

Well-Being and the Priority of Values

1. Introduction

Personal well-being or welfare is the thing at issue when we ask after someone's "best interests," when we debate what would be for someone's "own good," and when we discuss what would benefit or harm someone. The theory of welfare value attempts to explain what it is that constitutes being well off or badly off, that is, what facts about well-being turn on. Consequently, it aims to specify those states or conditions that are *directly* or *noninstrumentally* good for a person, as opposed to those states or activities that are merely *instrumentally* good for a person.

The concept of well-being figures prominently in contemporary consequentialist analyses of right action, in accounts of the virtues of benevolence and kindness, in accounts of the prima facie duties of beneficence and nonmaleficence, and in various proposals concerning distributive justice. The analysis of well-being is consequently important for illuminating these other concepts.

Theories of well-being are commonly classified according to their degree of objectivity or subjectivity. However, there is no agreement about how precisely to characterize these qualities. That said, paradigmatically objective theories of welfare assert that certain activities or ends—for example, knowledge, sociability, athletic accomplishment, the contemplation of beauty—are beneficial when pursued, even if the person who pursues them does not want to, or does not enjoy them.¹ Subjective theories of well-being give the subject much greater authority over what is good and bad for her. Paradigmatically subjective theories, such as actual desire satisfactionism, hedonism, and the self-assessment theory (also called the authentic happiness or life-satisfaction theory), eschew any reference to any form of objective value, and they tie well-being to a subject's actual concerns.

These paradigmatically subjective theories are thought to enjoy certain advantages over their more objective rivals, including not only the

¹See, for example, the "Objective List Theory" discussed by Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), p. 499.

paradigmatically objective theories mentioned above but also hybridized and idealized theories. Since subjective theories do not tie the existence of personal well-being to any other form of value—for example, the worthwhile, the excellent, the intrinsically good—they are more parsimonious than theories that do. They do more explanatory work with less theoretical machinery. For instance, unlike hybridized accounts of well-being, which award special welfare value to the pursuit or enjoyment of the worthwhile, paradigmatically subjective theories characterize well-being exclusively in terms of psychological states such as desire, belief, and enjoyment whose existence virtually everyone admits.² They also avoid the burden of providing an account of the worthwhile (etc.)—and the metaphysical commitments associated with such accounts. Moreover, since the paradigmatically subjective theories are actualist (they tie welfare value to a subject's actual concerns), they avoid defects associated with idealized theories such as rational desire satisfactionism. Versions of this theory base a person's well-being on the attitudes of the person's idealized counterpart—a counterpart who is more fully rational and better informed.³ However, a person's idealized counterpart might be so different from her actual self that the counterpart's preferences do not seem like a reliable guide to the actual person's good.⁴ For this reason, idealized theories are also thought to violate the “internalism requirement,” which states that “it is a condition on something being good for a person that she be capable of caring about it.”⁵

²Hybrid theories of well-being are proposed in Parfit, *Reasons and Persons*; Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986); Robert Adams, *Finite and Infinite Goods* (New York: Oxford University Press, 1999); Richard Arneson, “Human Flourishing versus Desire Satisfaction,” *Social Philosophy and Policy* 16, no. 1 (1999): 113-42; Stephen Darwall, *Welfare and Rational Care* (Princeton: Princeton University Press, 2002); and Kwame Anthony Appiah, *The Ethics of Identity* (Princeton: Princeton University Press, 2005). The “desert-adjusted intrinsic attitudinal hedonism” discussed in the final pages of Fred Feldman, *Pleasure and the Good Life* (New York: Oxford University Press, 2004), is also a hybrid view.

³Rational desire satisfactionism is proposed in Richard Brandt, *A Theory of the Right and the Good* (Oxford: Clarendon Press, 1979). For closely related views, see John Rawls, *A Theory of Justice*, revised ed. (Oxford: Oxford University Press, 1999); Peter Railton, “Moral Realism,” *The Philosophical Review* 95 (1986): 163-207; and Thomas Carson, *Value and the Good Life* (South Bend: Notre Dame University Press, 2000).

⁴See J. David Velleman, “Brandt's Definition of ‘Good’,” *The Philosophical Review* 97 (1988): 353-71; David Sobel, “Full Information Accounts of Well-Being,” *Ethics* 104 (1994): 784-810; Don Loeb, “Full-Information Accounts of Individual Good,” *Social Theory and Practice* 21 (1995): 1-30; Connie S. Rosati, “Persons, Perspectives, and Full Information Accounts of the Good,” *Ethics* 105 (1995): 296-325, and “Internalism and the Good for a Person,” *Ethics* 106 (1996): 297-326; Robert Noggle, “Integrity, the Self, and Desire-Based Accounts of the Good,” *Philosophical Studies* 96 (1999): 303-31; and Mark C. Murphy, “The Simple Desire-Fulfillment Theory,” *Noûs* 33 (1999): 247-72.

⁵Rosati, “Persons, Perspectives, and Full Information Accounts of the Good,” p. 300.

On account of their actualism as well as the particular psychological states on which they are based, paradigmatically subjective theories are also thought to satisfy what L.W. Sumner has defended as a “condition of adequacy” for theories of well-being. This condition states that a theory of welfare must “make your well-being depend on your own concerns: the things you care about, attach importance to, regard as mattering, and so on.”⁶ The paradigmatically subjective theories all link well-being with subjective states that are plausibly connected with the subject’s own concerns. The actual desire satisfactionist connects well-being with the subject’s desires, reasoning that if a person’s desires are fulfilled, then the things that she cares about will be preserved, protected, or promoted. The hedonist connects well-being with enjoyment or pleasure, with the thought that if a person is enjoying herself, her life must be going well with regard to the things that matter to her. Finally, the self-assessment theorist links positive welfare with the judgment that one is satisfied with one’s life—for surely one will be satisfied with one’s life just in case one has secured the things that are important to one.

In this paper, I will first show that the most sophisticated versions of actual desire satisfactionism and hedonism do *not* actually meet Sumner’s condition of adequacy—they do not make a person’s well-being depend on the things the person cares about, attaches importance to, or regards as mattering. Classic counterexamples to these theories involving addictions, compulsions, and certain intense pleasures can be understood as demonstrations that these theories fail to give priority to the subject’s own concerns. These counterexamples show that a life might be full of desire satisfaction or enjoyment while failing with respect to the things that really matter to the agent. This, I claim, is a serious problem.

I suggest that the way to avoid this problem is to base one’s theory of well-being not on an agent’s desires or enjoyments but on an agent’s *values*. I call my particular values-based approach “well-being as agential flourishing.” According to influential conceptions of agency, an *agent* is a person who orders her life and its component activities by reference to her own values.⁷ To flourish as an agent, then, is to steadfastly realize

⁶L.W. Sumner, *Welfare, Happiness, & Ethics* (New York: Oxford University Press, 1996), p. 42. Sumner believes that his condition of adequacy expresses the defining feature of subjective theories of well-being. Here, I wish to remain neutral on what distinguishes subjective from objective theories of well-being. However, my eventual arguments do seem to cast doubt on Sumner’s proposed taxonomy. (Cf. David Sobel, “Subjectivism and Idealization,” *Ethics* 119 (2009): 336-52.) Whether it should ultimately be classed as a subjective theory or not, I believe the theory I defend enjoys the virtues of paradigmatically subjective theories just described.

⁷See, e.g., Harry Frankfurt, *The Importance of What We Care About* (New York: Cambridge University Press, 1988), especially essays 2, 5, 7, and 12, and *Necessity, Volition, and Love* (New York: Cambridge University Press, 1999), essays 8 and 14; Michael

one's values. The maximally flourishing agent will not only be successful but will also be firmly disposed to succeed. That is, her successful activity will not be lucky or accidental but will be supported by the possession of capacities and traits that dispose her to success. A person is doing well, I propose, to the degree that she resembles this paradigm of the flourishing agent. In order to resemble this paradigm, a person must (a) have values, as opposed to mere desires or enjoyments, (b) actively realize these values, and (c) maintain the physical and psychological attributes that are the causal basis for the disposition to succeed.

In the final section of the paper, I discuss the advantages of this theory over the third form of subjectivism mentioned above, the self-assessment view. This view, too, is commonly thought to meet Sumner's adequacy condition. I will present reasons to think otherwise and show that it has additional defects that are not shared by well-being as agential flourishing. First, the self-assessment view has difficulty explaining the welfare of individuals who rarely or never pause to take stock of their life. Second, it has problems rendering verdicts about personal well-being over time. Third, it is compatible with the possibility of mistaken self-assessments.

I conclude that well-being as agential flourishing enjoys the most important advantages of paradigmatically subjective views without inheriting their problems.

