

Well-Being and Preferences

Abstract

The notion of well-being is an important one in normative ethics, political philosophy, economics, and public policy. In this paper, I defend an informed-preference account according to which one's well-being consists in the satisfaction of those preferences one would have if fully informed and rational. While I am not the first to entertain such an account, I try to fill it out in a convincing way and thus to eliminate some of the confusion and controversy that continues to surround it. On my version of the account: (a) one counts as fully informed when one knows relevant concepts, logical principles, and facts and likelihoods regarding what life would be like in each of the multiple futures open to one, where these things are vividly imagined and considered in a cool hour; (b) rationality is interpreted as consistency among one's fully-informed preferences; and (c) a preference counts as satisfied when what is preferred actually comes about, not when it is merely believed to have done so, although an individual may sometimes prefer not to know what has actually happened. Against various critics, I argue (d) that morally-motivated or other-regarding preferences should not be excluded, and (e) that the exogenous and adaptive nature of preferences does not pose an insuperable problem for the account.

Well-Being and Preferences*

I. Introduction

Is an afternoon at a philosophy lecture or one spent on a hike in the woods more conducive to a person's well-being? Is the well-being of a 35-year-old better promoted through a life of marriage and children or through swinging, childless independence? Would the well-being of a rural Indonesian be better enhanced by getting a water-sealed toilet or satellite television in her home? Most economists claim that each individual is the best judge of her own well-being, and so to answer these questions we should consult the preferences of the people involved in them.¹ This claim plays a fundamental role in an analytic tool typically favored by economists, namely cost-benefit analysis. I believe it is unconvincing without a better understanding of the notion of well-being.

In this paper, I defend a preference-based account of well-being, which at first glance seems congenial to the economists, since it maintains a connection between preference satisfaction and well-being. But in order to be plausible, this must be an informed-preference account, according to which one's well-being consists in the satisfaction of those preferences one would have if fully informed and rational.² While I am not the first to entertain such an account, I try to fill it out in a convincing way and thus to eliminate some of the serious confusion and controversy that continues to surround it.³

II: What is a Preference?

What is a preference? The notion is not quite a primitive, but what can be said as a characterization of it is limited. Behaviorist understandings of preference – including the one long favored by economists, whereby one has a preference for X over Y if and only if one is observed to buy X when one could have within one's budget bought Y – have rightly lost favor.⁴ Nonetheless, the notion should be understood to have *some* behavioral component. A claim to having a preference

for, say, doing philosophy will be dubious if the person making it never actually reads or writes any. The claim lacks all credibility if the person is never even disposed to read or write any. Furthermore, preferences are distinct from (on the one hand) desires and wants, and (on the other hand) beliefs and judgments – at least as those are crudely understood. Desires are sometimes thought to be mere urges or inclinations not connected in any way to reason, which includes being immune from criticism by reason. Many others who understand desires as more substantial than this still maintain that they differ in kind from, and are at most mediated by, beliefs or judgments. Beliefs and judgments, on the other hand, are taken to subsist in one's rational faculty, but to be motivationally impotent.⁵ Preferences as I see them have characteristics of both desires and beliefs.⁶ A preference for going for a walk over watching television, for example, can consist at least partly in belief that walking is healthier than watching television and that *ceteris paribus* it is good to do what is healthy. But it consists in more than beliefs. Having a preference for a thing involves being in some sense attracted to that thing, provided one does not hate, resist, or feel hostage to, this attraction. Rather, one must endorse (embrace, assent to) the attraction.⁷ Preferences so conceived have motivational power; we do not need to appeal to some other psychological notion in order to understand how someone could be moved to do/pursue that which she prefers.⁸

III. The Informed-Preference Account of Well-Being

It would undoubtedly be a mistake simply to identify well-being and preference satisfaction. Some people consistently prefer artery-clogging pork rinds to fresh fruit. Others prefer smoking two packs of cigarettes a day, or gambling under terrible odds with what little money they have, or wrestling alligators, or even playing Russian Roulette. However, it is hardly a conceptual truth, and indeed far from obviously true at all, that such an individual's well-being really is increased by doing these things. If you want to promote someone's well-being, you will probably not assist her in playing Russian Roulette, even if you know she prefers to play it. Maintaining a conceptual distinction between well-being and preference satisfaction is required in order to make sense of

patently reasonable statements like: “I did not get what I preferred, but I am better off as a result.” More sophisticated and plausible is what I will call an informed-preference account.

