edn, ed. Thomas Baldwin, Cambridge: Cambridge University Press.
- 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: Clarendon Press.
- Rabinowicz, Wlodek and Jan Österberg 1996. Value Based on Preferences: On Two Interpretations of Preference Utilitarianism, *Economics and Philosophy* 12: 1–27.
- Rawls, John 1971. *A Theory of Justice*, Cambridge MA: Harvard University Press.
- Schwartz, Thomas 1982. Human Welfare: What It Is Not, in *The Limits of Utilitarianism*, ed. H. B. Miller and W. H. Williams, Minneapolis: University of Minnesota Press: 195–206.
- Sidgwick, Henry 1907. *The Methods of Ethics*, 7th edn, London: Macmillan.
- Sumner, L.W. 1996. *Welfare, Happiness, and Ethics*, Oxford: Clarendon Press.

<sup>14</sup>Thanks to Ty Barnes, Fred Feldman, Maurice Goldsmith, Kristen Hine, Kris McDaniel, Jason Raibley, Jean-Paul Vessel, and two anonymous referees for the *Australasian Journal of Philosophy* for helpful comments on an earlier version of the paper. I am particularly indebted to Fred Feldman for extensive assistance and continued encouragement.