2. Actual Desire Satisfactionism

Actual desire satisfactionism traces its intellectual pedigree to Hobbes's *Leviathan*.⁸ It has recently been defended by Mark C. Murphy and Chris Heathwood.⁹ According to the most rudimentary version of this theory, it is good for a person to fulfill her desires, and it is bad for a person's desires to be frustrated. Actual desire satisfactionists disagree among themselves on whether, for desire fulfillment to occur, the object of a desire must simply obtain while it is desired,¹⁰ whether the subject must *know*

Bratman, "Identification, Decision, and Treating as a Reason," in *Faces of Intention* (New York: Cambridge University Press, 1996), pp. 185-208. See also n. 48 below.

⁸"But whatsoever is the object of any man's appetite or desire; that is it, which he for his part calleth good." Hobbes, *Leviathan*, VI:7 (New York: W.W. Norton, 1997).

⁹See Murphy, "The Simple Desire-Fulfillment Theory"; and Chris Heathwood, "Desire Satisfactionism and Hedonism," *Philosophical Studies* 128 (2006): 539-63, pp. 547-51. Note that Murphy defends a different view in his *Natural Law and Practical Rationality* (New York: Cambridge University Press, 2001).

¹⁰This is *objective* desire satisfactionism, called "the unrestricted theory" by Parfit in *Reasons and Persons*. It is defended by Murphy in "The Simple Desire-Fulfillment Theory." It is often claimed that this theory can avoid the counterexample to hedonism based on the case of "false pleasures" (see n. 20 below).

that the object obtains,¹¹ or whether the subject must simply *believe* that the object obtains.¹² But all versions of actual desire satisfactionism agree that if a person desires that *p* at *t*, and has a justified, true, belief that *p* at *t*, then this is directly beneficial to the degree that her desire is strong. If a person desires that *p*, and has a justified, true, belief that not-*p* at *t*, then this is directly harmful to the degree that her desire is strong. The welfare value of a life-segment (up to and including a whole life) is simply equal to the sum of the welfare values of all the desire satisfactions and desire frustrations that occur in it.

For actual desire satisfactionists, desire is a species of pro-attitude. A person has a pro-attitude towards a possible situation when, upon contemplating it, she favors it. Desire differs from other pro-attitudes like enjoyment or approval in that desire essentially involves a phenomenally felt yearning, wanting, wishing, or craving for a situation—a visceral motive “pull” towards it. The objects of desires are propositions describing situations or states of affairs (i.e., complex ways of being that might obtain).

According to Heathwood’s version of this theory, the desires relevant to well-being are, more specifically, occurrent, warm pro-attitudes of particular intensities. They are occurrent in that they are psychologically present or “activated.” Occurrent desires can be contrasted with merely dispositional desires, that is, desires that one would form if somehow prompted. They are warm in that they are not bloodless reflective preferences that reveal one’s moral and personal values, but urges whose prospect of fulfillment fills one with genuine enthusiasm.¹³ Finally, these desires come in different intensities: people want some things more than

¹¹This is one way of implementing James Griffin’s suggestion that the desire satisfactions must “enter the subject’s life.” See James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: Clarendon Press, 1986).

¹²This approach, called *subjective* desire satisfactionism, is defended by Heathwood in “Desire Satisfactionism and Hedonism.” Heathwood also defends actual desire satisfactionism in “The Reduction of Sensory Pleasure to Desire,” *Philosophical Studies* 133 (2007): 23-44. His official statement of his theory is provided in “Subjective Desire Satisfactionism” (unpublished ms., University of Colorado). This paper has already generated some discussion: see, e.g., Ben Bradley, *Well-Being and Death* (New York: Oxford University Press, 2008), pp. 38-45. Wayne Davis also defends a subjective desire-satisfactionist theory of happiness in “A Theory of Happiness,” *American Philosophical Quarterly* 18 (1981): 111-20, and “Pleasure and Happiness,” *Philosophical Studies* 39 (1981): 305-17.

¹³Note that the theory presented in Griffin’s *Well-Being* differs from many other versions of desire satisfactionism in that it is based on cold desires. Griffin writes that desires “do not have to have felt intensities; they need not be liked exclusively with appetitive states” (p. 307). Later, he writes: “One must assess their strength, not in the sense of felt intensity ... but rank in a cool preference ordering, an ordering that reflects appreciation of the nature of the objects of desire” (ibid.). On account of these features, my own values-based approach is similar to Griffin’s version of the theory, though there are still important differences, as I discuss below.

others. According to Heathwood, “‘intensity of desire’ is a nondispositional, occurrent strength of desire. It need not coincide with the desire’s location in the subject’s dispositional, rational preference ranking.”¹⁴

Despite its simplicity and considerable explanatory power, actual desire satisfactionism delivers some strange verdicts about possible cases. Consider this well-known example from Parfit:

I shall inject you with an addictive drug. From now on, you will wake each morning with an extremely strong desire to have another injection of this drug. Having this desire will be in itself neither pleasant nor painful, but if the desire is not fulfilled within an hour it will then become very painful. This is no cause for concern, since I shall give you ample supplies of this drug. Every morning, you will be able at once to fulfil this desire. The injection, and its after-effects, would also be neither pleasant nor painful. You will spend the rest of your days as you do now.¹⁵

If actual desire satisfactionism were true, then a person could reliably improve the quality of her life by letting Parfit inject her with the described drug. The drug causes a person to have intense, occurrent, warm desires *that she be injected*, and a person who then receives these injections will of course come to have a justified, true belief that she has been injected. Even if a person preferred not to become addicted to the drug, and even if she would always regret having become addicted to the drug, “it is likely that [her] initial desire not to become addicted, and [her] later regrets that [she] did, would not be as strong as the desires [she] ha[s] each morning for another injection.” And even if a person dislikes “the minor inconvenience that would be involved in remembering always to carry with [her] sufficient supplies, ... these desires might be far weaker than the desires [she] would have each morning for a fresh injection.”¹⁶ Consequently, a person could greatly improve her well-being simply by adding the drug to her daily regimen. But this is highly counterintuitive. I do not mean to deny the obvious importance of choosing realistic goals, but it is not good for a person to create and fulfill random desires.

In addition, actual desire satisfactionism holds that the intensity of a desire determines the welfare value of its satisfaction. But people often form very intense, occurrent, warm desires for the most trifling of ends, while the intensity of their occurrent, warm desires for long-term goals is much more faint. This is particularly true of desires born of annoyance. For example, a person who is dining in a restaurant might experience her desire that the waiter return with some ketchup as unusually intense. Actual desire satisfactionism implies—implausibly—that the fulfillment of

¹⁴Chris Heathwood, “Subjective Desire Satisfactionism,” §8.3.

¹⁵Parfit, *Reasons and Persons*, p. 497. Cf. Socrates’ discussion of itches in Plato’s *Philebus* (46a).

¹⁶Parfit, *Reasons and Persons*, p. 497.

this desire would be directly good for a person to a very high degree.

Or, consider the well-known phenomenon of “road rage.” Here, a person typically desires to conduct her vehicle in reckless and erratic ways, taking revenge on other drivers. While satisfying these desires may produce or lead to later desire frustrations (so that their fulfillment would not be net-beneficial), the theory implies that the relevant desire satisfactions are very, very beneficial at the time of their occurrence. The theory has similar implications for desires born of compulsions. It implies that compulsive shoppers greatly benefit from their purchases, that kleptomaniacs benefit from their acts of thievery, and that compulsive hand-washers benefit from washing their hands. Here, again, actual desire satisfactionists are not committed to the view that these things are *net-welfare-good* for the persons involved. But they are committed to the strange idea that these desire satisfactions are directly good to a high degree *when they occur*: the desires in question are very intense, and there can be no doubt that they are occurrent, warm desires, or that their objects are justifiedly and truly believed to be actual.¹⁷

These examples all illustrate the same basic point. A person’s desires do not always reflect her deepest concerns, that is, the things she cares about, attaches importance to, or regards as mattering. When they do not, their satisfaction is not necessarily beneficial. Actual desire satisfactionism says otherwise. Consequently, it fails to satisfy Sumner’s condition of adequacy.

3. Hedonism

In Plato’s *Protagoras*, Socrates introduces the hedonistic thesis that living pleasantly, over the course of one’s entire life, is to live well.¹⁸ Though attacked by Protagoras and characters in later Platonic dialogues, this approach survives to the present day. The contemporary hedonist claims that episodes of pleasure are the only things that are directly good for a person, while episodes of pain are the only things directly bad for a person. Though hedonists differ among themselves on the nature of pleasure and pain, most agree that episodes of pleasure and pain have intensities (their strength or vividness) and durations (their temporal extent). Most also espouse the following sort of additive theory: an episode of pleasure has positive welfare value that is a function of both its intensity and duration;

¹⁷Other interesting counterexamples to actual desire satisfactionism include the case of the lobotomized life in Carson, *Value and the Good Life*, p. 81, and the case of great dangers presented in Griffin, *Well-Being*, p. 307. Though there is not space to discuss these cases here, I believe that the theory defended below also generates more plausible verdicts about them.