(a) *Fully-Informed Preferences:*

We begin with the intuitive idea that one’s well-being *is* connected to the preferences one has, but then quickly note that the two can pull apart when – as is often the case – we operate in conditions of incomplete information. In such conditions, the satisfaction of an individual’s preference does not necessarily increase her well-being; in fact, it may diminish her well-being. Consider the following policy-relevant example: person *P* prefers drinking relatively cheap well water to paying for and drinking more expensive bottled water, but knows neither that the well water contains considerably more arsenic than the bottled water, nor that arsenic is a potent carcinogen. Again, one view is that, given *P*’s actual preferences, her well-being is better improved by drinking well water than by drinking bottled water (given their relative costs). More plausible, however, is a view that continues to link *P*’s well-being to her preferences, but to the preferences she *would have* if she had the aforementioned information regarding arsenic. If it is the case that *P*, were she aware of the arsenic, would prefer the bottled water (even given its relative costliness), then it is reasonable to say that the bottled water better contributes to her well-being than does the well water. The informed-preference account does not suggest that the satisfaction of an uninformed preference will in practice never increase one’s well-being; but it does maintain that the more information one has, the more plausible it is to say that the satisfaction of her preferences promotes her well-being. Well-being does not *consist in* the satisfaction of uninformed preferences, but rather (at least partly) in the satisfaction of fully-informed preferences.⁹

Despite what the arsenic example may suggest, the very notion of full information is not simple. Roughly, in order to have full information, one needs to have a clear picture of what life would be like – for oneself and for others – in each of the multiple futures open to one in accordance with the different choices one could make. But if well-being is to be defined in terms of the

satisfaction of fully-informed preferences, and having full information entails knowing one's level of well-being in the various states over which her preferences range, then the account is circular. Thus, being fully informed must not include being fully informed directly about one's well-being. Full information entails knowledge of facts, concepts, and logic. But it need not include more than those facts, concepts, and logical principles relevant in the particular context at hand. Information is relevant if knowledge of it might reasonably influence one's preferences. Most of us do not need to know the color of the seventh moon of Neptune in order to be possessed of complete information as the term is intended. In typical cases, though, a variety of facts within the purview of the physical, life, and social sciences will be relevant, as will certain facts about the past.

What facts about the future are required for full information is more complicated. On the one hand, full information is meant to entail more than hopes or reasonable expectations regarding what the various futures would be like. It is meant to preclude the possibility that someone could get what she with full information preferred but be worse off nonetheless because what she got was not all she expected it to be. But full information is not meant to include so much about the future that one always knows which of the various states over which her preferences range actually will obtain. This means that one should not know, e.g., whether she will actually end up drinking the well water or the bottled water. More interestingly, it also means that one need not know whether the arsenic would, if she drank the well water, actually end up giving her cancer, but she must know the relevant likelihoods. Someone with full information will also know how likely it is for various preferences, if satisfied, to make her contented feeling or happy, and what new preferences would be generated by the satisfaction of her different preferences.

An objection will be raised here. Suppose that P is offered a gamble the odds of which she knows to be in her favor: if a fair coin comes up heads either once or twice in two flips, she get \$50, but if it comes up heads neither time, she pays \$50. She does not know what the outcome will be, but she goes for the gamble. Alas, both flips land tails-up. This is a case, the objection goes, where

she got what she with full information preferred but is nonetheless worse off. I believe the objection fails. Insofar as she both preferred with full-information to take part in the gamble and was able to do so, she is better off, *ceteris paribus*. The intuition that she is worse off as a result is explained by the fact that she presumably also had an informed preference for *winning* the gamble, and this preference was frustrated. If the preference for winning was particularly strong, *P* might indeed, in virtue of getting to take part in the gamble but losing, count as worse off overall; but this is explained precisely by the frustration of this strong (informed) preference.

More must be done to fill out the informed-preference account's epistemic condition. Sidgwick suggests that "a man's future good on the whole is what he would now desire and seek on the whole if all the consequences of all the different lines of conduct open to him were accurately foreseen *and adequately realized in imagination at the present time.*"¹⁰ Adequate realization in the imagination is an important condition, and can be understood to have two parts. First, in order for *X* to better contribute to *P*'s well-being than does *Y*, *P* must not just prefer it in light of an abstract possession of the requisite knowledge of *X* and *Y*, but prefer it with this knowledge vividly portrayed in her mind. It is one thing to prefer having a cigarette when one knows but is not really thinking about the fact that cigarettes cause cancer; it is another thing to prefer a cigarette even when one vividly imagines what it would be like to have one's body wrecked by cigarette-induced cancer. Second, *P* must prefer *X* over *Y* in a "cool hour,"¹¹ which is to say that *P* must not be, as it were, in the grip of a passion. One must not, for example, be clouded by rage or raw animal lust. Now, it would clearly be too strong to say that one must be *entirely* unemotional or dispassionate. Indeed, such a person would probably not even count as having preferences; and in any case, well-being at least sometimes consists in getting/doing that which one passionately prefers. But the idea is simply that one's emotions must not be so strong as to entirely drown out one's judgments. Henceforth, I will have all of this in mind when I refer to the condition of full information.

(b) Rational Preferences:

The conditions that a compelling preference-based account of well-being will put on preferences do not stop here. Assume that someone gambles every day, genuinely prefers to do so, and prefers to even as she is fully informed (about, e.g., her tendencies, her financial situation, other ways her family could use this money, the poor odds involved, etc.). Still, it would be a mistake to conclude too quickly that this means that this individual's well-being is enhanced when she gambles. Even someone who has full information can have a preference the satisfaction of which would diminish rather than enhance her well-being. In short, this can happen if the preference is irrational.

By making contact with the notion of rationality, the informed-preference account risks trading one opaque normative notion, namely well-being, for another. In fact, the understanding of rationality that I recommend here is relatively simple. It is meant to be as thin as possible. It is just consistency among one's preferences. (If the full-information condition is satisfied, consistency among one's *beliefs* will be guaranteed.) One is not irrational simply in virtue of an inconsistency between an uninformed preference and a fully-informed one. On the other hand, if the actual preferences of someone who lacks relevant information are inconsistent with one another, she is irrational. Here, however, it is not just the irrationality but also the ill-informed nature of these preferences that keeps the individual's well-being from consisting in the satisfaction of these preferences. Most notable for the informed-preference account is irrationality in the form of inconsistency among one's fully informed preferences. One's preference, in order to be such that its satisfaction increases one's well-being, must be fully informed and also consistent with all of one's other fully-informed preferences. This includes one's long-term preferences, which it may be more natural to call the long-term *goals* one has set for oneself. One implication is that it is possible to act rationally against one's well-being. This occurs, e.g., when one pursues the satisfaction of an actual preference consistent with one's other actual preferences, but inconsistent with those she would have if given full information.

Appealing to only a thin notion of rationality is contentious. Some will insist, for example, that the preference of an anorexic individual to refrain from eating is necessarily irrational. She is hurting herself, the claim goes, and this is necessarily or categorically irrational; it is irrational regardless of what her other preferences are. Other candidates for categorically irrational preferences include those for terminating one's own life, as well as those for an outcome that yields less expected value than another that one knows is open to her. On my view, whether an individual with one of these preferences is irrational cannot be established without looking at her other preferences. Suppose the anorexic knows all about the significant dangers anorexia presents to her health, knows that she is already severely underweight, that many people care for her independent of how thin she is, and even that she is thinner than most people consider attractive.¹² Imagine that still she prefers to refrain from eating. She is irrational only if this preference conflicts with some other informed preference (such as for living a long and healthy life, for enjoying the taste and nourishment of delicious food, and so on). No thicker version of rationality should be appealed to; indeed, I have yet to see a thicker version that makes sense.

Suppose that the anorexic's informed preferences are inconsistent in that she prefers to be healthy but also prefers to be extremely skinny, and that such skinniness is incompatible with good health (and she knows this). Noting this inconsistency does not settle the issue of what her well-being consists in. She could remedy the irrationality by giving up on the preference for thinness, or the one for good health.¹³ Appealing to her second-order preferences – that is, her preferences regarding her preferences – is not the right move. People can have not just second-order preferences, but third-order preferences as well, and any reason that compels the move to the second-order will compel a move to the third-order, and then beyond that even. It is *ad hoc* to go second-order and no further.¹⁴ Instead, I advocate moving to longer-term and better-considered preferences. The relevant questions here should be: To which of her preferences has she given more, and more careful, consideration? Which has more claim to being counted as a long-term goal? Which is more central

to her considered conception of herself and her life plan?¹⁵ It is possible, however, that even at this level one's preferences may be inconsistent. If that is the case, then the inconsistency should be pointed out to her, and she should be pressed to resolve it. If she is unwilling or unable, then I have to admit that her well-being is, like a fraction whose denominator is zero, simply undefined. There is nothing for someone who wants to promote her well-being to promote.

(c) *On the Satisfaction of Preferences:*

The account defended thus far maintains that one's well-being consists in the satisfaction of one's fully informed and rational preferences. An additional question for a defender of an informed-preference account of well-being concerns how 'satisfaction' is to be understood. When does a preference count as satisfied, and when does one count as frustrated? Is it when the condition that one prefers actually obtains, or when one simply believes it obtains? In James Griffin's words, the question is whether satisfaction is a state of the world or a state of mind. Griffin quotes Nozick, who asks, "What else can matter to us other than how our lives feel from the inside?" Griffin says that "[Nozick] replies, surely rightly, that we also want to *do* certain things, to *be* certain things."¹⁶ While Griffin talks about desire fulfillment, he means by this essentially what I mean by preference satisfaction; and he maintains that "being fulfilled cannot be understood in a psychological way." He says, "A desire is 'fulfilled' in the sense in which a clause in a contract is fulfilled: namely, what was agreed (desired) comes about."¹⁷