¹⁸Plato, *Protagoras*, 351b.

an episode of pain has negative welfare value that is also proportional to its intensity and duration; the welfare value of a life-segment (up to and including a whole life) equals the sum of the welfare values of the episodes of pleasure and pain contained in that life-segment.¹⁹

Hedonism faces various objections, some of which look quite troubling.²⁰ But it might be thought that this form of subjectivism at least satisfies Sumner's criterion of adequacy and avoids the sort of objection just made against actual desire satisfactionism, for if a person is enjoying herself—if her life contains much pleasure and only minimal amounts of pain—it may seem that her life must be successful with respect to the things that matter to her. I will argue, though, that the most recent, most sophisticated, and most plausible versions of hedonism face an objection parallel to the one just made against actual desire satisfactionism, so that they, too, fail to satisfy Sumner's criterion of adequacy.

Here is the objection: Consider the episodes of intense enjoyment that might be brought on by the use of powerful opiates or the direct, electrical stimulation of one's own brain. Suppose that a particular person regards the prospect of such pleasures as repugnant and scary. This person simply does not wish to have these sorts of experiences. Moreover, she does not wish to be the sort of person who has such experiences. Suppose, though, that she is coerced or tricked into spending the rest of her life on a perpetual drug-induced high, or hooked up to an electrical stimulator, writhing in ecstasy. Standard versions of hedonism imply that her life would necessarily go fantastically well for her, even if she (at least previously) had the described attitudes towards this kind of life. Intuitively, this is just not so. Call this *the objection based on intense pleasures*.²¹

¹⁹This is precisely the structure presented in Feldman, *Pleasure and the Good Life*, p. 66.

²⁰A very popular objection to hedonistic accounts of well-being points to cases involving "false pleasures." According to this objection, if a person believes that some nonactual state of affairs obtains, and then takes pleasure in this state of affairs, this does not benefit her as much as it would if she took pleasure in the corresponding real situation. Famous thought experiments that have been used to make this point, including the "experience machine," described in Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 42-45. See also Thomas Nagel, "Death," in *Mortal Questions* (New York: Cambridge University Press, 1979), chap. 1, p. 4; and Shelly Kagan's description of "the deceived businessman" in "Me and My Life," *Proceedings of the Aristotelian Society* 94 (1994): 309-24, p. 311.

²¹Cf. J.J.C. Smart, "An Outline of a System of Utilitarian Ethics," in J.J.C. Smart and Bernard Williams (eds.), *Utilitarianism: For & Against* (New York: Cambridge University Press, 1973), pp. 18-20. Smart asks us to contemplate "the voluptuary of the future, a bald-headed man with a number of electrodes protruding from his skull, one to give the physical pleasure of sex, one for that of eating, one for that of drinking, and so on. Now is this the sort of life our ethical planning should culminate in? ... [H]ours spent at a switch, continually electrifying various regions of one's brain? Surely not. Men were made for higher things, one can't help wanting to say, even though one knows that men weren't made for anything, but are the product of evolution by natural selection." Similarly, Eliz-

The textbook hedonist cannot plausibly claim that people typically enjoy other activities and relationships to a greater degree than they would enjoy a perpetual opioid high. After all, the difference in quantity of enjoyment would have to come either from: (a) the number of episodes of enjoyment, (b) their intensity, or (c) their duration. A perpetual high seems well-positioned to outscore more familiar enjoyments on all three factors. In addition, such pleasures are reported to exercise a semi-hypnotic effect: they occlude one's mind, edging out other thoughts. It is therefore likely that people living such lives would not be simultaneously pained by their neglect of their spouse, children, career, or previous interests, because their attention would be entirely captured by their recurrent phenomenal experiences.

But as noted above, there are as many possible versions of hedonism as there are different views on the nature of pleasure. Both Roger Crisp and Fred Feldman have recently proposed their own versions of hedonism based on new accounts of pleasure.²² There is some reason to hope that these versions of the theory will be of assistance in answering the objection based on intense pleasures.

Crisp's form of hedonism involves a modified version of the *internalist* theory of pleasure. The traditional internalist theory says that all (and only) episodes of pleasure share a certain "hedonic tone"—an introspectively observable qualitative character.²³ Crisp explains his modified version of this view in the course of responding to the heterogeneity argument against the standard internalist theory, which asserts that two bona fide pleasures may have nothing in common with one another, phenomenally speaking. He writes:

[While] enjoyableness is indeed to be understood internally, there is a plurality of feeling tones ... there *is* a way that enjoyable experiences feel: they feel enjoyable. That is, there is something that it is *like* to be experiencing enjoyment, in the same way that there is something that it is like to be having an experience of colour. Likewise, there is something that it is like to be experiencing a particular kind of enjoyment ... in the same way that there is something that it is like to be having an experience of a particular colour. Enjoyment, then, is best understood using the determinable-determinate distinction, and the mistake in the heterogeneity argument is that it considers only determinates.²⁴

abeth Anderson has criticized hedonism on the grounds that it implies "that the ideal life would be that of a drug addict on a perpetual high, permanently absorbed in his own states of consciousness." *Value in Ethics and Economics* (Cambridge, Mass.: Harvard University Press, 1993), p. 124.

²²Roger Crisp, *Reasons and the Good* (New York: Oxford University Press, 2006); Feldman, *Pleasure and the Good Life*.

²³The *externalist* theory of pleasure denies this, holding that pleasures are simply experiences that the subject desires to experience for their own sake. If the externalist theory of pleasure is true, then there will be no real difference between hedonism and actual desire satisfactionism, so long as the latter is formulated in terms of intrinsic desires.

²⁴Crisp, *Reasons and the Good*, pp. 108-9.

Here, Crisp seems to propose that the relation between (e.g.) the pleasure of solving a Sudoku puzzle and the pleasure of eating a large Greek meal is analogous to the relation that holds between seeing a patch of red-orange and seeing a patch of deep, purplish red. Both patches are determinates of the determinable, *red*. However, to see a patch of red-orange is not to have exactly the same qualitative experience as one would have if one saw a patch of deep, purplish red.

Crisp's form of internalism will presumably count the phenomenal experiences associated with drug-induced highs as determinants of the determinable, *pleasure*. But Crisp makes a further innovation that may be helpful for answering the objection at hand. On his view, the welfare value of an episode of enjoyment depends entirely on *how enjoyable it is*, as opposed to its intensity or duration. Longer lasting episodes of enjoyment are not *ceteris paribus* better for an agent. The intensity and duration of an episode of enjoyment *may*—but *need not*—affect how enjoyable it is. Other factors, for example, the nobility of an experience, may also affect how enjoyable it is.²⁵ Consequently, Crisp might claim that the drug-induced high—while it may be very “intense” along certain phenomenal dimensions—cannot be especially enjoyable if the person desires not to experience it.

Even if the nobility of an experience can affect how enjoyable it is, whether it does so will surely depend on whether the subject retains the ability to critically reflect on the nature of the experience while it is occurring. But very intense pleasures impede the subject's ability to do precisely this. If the subject never recovers her critical faculties, then there may be no basis for denying that the drug-induced high is quite enjoyable. A subject might even rank the experiences associated with electrical or chemical stimulation as very enjoyable, while simultaneously desiring never to experience them again. Would these experiences be beneficial? Crisp's theory seems to imply that they would be: so long as the experiences are highly enjoyable at the time of their occurrence, they are proportionally good for the subject at that time. This seems problematic.

Perhaps, though, the intrinsic attitudinal hedonism recently proposed by Fred Feldman can avoid these difficulties. Feldman believes that the kind of pleasure that is relevant to well-being is *intrinsic attitudinal pleasure*. Attitudinal pleasures need not have any common sensory, qualitative, or affective character: “attitudinal pleasures need not have any ‘feel’.”²⁶ But attitudinal pleasures always have a propositional object—

²⁵Ibid., p. 110.

²⁶Feldman, *Pleasure and the Good Life*, p. 56. Leonard Katz, “Pleasure,” *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu/entries/pleasure/>, accessed July 10, 2009), criticizes this element of Feldman's theory. See also Daniel M. Haybron, *The Pursuit of Unhappiness: The Elusive Psychology of Well-Being* (New York: Oxford

some situation, state of affairs, or fact that the pleased person believes is actual. “A person takes attitudinal pleasure in some state of affairs,” Feldman writes, “if he enjoys it, is pleased about it, is glad that it is happening, is delighted by it.”²⁷ Feldman specifies that one’s welfare is directly affected only by episodes of *intrinsic* attitudinal pleasure that enter one’s life. These occur when a person takes pleasure in some situation “for its own sake” and not for the sake of what it causes or produces.²⁸

Feldman might claim that the pleasures associated with electrical or chemical stimulation are not typically very good for a person, because they are not genuine attitudinal pleasures. Attitudinal hedonism requires that pleasures be taken in situations, states of affairs, or facts that are described by full, subject-predicate sentences. So while the person hooked up to the electrical stimulator may enjoy the phenomenal feelings produced by it, she might not take attitudinal pleasure in *the fact that she is experiencing these feelings*. This fact might even disgust her.