I agree with Griffin, but the issue deserves more elaboration than he gives. The elaboration is needed to combat the appearance of unacceptable implications of this understanding of satisfaction (i.e., of being fulfilled). Imagine two cases. In both, *P*'s rational and fully-informed preference is that her spouse, *S*, not cheat on her, and in both, *S* is indeed faithful. In case 1 *P* rightly believes that *S* has not cheated, while in case 2 *P* mistakenly believes that *S* has. The question is: are 1 and 2 both cases where *P*'s preference that *S* be faithful has been satisfied? According to Griffin's interpretation of satisfaction, they are, and I agree. But it may initially appear that Griffin would

thus need to say one of two things, both of which are implausible. One, he might say that because *P*'s preference has been satisfied in both cases, *P* is no worse off in case 2 than in case 1. This move just seems implausible; *P*'s belief, mistaken though it is, should not be irrelevant to her well-being. *P*'s well-being is diminished by her coming to believe that *S* has cheated. Alternatively, Griffin could say that *P*'s level of well-being is lower in case 2 than case 1, and could explain this by jettisoning the connection between well-being and preferences. That *P* is less well off in 2 has nothing to do with the fact that her (or anyone's) preferences were not satisfied. From my perspective, this is obviously a dire move. It amounts to abandoning my account of well-being. In fact, neither of these moves is necessary. The fact that *P*'s preference that *S* be faithful was satisfied does not mean that case 2 does not involve the frustration of some of *P*'s other preferences that were in fact satisfied in case 1. We can suppose, reasonably enough, that in addition to having a preference that *S* not cheat, *P* had a preference for not *thinking* *S* was cheating. Additionally, *P* presumably had a preference not to be unhappy that was frustrated when she came to believe that *S* cheated. The fact that these two preferences were satisfied in case 1 but not in case 2 suffices to explain why *P* is worse off in case 2.

The story will now be modified so as to generate two new cases, 3 and 4. In both of these, *P* continues to have a rational and fully-informed preference for *S* to be faithful to her, but now *S* in fact cheats on *P*. In case 3, *P* rightly believes that *S* has cheated, while in case 4 *P* mistakenly believes that *S* has been faithful. Is *P*'s preference frustrated in both? Griffin will say yes, and I agree. But does that mean that *P*'s well-being is identical in these two cases? No. It depends on what *P*'s other preferences are. For example, would *P* prefer to know *S* has cheated on her, or to continue believing *S* is faithful? Suppose that even if *P* knew *S*'s behavior towards her will be no different than if *S* were not cheating,¹⁸ *P* would prefer to know that *S* has cheated. If so, *P* is worse off in case 4, because there *P*'s preference for knowing of *S*'s cheating has been frustrated. But nothing requires this to be *P*'s rational and fully-informed preference. Her informed preferences

would not have been inconsistent if she had preferred to remain ignorant of *S*'s cheating. She might have been someone who, as it were, finds ignorance bliss. For that matter, she might prefer knowing the truth here, but to remain ignorant on another occasion. In any case, the basic point is that Nozick's contention is too strong: most of us do prefer "to *do* certain things, and not just have the experience of doing them," to "*be* a certain way, to be a certain sort of person," and to make "*actual* contact with... reality," but not everyone does or must; and the well-being of those who are fully-informed and rational but do not have such preferences can consist in being deluded.¹⁹

The objection may be raised that if one can prefer not to be fully informed – as I have supposed is possible, though surely not required, for *P* – then it was a mistake for the account of well-being to have maintained up front that well-being consists in the satisfaction of informed preferences.²⁰ If the view is that fully-informed preferences are more important than uninformed preferences for fixing well-being because people prefer to be fully informed, then there is indeed a problem, for at least two reasons. First, as I have noted, not everyone does (or needs to) have this preference. Second, it would be circular to say fully-informed preferences are important because they are the kinds of preferences people prefer to have. The objection is an important one, and not one to which there is a short compelling response. A full response would involve discussing each competitor to a preference-based account, and showing the severe flaws with each. This would leave us looking at an account that connects well-being and preferences, where we would again confront the considerable shortcomings of an actual-preference account. We want an account that acknowledges without overstating the important role that information plays in linking preferences and well-being. This is achieved, precisely as I have done, by putting a structural constraint on preferences without constraining their content. One must be fully informed for one's preferences to count, but the content of one's fully informed preference may be to lack certain pieces of information.