This response to the objection comes, I think, at too high a price. Consider a lazy sunbather who enjoys an afternoon on the beach. She enjoys the warmth of the sun, the sea air mixed with the scent of coconut suntan oil, the sound of the surf. Surely the pleasures involved here are at least slightly beneficial. But for Feldman to count them as beneficial, there must be some fact or situation in which the sunbather takes attitudinal pleasure. But the sunbather is enjoying certain *experiences* directly, not *the fact that she is having certain experiences*. Consequently, if a version of hedonism is based exclusively on attitudinal pleasures with propositional objects, it misses some important pleasures.²⁹ It seems that Feldman should relax his approach so that it also counts towards one’s well-being enjoyment taken directly in activities and experiences. But if he does this, then he cannot appeal to the fact that the drug-user’s enjoyment is nonpropositional to deal with the objection based on intense

University Press, 2008), pp. 64-65.

²⁷Feldman, *Pleasure and the Good Life*, p. 56. Importantly, “a person may enjoy something he has never desired, and a person may desire a thing he never enjoys” (p. 70).

²⁸Thomas Blackson criticizes this aspect of Feldman’s theory in “On Feldman’s Theory of Happiness,” *Utilitas* 21 (2009): 393-400.

²⁹Feldman could also argue that since there is no state of affairs that is *consciously believed to obtain* by the person on the electrical stimulator, there is no corresponding episode of attitudinal pleasure. But if he opts for this sort of response, here, he will again have difficulty accounting for the benefits enjoyed by the lazy sunbather. For unless our sunbather is practicing Vipassana meditation while sunbathing, it is unlikely that she is consciously noting her experiences. However, if something less than occurrent belief that *p* is required for being intrinsically attitudinally pleased that *p*, then it seems Feldman must also allow that the person undergoing electrical or chemical stimulation takes attitudinal pleasure in these states of affairs, too. Perhaps for similar reasons, Feldman appears to allow that pleasures involving less-than-occurrent beliefs count as attitudinal pleasures. See his discussion of “unconscious pleasures” in *Pleasure and the Good Life*, pp. 114-17.

pleasures.³⁰

However, Feldman might still respond to the objection in a different way. He might claim that while the person may be pleased to be experiencing the feelings caused by the electrical stimulator or the drug, she may also simultaneously be displeased that she is having such feelings, or displeased that she is enjoying such feelings. Consequently, the experience may not be net-beneficial, provided that the subject also takes attitudinal pain in some aspect of her condition.

But Feldman's attitudinal hedonism still dramatically overestimates the welfare value of the chemically or electrically induced experiences *while they are occurring*, for the theory must say that those experiences—since they are enjoyed to a high degree—benefit the person tremendously while they are happening. If the subject of electrical or chemical stimulation takes attitudinal pain in some aspect of her condition, the welfare-disvalue this pain represents can come only later, when the person recovers her reflective capacities and considers whether she is pleased or pained to have had such feelings.³¹ So on Feldman's theory—just like on Crisp's—it seems to follow that it would be incredibly beneficial for a person to remain in this euphoric condition and never recover her reflective capacities. This seems false.

The theories of Crisp and Feldman fail for precisely the same reason as actual desire satisfactionism: they do not tie well-being closely enough to a person's concerns, the things that matter to her, the things she cares about. Just as a person may have an intense desire for an outcome that she does not truly care about, a person may find an experience highly

³⁰Relaxing the approach in this way may lead to problems of double counting when a person enjoys an experience *and* the activity associated with the experience *and* the fact that she is having the experience. But perhaps these can be addressed in some other way.

³¹Feldman also presents a version of attitudinal hedonism that is designed to meet the objection based on "false pleasures" (see n. 20 above). Readers who are swayed by this objection are invited by Feldman to consider "truth-adjusted" attitudinal hedonism. (Note that Feldman does *not* seem swayed by this objection; see *Pleasure and the Good Life*, pp. 42-43, 110-11. He seems to favor the simple intrinsic attitudinal hedonism already described.) The key idea, here, is that if the objects of one's enjoyment *do not exist*, then the welfare value of one's enjoyment should be multiplied by .01 (p. 112). It might be thought that this alternative theory can answer the objection based on intense pleasures. But this is not so. For in these cases, there is no question about the existence of the object of pleasure. One is enjoying the feelings caused by the electrical stimulator or the drug. And one is really having those feelings. Consequently, their "truth-adjusted" welfare value is not discounted—it is exactly what it would be on the regular version of attitudinal hedonism. L.W. Sumner presents several further interesting and persuasive criticisms of the truth-adjusted version of the theory in "Feldman's Hedonism," in Kris McDaniel, Jason R. Raibley, Richard Feldman, and Michael J. Zimmerman (eds.), *The Good, the Right, Life and Death: Essays in Honor of Fred Feldman* (Aldershot: Ashgate Publishing, 2006), pp. 83-100. He concludes, as I do, that the non-truth-adjusted version of the theory is more promising.

enjoyable without it really mattering to her—or even while actively devaluing it! Such enjoyments are not tremendously beneficial. A life consisting entirely of such enjoyments would not be ideal in terms of personal welfare. Since both these versions of hedonism say otherwise, they fail to satisfy Sumner's condition of adequacy, and we have good reason to reject them.

4. Well-Being as Agential Flourishing

We have seen that the fulfillment of addictive or compulsive desires is directly beneficial to a very small degree, if at all. We have also seen that episodes of very intense pleasure are not proportionally good for a person, particularly if the person desires not to have these experiences—or not to be the type of person who enjoys them.

Many years ago, John Dewey argued that we should distinguish valuing—that is, caring for, esteeming, prizing, holding dear, admiring—from both evaluative judgment (belief that something is valuable) and simpler subjective states such as enjoyment, liking, and desire.³² Several contemporary theorists have proposed similar distinctions.³³ Depending on exactly how these distinctions are drawn, they may or may not be useful or tenable. But I believe that there is a concept of valuing that is not unfamiliar—a concept that opens up promising avenues of thought for the theory of well-being. I will attempt to explain this concept. My hope is that an account of well-being based on the realization of one's values will, paraphrasing Sumner, make one's well-being depend on one's own concerns: the things one cares about, attaches importance to, and regards as mattering. Consequently, such an account will not face the objections that threaten actual desire satisfactionism and hedonism.³⁴

³²See John Dewey, *The Quest for Certainty* (New York: Menton, Balch, and Co., 1929), pp. 254-86, as well as his "Moral Judgment and Knowledge," in Larry A. Hickman and Thomas M. Alexander (eds.), *The Essential Dewey*, vol. 2 (Bloomington: Indiana University Press, 1998 [1932]), pp. 328-40.

³³See Harry Frankfurt, *The Importance of What We Care About, and Necessity, Volition, and Love*; Gary Watson, "Free Agency," *The Journal of Philosophy* 72 (1975): 205-20; David Lewis, "Dispositional Theories of Value," *Proceedings of the Aristotelian Society*, suppl. vol. 63 (1989): 113-37; Gerald Gaus, *Value and Justification* (New York: Cambridge University Press, 1990); David Copp, *Morality, Normativity, and Society* (New York: Oxford University Press, 1995); Bratman, "Identification, Decision, and Treating as a Reason"; Noggle, "Integrity, the Self, and Desire-Based Accounts of the Good"; Valerie Tiberius, "Humean Heroism: Value Commitments and the Source of Normativity," *Pacific Philosophical Quarterly* 81 (2000): 326-46; and Donald C. Hubin, "Desires, Whims, and Values," *The Journal of Ethics* 7 (2003): 315-35.

³⁴Richard Kraut, "Desire and the Human Good," *Proceedings and Addresses of the American Philosophical Association* 68, no. 2 (1994): 39-54, and Haybron, *The Pursuit of Unhappiness*, p. 195, both note that desire satisfactionism is more plausible if under-

Values are often contrasted with desires born of addiction or compulsion. To use a well-worn example, addictive smokers habitually desire to smoke, though they may not value smoking. Suppose we ask such a smoker, “Is smoking one of the things that you value?” It is likely (though not certain) that the smoker will say, “No—though I really want a cigarette right now, I wish I did not. I would rather not be a smoker. I do not really value smoking.” Reflection on this sort of case has inspired two well-known approaches to understanding valuation: the view that to value something is to desire to desire it, and the view that to value something is to appraise it as good or worthwhile.³⁵

Both of these approaches face well-known objections. Against the first approach, it is often claimed that since there are even higher orders of desire (e.g., the third-order desire not to desire to desire to smoke), there is no good reason to equate values with desires that are approved at the second order. Against the second approach, it can be pointed out that a person might value something without thereby believing that it is objectively worthwhile or meritorious. For example, a person might value completing a marathon, owning a home near the water, or seeing a certain musician in concert, without taking any position on the objective value of these activities and situations.³⁶ Moreover, the second approach implies that a person can value something without having any pro-attitude towards it—that is, that a judgment of merit suffices for valuation. But if a repressed homosexual judges that physical intimacy with members of the opposite sex is good and worthwhile, without having any of the relevant desires, we would not say that the person values heterosexual sexual activity.³⁷ Judgments of merit do not seem to be either necessary or sufficient for valuing.

My proposal is that valuing involves stable identification with one’s

stood in terms of wholehearted or endorsed desires, before going on to reject this sort of theory.

³⁵For the first sort of view, see Frankfurt, *The Importance of What We Care About*, essay 2, and Lewis, “Dispositional Theories of Value.” For the second sort of view, see Watson, “Free Agency.”

³⁶See Frankfurt, *Necessity, Volition, and Love*, essay 14.

³⁷Cf. Nomy Arpaly, *Unprincipled Virtue* (New York: Oxford University Press, 2003), chap. 1. Arpaly argues that ego-syntonic desires are not necessarily representative of an individual’s interests. For example, the ego-syntonic desires of repressed homosexuals and anorexics should be reformed. My own view—discussed in greater detail below—is that *wholehearted* ego-syntonic desires—i.e., those that are (a) coherent with a person’s lower-level appetites, urges, and wants, and (b) not contradicted at any higher level of self-evaluation—are correct guides to the person’s own good. The fact that these desires are confirmed at the second-order level is a testament to their wholeheartedness, even though second-order desires are not necessarily authoritative. What about a person whose desires are all in a jumble, who has no wholehearted desires? For her own good, she should get some.

pro-attitudes: when one has a pro-attitude towards an object, and also stably identifies with this pro-attitude, then one values the object.³⁸ Values in this sense represent a person's true concerns and perhaps define her practical identity.³⁹

Several components of this account are relatively straightforward. The possible objects of valuing include persons, things, situations, relationships, activities, experiences, and ideals. To give some examples, one can value one's spouse, one's business, one's antique boat, one's relationship with one's brother, advancing in one's profession, ice hockey, tasting fine wines, honesty, or freedom. The relevant pro-attitudes include desire, enjoyment, approval, liking, and love. Some of the objects just mentioned are situations that are typically desired (advancement in one's profession), others are activities that are typically enjoyed (playing ice hockey, tasting fine wines), others are ideals that one typically approves of (honesty and freedom), and still others are liked objects (one's antique boat) or loved persons (one's spouse). Stable identification is identification that persists over a longer stretch of time, though it may also temporarily diminish or even vanish.⁴⁰

Identification, itself, requires a more detailed explanation. Identification characteristically involves at least the following three features, though borderline cases that lack one of these features might also be possible. First, if a person identifies with a pro-attitude, she tends not to experience the attitude as alien or as "coming from the outside": valuing has a particular phenomenology. Second, if a person identifies with a pro-attitude, she will be "wholehearted" with respect to it.⁴¹ This means that she will be disposed to approve of the attitude—to take it as representative of who she is and who she wants to be. It also means that there is no conflict between the attitude and her attitudes involving the same object at other (lower or higher) levels. For these reasons, the person will be disposed not to question or reconsider the attitude. Third, if a person identifies with a pro-attitude, she will be disposed to treat the attitude as reason-giving in the sense of justifying her actions.

Valuing requires some measure of awareness of the valued object's real characteristics. One cannot be said to have a pro-attitude towards a

³⁸My account—like Frankfurt's, Bratman's, and Tiberius's—is "noncognitivist," as R. Jay Wallace puts it. See his "Caring, Reflexivity, and the Structure of Volition," in *Normativity and the Will* (New York: Oxford University Press, 2006), pp. 190-211.

³⁹Cf. Christine Korsgaard, *The Sources of Normativity* (New York: Cambridge University Press, 1996), pp. 100-112.

⁴⁰The account of valuing presented in Valerie Tiberius, *The Reflective Life* (New York: Oxford University Press, 2008), pp. 23-61, also requires that values be stable in this way.

⁴¹This is Frankfurt's response to Watson's criticism of the second-order desire approach to valuation. See Frankfurt, *Necessity, Volition, and Love*, essay 8.

thing unless one has a basic understanding of what the thing is. That said, the requisite degree of awareness is not easy to specify. One might value something because it is somewhat mysterious or intriguing—that is, partly in virtue of the fact that one does not have a complete understanding of it. Still, if “valuing” is sustained by a thoroughgoing ignorance of the “valued” item, or if it is sustained mainly by false beliefs about the “valued” item, then it is not genuine valuing.

Valuation comes in degrees, of course. It is natural to expect that how important the item is to the subject will depend both on the strength or intensity of the pro-attitude in question and the stability of the subject’s identification with this pro-attitude.⁴² The more important to a person a particular value is, the more significant it is for her well-being.

The stability of pro-attitudes—and the stability of identifications with these pro-attitudes—is typically enhanced when the subject has a more or less complete understanding of the attitude’s object, as well as approximately true beliefs about adjacent parts of the world. A person’s pro-attitudes towards an item will not hold up in the long-run if they are based on significant factual misunderstandings. In reflective agents like us, stability is also enhanced by the belief that both the object *and* the means necessary for its realization cohere with the other things that one values. Stability is further increased when a particular value is implied by the agent’s other values so that its pursuit complements and reinforces them. This is not to say that it is impossible to discover conflicts among one’s values: this can surely happen, especially when the means of promoting or safeguarding one value turn out not to be compatible with the means required to promote or safeguard a second value. But valuing things that are in tension with one another risks both cognitive dissonance and failure in action, both of which threaten the stability of one’s pro-attitudes and one’s identification with them.

Certain sorts of values tend to be particularly stable and resilient. This is true, for example, of what I call one’s *ownmost values*.⁴³ When something is one of one’s ownmost values, one is stably disposed to nurture or maintain one’s pro-attitude for it, as well as one’s identification with this pro-attitude. It is not merely that one has (e.g.) a desire and identifies with it; rather, one actively strives to keep this desire—and one’s identification with it—alive.

The sort of activity appropriate to a particular value will vary with the nature of its object, as well as with the nature of the pro-attitude that con-

⁴²Other factors may be important here, too. It is beyond the scope of this paper to provide a full theory of weighting for values. For doubts about the possibility of providing such an account, see Mark Johnston, “Dispositional Theories of Value,” *Proceedings of the Aristotelian Society*, suppl. vol. 63 (1989): 139-74.

⁴³Frankfurt, *Necessity, Volition, and Love*, essay 14, explores values of this sort.

stitutes it. Enjoyed activities like fishing, dancing, or contemplating poetry are properly *pursued*, not in the sense that one strives after them but in the sense that one participates in them. By contrast, loved persons, important relationships, and cherished objects must be *maintained*, that is, safeguarded, preserved, nurtured, or protected. Desired ends, however, are to be *achieved*, while ideals such as freedom and honesty are to be promoted or exemplified. I use the term “realized” to cover all these forms of activity.⁴⁴ It should be noted that some activities, such as pondering a philosophical theory, inventing a new sort of bridge, or devising a scenario for use in a novel do not mainly involve the movement of one’s limbs. But mental actions are still actions.⁴⁵

Even if a theory of well-being should be centered on the realization of one’s most important values, it must also accommodate other sources of direct benefit and harm. It cannot be denied that leading a cheerful and joyful life is also good for one. For example, if two lives are equal in the amount of value-oriented accomplishment they contain, but the first is experienced by its subject as joyous and rewarding, while the second is experienced merely as a hard slog, we would judge that the first life is better for the one who lives it. Intuitively, other emotional states can also augment or diminish one’s welfare: other things being equal, it is good to feel confident, optimistic, and energized; it is bad to feel impotent, anguished, and alienated.⁴⁶ This suggests that our account of welfare must recognize the direct welfare value of affective states.⁴⁷

⁴⁴Since aims are plausibly thought of as involving desires for ends and intentions to realize these ends through one’s own action, the theory proposed here partially resembles the aim achievementism of T.M. Scanlon and Simon Keller. See T.M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998), pp. 108-43; Simon Keller, “Welfare and the Achievement of Goals,” *Philosophical Studies* 121 (2004): 27-41, and “Welfare as Success,” *Noûs* 43 (2009): 656-83. On my view, though, aims are only one species of value, and achievement is only one kind of value realization. Additionally, I do not claim—as Scanlon does—that the contribution of successful aims to one’s well-being varies with the degree to which they are objectively “worthwhile.”

⁴⁵This focus on activity enables the theory to handle an objection from Parfit, *Reasons and Persons*. When Parfit meets the stranger, and forms the desire that he be cured—and then never meets or thinks about him again—the stranger’s being cured is not one of Parfit’s aims. Parfit never acts or even intends to promote this state of affairs; indeed, he never *does* anything relevant to this person’s condition. Cf. Scanlon, *What We Owe to Each Other*, p. 120.

⁴⁶See Haybron, *The Pursuit of Unhappiness*, for an extended discussion of positive emotions and related mental states, which he takes to be components of happiness.