An additional objection is that if satisfaction is understood in something other than a psychological sense, certain of *P*'s future-oriented preferences – like one for her grandchildren to grow old in good health – could be satisfied or frustrated after *P* has died. This means, implausibly, that *P*'s well-being can be affected after she has died. Several responses can be given. Firstly, I see nothing unacceptable about the stipulation that one's well-being consists in the satisfaction of one's well-informed and rational preferences, up to but not beyond the time when one dies. Secondly, though, the idea that one's well-being can be affected after one has died is not so obviously wrong. I do not have a strong view here, but the idea does not strike me as crazy. Even if this is the case, however, one's well-being would not be liable to change in all the ways it could when one was alive. For example, any preference *P* has for knowing something will not be one that could be satisfied after she is dead.²¹

(d) *Morally-Motivated and Other-Regarding Preferences:*

It is often supposed that a preference-based account of well-being needs to restrict the relevant preferences to those not tied to a sense of moral obligation to prefer/do one thing rather than another. According to J. S. Mill, "Of two pleasures, if there be one to which all or almost all who have experience of both give a decided preference, irrespective of any feeling of moral obligation to prefer it, that is the more desirable pleasure."²² David Sobel claims that "the root idea behind this test is that if a person intrinsically prefers X to Y, independent of moral considerations, while fully acquainted with both options and in the 'cool hour,' then X is more conducive to the agent's well-being than Y no matter what other properties X and Y have."²³

I believe it is a mistake to add this condition to an informed-preference account. Suppose that *P* is rational and equally well-acquainted with stealing and not stealing from others: *P* has done both, knows vividly how each feels, knows the punishments for stealing and the chances of getting caught, knows the harm she has caused to those from who she has stolen, and so on. Imagine that at this stage in *P*'s life, she prefers not to steal, and this preference is based on a belief that it is morally

wrong to steal. Absent this belief, she would (at least sometimes) prefer to steal (meaning that she sometimes enjoys the feeling of doing so, is not worried about getting caught; but given the belief, she prefers not stealing. Some want to say that on those occasions where *P* would prefer stealing over not stealing if not for the feeling of moral obligation, her well-being would be better promoted by her stealing. What should be said about this case? Perhaps there is a way of filling in the details so as to make it most plausible to say that *P* does *not* actually prefer not stealing. Since preferences are not understood in an entirely behaviorist way, this is possible even if *P* actually never does steal. But if *P* is fully informed and rational, then so long as she counts as preferring not to steal – and the fact that she feels morally obligated not to steal does not preclude this – it should be conceded that her well-being consists in not stealing.²⁴

If we were trying to give a preference-based account of moral rightness, the requirement of independence from a feeling of moral obligation to prefer *would* be needed. We would want to say that if what is preferred is preferred due to moral considerations, i.e., feelings of moral obligation, then the fact that it is preferred (even where the other conditions are satisfied) is not a reliable guide to its being morally right. But I am aiming for an account of individual well-being, not moral rightness, and so appealing to preferences that are not independent of a feeling of moral obligation is not circular. If this condition *is* included in an account of well-being, we flatly rule out the possibility that *P*'s well-being consists at least partly in her sense of herself as a morally decent person, and in doing what she believes to be morally obligatory. It negates the possibility that *P*'s well-being consists even partly. To do so would be a mistake. To require that preferences be independent of feelings of moral obligation is to build too much into the informed-preference account of well-being.

This is the same basic response that should be given to those who build into an informed-preference account the requirement that preferences be self-interested.²⁵ But here my response is even stronger. A self-interested preference seems to be based only on considerations of one's own

well-being. But if that is so, those who define the notion of well-being in terms of satisfaction of self-interested preferences invoke in their definition of well-being the very notion that they are trying to define. Nonetheless, there is a reasonable concern that can underpin the move to restrict preferences only to self-interested ones. The worry is that if the preferences whose satisfaction well-being consists in are not so restricted, it will be logically impossible for someone who is rational and fully-informed to prefer to act in a way that diminishes her well-being. This seems to render altruism impossible, since altruism is reasonably understood as voluntary behavior sacrificing at least some of one's own well-being for the sake of the well-being of another.

My account does indeed deny that one can rationally and with full information prefer to sacrifice (even a little of) one's well-being. This does not mean, however, that it denies the possibility of altruism. Why not? On the view I have urged, well-being is not identical to any actual mental or psychological state. One's well-being can be increased without one actually feeling better, and without one actually counting oneself as better off. It is possible, in fact, for one to count as generally well-off without ever experiencing any sensation of satisfaction. This might be the case if none of one's actual preferences is satisfied, but what one actually prefers is always opposite of what one would prefer if one were rational and fully informed. Well-being is distinct from happiness, which, whatever it is exactly, surely must involve some actual psychological good feeling. An individual can feel great, sincerely claim to be contented, frequently and genuinely laugh, etc. – in short, be happy – but unknowingly have a terminal disease (something she surely would rationally and with full information prefer not to have). Though the disease-stricken individual will presumably eventually be unhappy, well-being is also distinct from long-term happiness. I suspect that many people are constituted such that the same things that promote their well-being will promote their long-term happiness; but nothing guarantees that one could not rationally and with full information prefer to commit oneself over the course of one's life to projects that one deems to be worthwhile but that will not make one happy.