⁴⁷This is particularly important for my theory, if it is to count as an improvement on Feldman’s attitudinal hedonism. Earlier, I argued that if Feldman does not expand his theory so that it counts enjoyed activities and experiences in addition to enjoyed propositions or states of affairs, it will not be able to explain the benefit of the positive affective states associated with, e.g., the pleasure of sunbathing. It would be problematic if my

Not only must the theory make room for the importance of affect. If it is to predict and systemize our commonsense judgments, it must also place some limits on the sorts of ends that can be legitimately pursued, while remaining basically ecumenical. Here is what I mean. The true theory of welfare must recognize that different ways of life can be equally good for different people. Some thrive on the dangers of mountain-climbing and downhill skiing, while others prefer to quietly exercise their minds with a mystery novel or a game of chess; some have children and put family above all else, while others remain single and seek out the excitements of urban life. At the same time, though, certain activities seem directly harmful, even if they are valued and pursued. Suppose one were to learn that a loved one was contemplating elective bodily dismemberment, or becoming interested in extreme forms of masochism, or using dangerous, brain-damaging substances like phencyclidine (PCP). One would surely be concerned, and the basis for one's concern would be more than a puritanical aversion to deviance.

In order to explain the welfare value of positive affect and simultaneously explain why some values are best left unpursued, our theory must recognize sources of benefit and harm beyond success and failure with respect to activities governed by one's values. If we construe well-being as agential flourishing, I think that our theory will be able to do just this.

An agent, we have said, is one who regulates her activities by reference to her own values.⁴⁸ The successful exercise of one's agency, then, will involve actively realizing one's values through one's own effort over the course of a normal human life-span. But exercising one's agency in a way that is not merely successful, but robustly so, requires something more. It requires being stably disposed to realize one's values in those nonactual situations that are likely to occur. To be stably disposed in this way, one must have redundant strategies, capabilities, and resources in place. In particular, there is good reason to suppose that one must be aware of how one is doing with respect to one's values, and that one must have many of the properties commonly associated with physical and psychological health, including an adequate amount of positive affect and functionally appropriate emotions. If one realizes one's values and furthermore has these other properties that constitute the causal basis for the disposition to realize one's values, then one is *flourishing as an*

own theory could not explain the welfare value associated with such states. But I believe it can (see below).

⁴⁸Though I take this to be a plausible characterization of agency, my main contention is that the process I describe as agential flourishing is what well-being consists in, i.e., what true judgments about well-being turn on. Strictly speaking, one could accept my theory of well-being while rejecting this thesis about the nature of agency—although one ought then to use some other name for the theory.

agent, that is, realizing one's values in a stable, unswerving, and steadfast way that represents the optimal functioning of the human action-guiding system. This is the paradigm case of agential flourishing. To the degree that an adult human person resembles or matches this paradigm case at a time, she is doing well at that time.⁴⁹ To the degree that a human life approximates this paradigm over the course of a normal human life-span, it goes well for the one who lives it; to the degree that a life departs from this paradigm, its welfare value is compromised. Any change in a person's condition that moves her further away from this paradigm is directly bad for her at the time of its occurrence; any change that moves a person closer to this paradigm is directly good for a person at the time of its occurrence. The degree of harm or benefit, here, is proportional to the degree to which the person has been moved closer to—or farther from—the paradigm.⁵⁰

For example, if a person actively realizes her most important values, but her disposition to realize these values is occasionally weakened because she is unhappy and inappropriately angry, she may still be doing quite well on balance (though less well than she might be). But if a person's life departs to a great degree from this paradigm—if the person does not really have any values, or if she fails with respect to the things that matter most to her, or if she is physically infirm, chronically unhappy, and in poor mental health—her life may be positively bad for her.

This theory construes well-being as a partly dispositional property: the person who flourishes fully as an agent must not only realize her values but be stably disposed to realize them.⁵¹ The dispositionality of well-being allows us to claim that some things that have traditionally been thought to be merely instrumentally good are actually directly good. For instance, functionally appropriate emotions turn out to be directly good because they are part of the categorical basis for the disposition to realize one's values; they help constitute this disposition. This aspect of the theory also enables us to explain the direct benefit of positive affect, as

⁴⁹Though I do not have space to discuss the matter in detail, I would extend this theory to cover children in the following way. We should look not to see whether they currently resemble the flourishing agent, but whether they are well-equipped to become such an agent as adults.

⁵⁰As I have formulated it, well-being as agential flourishing is a *holistic* theory of well-being: we cannot learn whether a change in a person's condition is directly good or bad for the person merely by inspecting its intrinsic character, or by looking at the person's momentary state. We must look to see whether, in the context, it moves an agent nearer to or further from the paradigm of the flourishing agent. Consequently, I do not believe there is an accurate paraphrase of it into the atomistic formal structure proposed by Bradley (*Well-Being and Death*) as a template for all theories of well-being.

⁵¹Cf. Gilbert Ryle, *The Concept of Mind* (Chicago: University of Chicago Press, 1949), p. 43.

well as why goals that undermine one's disposition to realize one's values are best forsaken.

The question of exactly which states support this disposition is ultimately an empirical one.⁵² However, widely known facts support the idea that many states ordinarily associated with physical health partially constitute the disposition to realize one's values. If one's body is damaged or missing crucial parts, one will not be positioned to cope with the various threats to one's values that are likely to materialize. But an intact body with properly functioning senses, locomotive abilities, and organs will fortify one's ability to realize one's values. Physically impaired individuals can of course lead lives that go quite well on balance, but if states associated with physical health are part of the categorical basis for dispositional success, then bodily damage is almost always directly bad.

States linked with emotional health are also good candidates for the categorical basis for the disposition to succeed. One such state is happiness—or rather, adequate or minimal happiness.⁵³ For one to be adequately happy, positive affect must predominate over the course of one's life. This state is arguably part of the categorical basis for dispositional success because one must be in good spirits or good cheer (though not necessarily elated or ecstatic) in order to be engaged with the world and ready for action. Feeling happy (but not *too* happy) also improves the operation of cognitive abilities that are relevant to reliable introspection, the selection of values, and the formulation of rational plans of action.⁵⁴ For this reason, if one's values are ill-suited to one's affective nature or one's implicit self—so that one is often unhappy, despite one's successes—one will not be flourishing as an agent.

However, it can also be directly beneficial to respond with sadness

⁵²Here, I emphasize states associated with physical and emotional health as important causal bases for this disposition. However, future empirical inquiry might establish that other factors also underwrite the disposition to realize one's values. The testimony of people who are very successful supports the view that one must be confident in one's abilities and trusting of one's own judgment in order to pursue, maintain, and achieve one's values over the long haul. In addition, several nonmoral virtues appear to further reinforce one's ability to cope with adversity. These include rationality, courage, temperance, fortitude, perseverance, loyalty to one's values, and adeptness at optimistic sense-making of one's life experiences. For a good discussion of sense-making, see Jonathan Haidt, *The Happiness Hypothesis* (New York: Basic Books, 2006), pp. 145-49. Cf. Timothy D. Wilson, *Strangers to Ourselves* (Cambridge, Mass.: Harvard University Press, 2002), pp. 203-21. Furthermore, Tiberius discusses various habits of mind that, according to her, constitute *reflective wisdom*. These include knowing when and when not to adopt a reflective point of view, maintaining proper perspective, self-awareness, and optimism. See *The Reflective Life*, pp. 65-157.

⁵³I discuss happiness and its connection with well-being at much greater length in my forthcoming paper, "Happiness is not Well-Being" (unpublished ms., California State University, Long Beach).

⁵⁴See Jennifer Hawkins, "The Subjective Intuition" (unpublished ms., Duke University).

and grief to loss—or with anxiety to looming threats. Another aspect of emotional health, accordingly, is the experience of functionally appropriate emotions. An emotional experience is functionally appropriate when, in the circumstances, it disposes one to successful action ordered by one's values. Negative emotions are appropriate when one's values are threatened or harmed, because they ordinarily dispose one to neutralize the threat or rectify the harm.⁵⁵ The “flow” experiences famously described by Mihaly Csikszentmihalyi—experiences of piqued interest, increased attentional focus, and continuous joy as one carries out a task that is well-matched to one's abilities—are appropriate to the strivings through which one pursues one's values.⁵⁶ Enjoyment of one's successes is also appropriate. This form of emotional feedback is, first and foremost, a motivator of further success. Without such feedback, there is a significant danger that one's pro-attitudes for the relevant object may weaken—or that one may cease to identify with these attitudes.

Once again, I am not proposing that states associated with physical health, emotional health, and happiness are merely instrumentally good. They are instrumentally good, to be sure. But since well-being is partly dispositional, and since physical health, emotional health, and happiness underwrite the disposition to realize one's values, these states are partially constitutive of agential flourishing and therefore *directly* beneficial.⁵⁷

This theory illuminates the counterexamples to actual desire satisfactionism and hedonism considered earlier. It explains, first, why it would not be advantageous to accede to Parfit's proposal that he inject you with an addictive drug. Unless you would stably identify with your desires for further injections (these desires are created by the drug), such injections would not benefit you directly, because they would not be valued, but

⁵⁵It is ideal to have *both* a preponderance of positive emotional states *and* to experience appropriate emotions. These two desiderata are potentially in conflict with one another. Consequently, there is a point at which it becomes beneficial for a very unlucky person—i.e., a person who is experiencing one failure after another—to have somewhat “inappropriate” emotions, i.e., to be less sad or aggrieved than her situation would ordinarily warrant.