Notice that preferring to sacrifice one's happiness for the sake of another – either on a certain occasion or over the course of one's entire life – is a fine description of altruism. It is because my account permits such preferences that the account allows that rational and fully informed people can be altruistic. Once the possibility of altruism is saved, I see no grounds for taking the fact that my account precludes the possibility of a rational and fully-informed preference for sacrificing one's well-being to be a reason for objecting to the account.

(e) *The Exogenous and Adaptive Nature of Preferences:*

An additional objection to an informed-preference account is that the exogenous and adaptive nature of preferences makes them ill-suited to play such a key role in an account of well-being.²⁶ Why should we think that one's well-being consists in what one would rationally and with full information prefer when it is the case that one's preferences are shaped, sometimes very significantly, by a whole variety of external forces, including advertising, social norms and pressures, one's upbringing, and possibly even factors like a desire to impress others (including others who are cruel, vain, excessively materialistic, etc.)? Furthermore, we know that people's preferences adapt to their circumstances. If, for example, you are told repeatedly and from an early age that you are not worthy of, and have no chance to possess, political power, then you may well come to prefer not to have political power. As Rousseau famously says, "Slaves, in their bondage, lose everything, even the desire to be free. They love their servitude."²⁷ This adaptation of preferences is intelligible – it is a critical coping mechanism – but surely the well-being of such an individual does not consist in being a slave.

These are serious challenges to an informed preference account of well-being, though not so powerful as to prove that this sort of account must be abandoned. At the risk of trivializing the issue, consider my extreme love for Bob Dylan's music. It is a bit embarrassing to admit to having collected and listened to recordings of over 1000 complete concerts of his. A wide variety of factors have worked to influence my preference for listening to Dylan.²⁸ A long (and true) story could be

told about how I came to have this preference: it might mention what my parents listened to when I was young, the musicians to whom my friends exposed me, the English classes I took that fostered in me an appreciation for kinds of poetry to which Dylan's lyrics are akin, and so on. The story would show that my preference for listening to Dylan is exogenous, but it would not make it any less plausible to say that my well-being is increased when my preference for listening to Dylan is satisfied. To the contrary, telling this story would actually help *explain* why my well-being is increased when I get to listen to Dylan. My wife was not a big Dylan fan when she met me, but spending time with me meant spending time listening to Dylan. She could have avoided Dylan by spending less time with me, or she could have spent time with me but let the music make her miserable. What actually happened was that as I continued to bombard her with Dylan, her preferences adapted such that she too came to enjoy and prefer this music. That her preferences adapted does not prevent us from saying that her well-being is now enhanced when her preference for listening to Dylan is satisfied.

These stories show that it would be a mistake to require that a preference not be exogenous or adapted in order for an individual's well-being to consist in the satisfaction of that preference. Still, serious concerns remain. What should be said, for example, about the person whose preferences were significantly influenced by her parents vigorously pushing her to play the piano, or be a gymnast, or be a beauty queen? What about the well-being of an individual who through abuse has developed extremely low self-esteem and as a result prefers to stay in a relationship where she is not respected, or worse? What about a person who develops certain preferences through brainwashing or hypnosis? One way of handling these cases is simply to add to the informed-preference account the requirement that preferences not arise from hard pushing by parents, abuse, low self-esteem, brainwashing, or hypnosis. But simply compiling a list of disqualifying influences on preferences is unacceptably *ad hoc*. A more elegant move would be to identify what it is that these putatively problematic influences have in common, and to explain why this common element is

problematic. One such attempt would be to say that they are cases where the preferences are not autonomous. The worry is that as autonomy is most intuitively understood, neither my own nor my wife's preferences for Dylan are autonomous. Indeed, 'autonomous' and 'not exogenous' are essentially synonyms. I see no good way of filling out autonomy – or any other condition, for that matter – so as to *exclude* the preferences of the brainwashed, hypnotized, and abused individuals, but to *include* my own and my wife's preferences to listen to Dylan.

This is fine. It is not the exogenous and adaptive nature of the preferences involved in the problematic cases that makes the cases problematic, but rather two other features of the cases. The first is that we are unsure whether the individuals in question – those who were pushed extremely hard by others, or were abused, brainwashed, hypnotized, etc. – really count as having the preferences that their behavior seems to suggest they have. For a preference to be such that its satisfaction increases *P*'s well-being, it must genuinely count as *P*'s preference, not the parents', the abuser's, brainwasher's, hypnotist's, or anyone else's. And it must of course also be a *preference*. It must be more than a mindless pull, must be constituted by particular beliefs, and must involve an attraction that one endorses. If we were sure the individuals in the problem cases really had preferences, the grounds for concern would be diminished.

The second feature of the problem cases that makes them troublesome is that the individuals involved seem not to have, as it were, a clear view of the requisite information. A particularly important resource of the informed-preference account is that someone who satisfies its conditions will have full information about her reasons for preferring what she does (where these reasons could be understood in psychological as well as causal terms). Insight into one's reasons for preferring *X* to *Y* may alter one's preference for *X* over *Y*. If the abused individual fulfills the informed-preference account's conditions, she will know (among other things) that she was abused, that she now has low self-esteem, that and how the abuse contributes to her low self-esteem, and that and how her low self-esteem contributes to her staying in this relationship. It is possible, though not

guaranteed, that were she to know this she would prefer not to stay in the relationship. If so, then her well-being is diminished by staying in the relationship. The hypnotized person will know that she was hypnotized, will know what was said to her under hypnosis, and will know how the hypnosis modified her preferences. I believe that it is the fact that these things are typically unknown that makes it so uncomfortable to have preferences generated under hypnosis, through past abuse, etc., play a constitutive role in an account of well-being.²⁹

NOTES:

* My considerable gratitude is extended to Douglas MacLean, Richard Andrews, Thomas Hill Jr., Marc Lange, C. D. C. Reeve, and Susan Wolf for generous and helpful feedback on earlier drafts of this paper. Their criticisms and suggestions helped to improve the paper tremendously.

¹ See, for example, Edith Stokey and Richard Zeckhauser, *Primer for Policy Analysis*, (New York: W.W. Norton & Co., 1978), esp. p. 257-263. The claim is also made by A. Myrick Freeman, "The Ethical Basis of the Economic View of the Environment," as reprinted in *The Environmental Ethics & Policy Book*, 2nd Ed., Donald VanDeVeer and Christine Pierce (eds.), (Belmont, CA: Wadsworth Publishing, 1998), pp. 293-301 (p. 294). See also "Economics, Incentives, and Environmental Regulation," in *Environmental Policy*, 4th Ed., Norman J. Vig and Michael E. Kraft (eds.), (Washington, D.C.: CQ Press, 2000), pp. 190-209. We find the claim again in Richard Layard and Stephen Glaister, *Cost-Benefit Analysis*, (Cambridge, UK: Cambridge University Press, 1994), p. 2.

² Some people refer to this general kind of account of well-being as a full-information account. Others who have the same basic approach in mind say informed-*desire* account. I explain later in the paper why I eschew 'desire' in favor of 'preference.'

³ Some people talk as if everyone basically agrees that an informed-preference (or "informed-desire") account is the right kind of account of well-being, and that it is just the finest of details of this account that need working out. Would that that were the case. Those objecting to such an account include scholars as different from one another in terms of their theoretical orientation to morality as T. M. Scanlon and Richard Brandt. Make no mistake: that this is the right kind of account has by no means been clearly established.

⁴ It seems to me that the behaviorist conception has lost favor even among economists, though this is just a sense I have. (I have not, e.g., taken a poll.) For a seminal treatment of preferences in economics, see Paul Samuelson, "Consumption Theories in Terms of Revealed Preference," *Economica*, Vol. 15 (1948), pp. 243-253. For a discussion of problems with Samuelson's approach, see Amartya Sen, "Behaviour and the Concept of a Preference," in *Choice, Welfare, and Measurement* (Oxford, UK: Basil Blackwell Publisher, 1982), pp. 54-73. I think that particularly forthcoming economists who currently work with the notion of a preference will admit that it is a somewhat fuzzy concept, but clear enough to be passable. I heard one well-known economist say of preferences – only partly in jest, I suspect – "You know them when you see them."

⁵ I claim neither that everyone subscribes to this sharp distinction between desires and beliefs/judgments, nor that it is a convincing distinction; but it is clearly influential.

⁶ It should be noted that there is probably an understanding of desire that is less crude – and more specifically, that is more cognitivist – than the one I have described, and that could thus be treated as synonymous with preference. Even so, I would use 'preference' rather than 'desire' in order to keep contact with the terminology that proponents of cost-benefit analysis (CBA) favor, since CBA is at the center of a great deal of my work.

⁷ That preferences consist partly in beliefs, and that they involve attraction that one endorses, together provide the basis for a distinction between preferring and being physically addicted.

⁸ It may be asked here whether I take the notion of a preference to be essentially comparative. I frankly do not have a strong view. I am comfortable talking about a preference for *X simpliciter*, i.e., without invoking some alternative or comparative thing *Y*; but typically there are contextual cues that indicate what less-favored alternatives exist, and if pressed, I could make these explicit. Talking simply about a preference for *X* – as opposed to a preference for *X* over *Y* – is often less cumbersome.

⁹ I say "at least partly" because this is just a preliminary articulation of the informed-preference account of well-being, not the final version. Several modifications are yet to come.