⁵⁶Mihaly Csikszentmihalyi, *Flow: The Psychology of Optimal Experience* (New York: Harper and Row, 1990).

⁵⁷This feature of the theory gives us a response to a worry that has plagued conative theories based on cold desires (see, e.g., Mark C. Overvold, “Self-Interest and the Concept of Self-Sacrifice,” *Canadian Journal of Philosophy* 10 (1980):105-18). If a conative theory of well-being based on cold desires is true, some say, then it is impossible that the realization of a value could ever be self-sacrificial. But suppose a parent values the education of her child, and takes a second job to pay for private schooling. Surely, it is objected, this is self-sacrifice. In brief, my response is this: if the realization of this value detracts from the degree to which the person's life resembles the paradigm case of agential flourishing, it is self-sacrificial. It might do so indirectly, by making it impossible for the person to pursue other things that she cares about more. Or, it might do so directly, if it has a negative impact on the person's health or happiness.

merely desired. And since Parfit stipulates that ingestion of the drug has no positive phenomenal character or “feel,” it is unlikely that you would stably identify with your desire for it. Also, if it has no “feel,” its use could not be productive of positive affect or happiness. Use of the drug is not even likely to be instrumentally valuable, for how could the satisfaction of these desires cause, promote, or lead to other states that are components of agential flourishing?⁵⁸

Well-being as agential flourishing can also explain why the fulfillment of very intense desires is not proportionally beneficial: the importance to an agent of a particular pro-attitude depends on the associated pro-attitude’s intensity, but also on the degree and the stability of the agent’s identification with the pro-attitude. It is implausible to think that agents stably identify with the intense desires discussed earlier: no one consistently takes a momentary desire for ketchup to be representative of who they are and who they want to be. Similarly, because desires associated with compulsions are nonvalues, fulfillment of these desires is beneficial only if it is happiness-constituting, as minimal happiness is part of the categorical basis for the disposition to realize one’s values. But many compulsives report that their actions do not make them happy or improve their mood. Even if the fulfillment of such desires were associated with momentary positive affect, well-being as agential flourishing plausibly implies that it is simultaneously instrumentally bad to a high degree, since the satisfaction of compulsive desires typically distracts a person from the things that she really cares about. In summary, well-

⁵⁸What, though, if Parfit offered to inject you with a drug that would make you *value* getting another dose of the drug every day? This drug would have to make you (a) desire repeated, periodic injections of itself, (b) stably consider these desires to be representative of who you are and want to be, and (c) treat these desires as reason-providing. If such a drug existed, my theory implies that it could benefit you to receive daily injections of it. Note, though, that a similar difficulty affects Parfit’s own proposed solution to this problem. Parfit writes that we should adopt a version of *global* desire satisfactionism, which “gives us the right answer in the case where I make you an addict. You would prefer not to become addicted, and you would later prefer to cease to be addicted” (*Reasons and Persons*, p. 498). A desire is global “if it is about some part of one’s life considered as a whole, or is about one’s whole life” (*Reasons and Persons*, p. 497). But suppose Parfit offered to inject you with a drug that would make you form a global desire for getting another dose of it every day. On his own theory, it would benefit you to accept this offer. It seems that we must accept some such implication, whichever way we go. On any subjective approach, there must be *some* form of pro-attitude, such that, if a drug could make you take that pro-attitude towards further injections of the same drug (which you could reliably get), it could make your life go well for you. I believe I have demonstrated that the pro-attitude in question cannot simply be warm, occurrent desire. Though valuing (as construed here) and Parfit’s global desires are probably related, there is reason to favor a subjective theory based on valuing, as the concept of a global desire is somewhat difficult to pin down. Suppose I want my life to be such that I eat eggs for breakfast on January 1, 2011. Is that a global or a local desire?

being as agential flourishing connects one's well-being with one's own concerns, whereas actual desire satisfactionism connects it with one's momentary urges.

The theory can also explain why intense, chemically or electrically induced pleasures are not necessarily either directly or indirectly beneficial. According to well-being as agential flourishing, an episode of enjoyment is directly beneficial only if either (a) it is a value, that is, the agent stably identifies with the enjoyment, (b) the episode constitutes the emotional state of happiness, or (c) the episode represents appropriate emotional feedback on agential success. If the person who enjoys being hooked up to the electrical stimulator does not stably identify with this enjoyment—as was stipulated—then the pleasure is not a value for her. While her experiences are certainly pleasurable, they do not seem constitutive of happiness in the sense of broadly positive affect, good cheer, or an upbeat attitude. Finally, these episodes of enjoyment are certainly not appropriate emotional feedback on agential success. In fact, being hooked up to the machine is almost certain to damage a person's motivational architecture, thereby undermining her disposition to realize her values. After experiencing pleasure like this, a person will probably find it difficult—at least temporarily—to hold or pursue other values. For this reason, such experiences may even be directly harmful.

In order for an enjoyed activity to be indirectly beneficial, as many “just fun” activities are, it must promote a person's agential success, her health, her happiness, or some other component of her disposition to realize her values. Fun activities typically restore a person to a mentally balanced state and enable her to recuperate energy and motivation. While there may be a case to be made for alcohol, nicotine, or other “soft” drugs, here, the very intense chemically or electrically induced pleasures discussed earlier manifestly do not do this. These pleasures make it more difficult to find the motivation to act, and they make it more difficult to care about anything else. It is hard to see how they could be fitted into an active life organized around the things that are really important to a person.

What, though, should be said about the case of the eager and willing addict—the person who stably identifies with her enjoyment of intense chemical or electrical stimulation? Suppose that a person simply values the feelings brought on by chemical or electrical stimulation, and nothing else. Suppose that the person is actually successful in bringing about these feelings, that her successes are counterfactually stable, and that she is disposed to bring about such feelings. Suppose she is living in a concrete bunker, safely sheltered from the elements. Her finances are in good order. She is plugged into tubes for feeding and excretion that are attached to large-capacity tanks. Her only action is to periodically press

her finger to administer further stimulation of slightly increasing dosages. The stimulation-delivery system is set up and calibrated so that it will be possible for her to continue doing this for many years before death.

At the very least, we can say this: this person is very strange, so it should not come as a great surprise that her well-being is also quite strange. But well-being as agential flourishing enables us to say something more. This scenario seems quite difficult to sustain. The problem is not that the person's feeding solution might expire, or that the power might get shut off, or that there might be an earthquake. It is that the person leading this life may not be able to stably identify with her desire for the feelings brought on by chemical or electrical stimulation. Without a complex, changing environment that stimulates her mind, and without other persons to give her feedback on her thoughts and actions, this person may sooner or later not be able even to think about who she is, or who she wishes to be, or what she has reason to do. If she cannot do these things, she cannot stably identify with her desires, and so she cannot have values. If she has no values, she cannot flourish as an agent.

5. The Self-Assessment Theory

There is a competing subjective theory of well-being that I have not yet considered: the self-assessment theory proposed by Sumner himself and recently extended by Valerie Tiberius and Alexandra Plakias. The basic idea behind this theory is that a person is the final authority on her own welfare; if a person judges that her life is going well, she cannot be mistaken. This approach appears to satisfy Sumner's adequacy condition, so that it does not face the sort of objections made earlier against actual desire satisfactionism and hedonism. This appearance, I will argue, is illusory. Furthermore, the self-assessment theory has its own difficulties. I will also illustrate these, thereby completing my case that well-being as agential flourishing enjoys the most important advantages of paradigmatically subjective views while avoiding the problems associated with them.

A rudimentary version of the self-assessment theory holds that a person is doing well (to some degree) at a time just in case she *believes* at that time that her life is satisfactory (to that degree). But this version of the theory is now known to face several serious problems.⁵⁹ First, we can imagine individuals who are misinformed or manipulated to such a degree that while they *believe* they are doing well, commonsense would say

⁵⁹See Haybron, *The Pursuit of Unhappiness*, pp. 79-103, 199-223; Fred Feldman, *What is This Thing Called Happiness?* (New York: Oxford University Press, 2010), pp. 70-97.

that they are not. Citizens of a totalitarian state might believe that their life is highly satisfactory since they are able to eat boiled grass, if they have been told that people in other countries are even worse off. Second, it is not uncommon to encounter people who *say* that they are doing quite well—and who even appear to believe that they are doing quite well—though their behavior tells otherwise.⁶⁰ A person's conscious self-assessments can diverge from her implicit beliefs and attitudes concerning her life. The self-assessment theorist should acknowledge that a person who is self-deceived or conflicted in this way is not doing as well as a person who is not. Third, self-assessments are vulnerable to various cognitive biases. For example, they are susceptible to the contrast effect. If I were told to contemplate in detail the plight of people in an impoverished, war-torn country with terrible environmental conditions, I would likely raise my assessment of my own circumstances.⁶¹ If I were instead encouraged to contemplate in detail the material circumstances of the residents of Newport Beach, California, I would probably take a dimmer view of my own lot. Psychologists who understood the mechanics of such cognitive biases could manipulate a subject's self-assessment so that it changed dramatically over the course of a few hours, without effecting proportional changes in her well-being. For all these reasons, self-assessments seem like unreliable indicators of personal well-being.