¹⁰ Henry Sidgwick, *The Methods of Ethics*, 7th Ed., (Indianapolis: Hackett Publishing, 1981), pp. 111-112. Emphasis added.

¹¹ The notion of a "cool hour" is probably best known from "Sermon XI" of Bishop Joseph Butler's famous *15 Sermons*.

¹² Notice that I do not say that she knows that she is so thin as to be unattractive, since I do not want to commit myself to the view that there is a fact of the matter about this. There may be, but I do not need to contentiously suppose so here. How attractive others find someone is clearly inbounds, since it is the kind of question that sociologists, psychologists, neuroscientists, etc., can fairly easily test.

¹³ This is on the assumption, of course, that giving up either of these would not create other inconsistencies.

¹⁴ It should also be noted that just as her first-order preferences conflict, so too can her second-order preferences conflict. She could prefer to prefer to be so skinny and also prefer to prefer to have good

health.

¹⁵ This is the same basic move I would make to deal with the problem posed by preferences that change with time. See, for example, James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance*, (Oxford, UK: Clarendon Press, 1986), p. 16. Griffin's example is an individual who "through much of his life... wanted his friends to keep him from vegetating when he retired but, now that he is retired, wants to be left to vegetate." Being sedentary is not in his interest the moment he develops the preference for being sedentary; but nor should we insist the fact that he long preferred to be active in his retirement means that just because he has reached retirement being active *must* continue to be in his interest. Being sedentary can be in his interest so long as his preference for being sedentary is based in complete information and is consistent with whatever well-considered, long-term goals he maintains for himself and his life in retirement.

¹⁶ Griffin, *ibid*, p. 9.

¹⁷ Griffin, *ibid*, p. 14.

¹⁸ This means, for example, that *S* is as affectionate, as interested in hearing about *P*'s day, as willing to help around the house, etc., as *S* would be if *S* were not cheating

¹⁹ The Nozick quotes come from *Anarchy, State, and Utopia*, (New York: Basic Books, 1974), pp. 42-45.

²⁰ Thanks to Marc Lange for pressing me to think harder about this issue.

²¹ This is true, of course, unless one invokes certain metaphysical claims about the possibility of *P* existing in another life after this one.

²² Sobel, "Full Information Accounts of Well-Being," *Ethics*, Vol. 104, No. 4 (July 1994), pp. 784-810 (p. 790).

²³ Sobel, *ibid*. I actually understand Mill, in the aforementioned passage, to be proposing a test for distinguishing higher from lower pleasures, which is not the same as giving an account of well-being. Still, some find this to be a reasonable condition to build in to an informed-preference account of well-being. See, e.g., Richard Brandt, *A Theory of the Good and the Right*, (Amherst, NY: Prometheus Books, 1998), especially Ch. XIII, "Welfare: The Concept, Measurement, and Interpersonal Comparisons."

²⁴ If there are moral facts, one who has full information would then necessarily need to know at least the relevant ones. And if there are such facts, it seems likely that its being wrong to steal from others (except perhaps in certain exceptional circumstances) would be among them. If there are not moral facts, or moral facts are (at least partly) subjective in that they depend (at least partly) on a person's preferences – perhaps her rational and fully empirically informed preferences – then the mere fact that someone (rationally and with full information) prefers *X* over *Y* on the basis of a feeling or judgment that she is morally obligated to do so is sufficient to make doing/getting *X* more conducive to this person's well-being than doing/getting *Y*.

²⁵ Such a move is made by, e.g., Matthew Adler and Eric A. Posner in "Implementing Cost-Benefit Analysis When Preferences Are Distorted," in *Cost-Benefit Analysis: Legal, Economic, and Philosophical Perspectives*, Adler and Posner (eds.), (Chicago: University of Chicago Press, 2001), pp. 269-311.

²⁶ See, e.g., Martha Nussbaum, *Women and Human Development: The Capabilities Approach*, (Cambridge, UK: Cambridge University Press, 2001), especially Chapter 2, "Adaptive Preferences and Human Development."

²⁷ Jean-Jacques Rousseau, *The Social Contract*, Maurice Cranston (trans.), (New York: Penguin Putnam Inc., 1968), Book I, Chapter II, p. 52.

²⁸ It might even plausibly be said that these factors have *entirely created* my preference for listening to Dylan.

²⁹ There are reasonable reservations about the informed-preference account based in a concern that it does not seem to be a plausible account of the well-being of children. The problem, I think, is not with the account, but with the fact that it is so hard to even imagine children having and processing all the information the account requires, and formulating an internally consistent set of preferences in response to it. But in any case, I will be happy if it is granted to me that my account is plausible with respect to adults. I am not entirely averse to the idea of giving a different type of account of well-being for children and animals, (and even possibly plants and ecosystems).