These objections have prompted the development of more sophisticated versions of the theory. Sumner himself requires that a person's self-assessment be *authentic*—both autonomous and informed—in order for it to be constitutive of personal well-being. For the assessment to be autonomous, it must not be influenced by processes that “[deny] the subject the opportunity for critical reflection on the process itself and its outcome.”⁶² Similarly, Sumner contends, we have reason to doubt the relevance of a self-assessment that is not informed.⁶³

⁶⁰Haybron provides compelling illustrations of this sort of phenomenon, though his topic is happiness, not well-being. One illustration concerns Glen, who lives in Detroit, though he seems to be in better spirits when he is in the country. When Glen's sister asks him how happy he is, he says, “What kind of question is that? Who gives a goddamn how I feel? If you have to know, I suppose I feel fine. Got nothing to complain about. Yeah, sure, I'm happy.” See Haybron, *The Pursuit of Unhappiness*, p. 199. Glen's self-assessment should probably not be taken at face value.

⁶¹Haybron describes several other such biases. Cultural norms may lead us to ignore certain forms of affect; this is “affect-type bias.” In addition, our theories about how we should feel in certain circumstances may lead us to discount certain dimensions of our experience; self-assessments might be distorted by “expectation effects” (*The Pursuit of Unhappiness*, p. 209).

⁶²Sumner, *Welfare, Happiness, and Ethics*, p. 170.

⁶³However, Sumner writes: “The relevance of information for a person's well-being is a personal matter to be decided by personal priorities; there is here no authoritative public standard” (*ibid.*, p. 161).

Tiberius and Plakias propose their own amendments to the basic self-assessment approach. They call their view the “Value-Based Life Satisfaction Account.” On this view, only positive self-assessments that are made with a view to a person’s own values are truly well-being constituting.⁶⁴ We are to imagine asking a person, “How satisfied are you with how your life is going with respect to your own values?” The person is doing well provided that she is very satisfied across those various domains of life that correspond to her values; if she is significantly dissatisfied, she is doing badly.

Tiberius and Plakias’s account of values is not dissimilar from my own account. They write that “values can include activities, relationships, goals, aims, ideals, principles, and so on.” Values “must be normative from the point of view of the person who has them: that is, a person takes her values to provide good reasons for doing things.”⁶⁵ In addition, “values include an affective component: part of what it is to care about something in the way distinctive of valuing is to have some positive emotional response toward it.”⁶⁶ Values must also be stable and adequately well-informed; if they are ill-informed, then they do not count as the subject’s real values. Finally, one’s activities, relations, goals, and so forth are not really one’s values if they are ill-suited to one’s affective nature.⁶⁷

Perhaps a self-assessment approach that requires, as Tiberius and Plakias do, that (a) welfare-constituting self-assessments be based on values (which cannot be ill-informed), or, as Sumner does, that (b) such self-assessments be autonomous, will be able to avoid the problems posed by manipulation. A version of this theory that also requires that (c) the values on which the assessments are based be affectively appropriate will presumably not overestimate the well-being of a person whose explicit self is out of sync with her emotional self.

It is less clear that these modifications (a)-(c) will help with the manipulability of self-assessments due to cognitive biases, for even self-assessments guided by informed, affectively appropriate, stable values might be subject to the contrast effect. Suppose I value eating well. I may be quite satisfied with how my life is going with respect to this value if I

⁶⁴Valerie Tiberius and Alexandra Plakias, “Well-Being,” in John Doris (ed.), *The Moral Psychology Handbook* (New York: Oxford University Press, 2010), pp. 402-32, at pp. 421-27.

⁶⁵*Ibid.*, pp. 421-22; cf. Bratman, “Identification, Decision, and Treating as a Reason,” p. 197.

⁶⁶Tiberius and Plakias, “Well-Being,” p. 422; cf. Gaus, *Value and Justification*, p. 109.

⁶⁷Suppose a person claims to value being a college teacher. Suppose, though, that teaching students, conducting research, and serving on committees makes this person progressively more miserable and cantankerous. In this case, according to Tiberius and Plakias, the *prima facie* authority of the person’s self-proclaimed value is defeated by the fact that the activity is not affectively appropriate for them.

have recently pondered the diverse, high-quality grocery stores and farmers' markets in my city. But if I then sit down to watch a television show about Mediterranean cooking with long shots of colorful, appetizing dishes prepared by professional chefs, I might produce less positive self-assessments. I am not sure that it is plausible to think that my well-being could be much affected by viewing a cooking show on television.

Another problem is that a person might be sufficiently unreflective that she never takes stock of her life, so that she simply has no beliefs about how satisfying her life is. Although admittedly unlikely, this sort of scenario is at least possible. Are we to say that such a person is necessarily doing badly? Or that she has a well-being level of zero? Or that there is no fact of the matter about her well-being? It is not even clear how a self-assessment theory ought to be applied to such a case. The versions of the theory that we have considered would each need to be extended in some way.

A more serious problem is that neither of these sophisticated forms of the self-assessment view contain any conceptual machinery for evaluating welfare over time, that is, for estimating how well a person is doing over a period of days, weeks, or months. To supplement the above theory so that it is capable of evaluating life-segments, we must add to it some clause such as this: a person is doing well to degree d over a particular period of time just in case the *average* degree of (judged) satisfactoriness over this period of time is d .⁶⁸ But this extension is not unproblematic. Suppose that there are several months during which a person repeatedly assesses her life as quite satisfactory. Then, some years later, this person comes to believe (she might even say, "recognize") that she was actually doing quite badly during those months. The version of the theory just described implies that the person's later judgments are simply mistaken. And if a person can be mistaken about her well-being at some point in the past, why should we not also acknowledge that she might be mistaken about her own welfare in the present? Once the self-assessment theory is expanded to deal with welfare over time, one of the basic rationales for the theory—viz., that a properly informed and autonomous person cannot be mistaken about how well off or badly off she is—is undermined.

The only way that I can see to avoid this implication is to say that facts about how well a person is doing over a period of time—as well as facts about an individual's degree of well-being at past times—must all be relativized to the instants at which the person assesses her life. On this view, there is no aperspectival fact of the matter about how well off or badly off

⁶⁸This is a simplification. There are an infinite number of instants in any life-segment; some more sophisticated analysis involving integrals is required.

a person is at a point in time. A person could be both doing well (from her current self's perspective) and badly (from her future self's perspective) at the very same time. It suffices to say that this approach would require us to revise quite a lot of our ordinary thinking about welfare.

Finally, if the self-assessment approach—with modifications (a)-(c) in place—is true, then it is not possible that a person be mistakenly dissatisfied (or mistakenly satisfied) with her life, provided that her self-assessment is informed, autonomous, and based on stable, affectively appropriate values. And yet even if a person is evaluating her life with a view to her stable, affectively appropriate values, it seems that inappropriate negative emotions might still corrupt her self-assessment. People sometimes get “down on themselves” or “down on the world” so that they assess their situation quite negatively, even though they are basically successful in realizing their values and quite happy the majority of the time. It is not clear in these cases that more information or greater autonomy would necessarily help them regain proper perspective. Even if their current negative affect is bad for them, it seems that we can correctly—that is, truthfully—say to such people: “Hey, cheer up! You’re actually doing quite well, on the whole!” And if this is so, then even the modified self-assessment theory may not satisfy Sumner’s criterion of subjectivity, for it makes a person’s well-being depend more on her current emotional state than on the things that she really cares about.

Well-being as agential flourishing does not face these problems concerning manipulability due to cognitive bias: whether you are realizing your values (and are stably disposed to do so) is a matter of fact, not opinion. It can generate estimates of well-being over time: to the degree that a person’s life, over a particular period of time, resembles the paradigmatically flourishing agent’s life, it is going well for her. It can explain the welfare of people who have not paused to assess their own lives. It both allows for mistaken self-assessment and illuminates the grounds of correct self-assessment. For these reasons, as well as the reasons set out in earlier sections, I think it is the most promising subjective theory of well-being currently on offer.⁶⁹

Jason Raibley

Department of Philosophy
California State University, Long Beach
jraibley@csulb.edu

⁶⁹I thank Fred Feldman, Hilary Kornblith, Phillip Bricker, Valerie Tiberius, Michael Rubin, Alex Sarch, Marcy Lascano, Teresa Chandler, Alan Goldman, Noah Lemos, Hud Hudson, Ned Markosian, Erik Angner, Harold Kincaid, Gregory Pence, and Sam Wren-Lewis for useful feedback on an earlier draft.