Well-Being: A Millian Hybrid View

Dan Haybron, Saint Louis University[1]
Draft: April 28, 2021

[This is an extensive revision of a 2016 paper, part of which split off into a forthcoming publication and which is currently being fashioned into a pair of chapters of a book ms, *Good Lives*. So it is now a bit of a Franken-paper. While there may be some rough edges and references are very incomplete, it should get the basic arguments across well enough. Please do not cite without permission.]

1. Introduction

In thinking about well-being, both lay and philosophical, two sorts of good seem recurrently to turn up as candidate elements of human welfare: pleasure, and nature-fulfillment. That pleasure has some sort of fundamental role in well-being seems especially hard to avoid when we think about its opposite, suffering. Paradigm cases of extreme suffering seem obviously to constitute a detriment to welfare, and to do so for a distinctive sort of reason: suffering appears to matter because of *what it is like* to suffer. The badness of suffering has to do with its phenomenological character, or at least that's how it seems. Relatedly, the goodness of pleasure, of pleasant experience, also appears to be grounded in its character *qua* conscious experience. Call this "the phenomenological intuition." Many philosophers have denied this appearance, and perhaps they are right to do so. But it seems to me bizarre to deny that there is at least the appearance.[2]

The apparent goodness of nature-fulfillment needs a little more explaining, but can be crudely summarized in a pair of popular slogans: "be yourself," and "be all you can be." The second slogan essentially says, don't just waste your life lazing about doing nothing, learning nothing, going through life with the mind of a child and the passivity of a cow. Realize your potential; let the seed come to full fruition. This sort of point is neither new nor peculiarly Western: we find it in Homer's depiction of the lotus eaters, in Hindu notions of self-realization, in Mencius' talk of cultivating human nature [Kim 2014], and in many other places. Nor is it particularly highfalutin': an American bow-hunting shirt exhorts us to "live life at full draw."[3] I know of no data on the question, but suspect that such sentiments are more or less universal: were you to show a photograph of the whacked-out denizens of a 19th century opium den to people just about anywhere, even were you to convince them that the experience was lastingly pleasant, you might well find widespread agreement that there's something sad, diminished, about such a life. In general, human beings seem to take pity on those whose lives they see as *impoverished*, and celebrate lives perceived to be rich, full, and fully realized. Liberal-minded moderns might often be loath to claim as much—and 'pity' is not a popular word these days—but this may owe mainly to values of tolerance and anti-paternalism, which can make us reluctant to pass judgment in such cases; but

---

[1] 2016: An earlier version of this paper was presented at the Kansas Workshop on Well-Being, 2015. I am grateful to the other participants, especially Connie Rosati, for feedback. I've only added some references, at the last minute (in the midst of changing to a new reference management program), but at least mentioned the names that have come to mind and welcome suggestions for others.

[2] For more discussion, see (Haybron, 2001; 2008b; 2016a).

[3] I have discussed nature-fulfillment ideals at some length in (Haybron, 2008a; 2008b), (Haybron, 2016b), (Haybron, 2013).

when thinking of our own lives, or our children's, there probably aren't many of us who would really be pleased at the prospect of whiling life away prone, on the sofa.

The first slogan, "be yourself," invokes a somewhat different sentiment but still, I would suggest, relates to a fundamental concern for nature-fulfillment. The proximate concern appears to be for *authenticity*, the idea being that how we think, act and live should reflect who we truly are. Concerns about inauthenticity are often expressed in relation to mindless conformity, lives spent fitting someone else's idea of who one should be or how one should live, self-deception, and so forth. In the modern era they have also become rather poignant in relation to various technologies for mental adjustment or self-enhancement, antidepressants being a familiar case, but recreational drugs and science fiction scenarios as well (*soma…*), and the prospect of drugs or treatments to enact moral or cognitive improvement worries many—will the resulting happiness or altruistic behavior, for example, really be authentic? One might think authenticity a peculiar concern of individualistic Westerners, and again I do not know of data on the question, but there is no reason why members of more collectivist cultures should lack concern for whether, say, Granny is really being herself. A collectivist, relational self can still be an authentic self, or so it seems. Were you to propose to the folk of Papua New Guinea that they might arrive at more agreeable children by running them through a brain-scrambling device that "improved" and homogenized their personalities, leaving them unrecognizable as the individuals they started out with, I suspect you'd get little consent. I could be wrong, but it seems plausible that people generally like the people they care about to actually *be* the people they care about—the authentic thing. More to the point, it is doubtful that many of us would envy someone who, though happy, is happy only because her personality has been surgically or pharmacologically altered beyond recognition. It might seem, rather, as if the person has died and been replaced by someone else. There's an appearance, at least, that authenticity matters for well-being, and concerns for authenticity seem naturally to cluster with the notion of nature-fulfillment: it is important to be yourself, arguably, because you can't realize your potential or achieve self-fulfillment, say, if you are inauthentic. Authenticity matters, if it does, because it is essential to living in accord with who one is.

Again, many philosophers would reject these appearances, and again they might be correct. But I think they are the appearances nonetheless. I'm not sure we can plausibly refer to a single intuitive appearance here, so let's just call these the "eudaimonic intuitions"—where 'eudaimonic' refers to the teleological ideal of nature-fulfillment that arguably characterized theories of well-being in the ancient eudaimonist tradition.[4] The basic schema for a nature-fulfillment theory is that it grounds a conception of well-being on some understanding of a person's or organism's nature: some sort of ideals or goals are implicit in the individual's nature, and well-being—often rendered as 'flourishing' in this genre—consists the achievement of those ideals. For instance, perfecting or fully realizing one's nature as a human being, leading a fully human life, say.

The phenomenological intuition and the eudaimonic intuitions seem to me among the most compelling intuitions about the substantive content of well-being: a theory of well-being that is going to comport at all well with our experience of well-being shall need to give each their due, and theories that fail to do so will do so only at great cost. They are not likely to seem terribly consistent with our lived experience of well-being, of what matters in human life. Or, for that matter, in other kinds of life: many of us feel uneasy about keeping a tiger caged, a cat indoors, or breeding a dog to be a stunted little version of a hunting animal. Eudaimonic intuitions inform ordinary thinking about the living world quite generally.

[4] I generally use 'eudaimonistic' to refer to a type of theory or school of thought, and 'eudaimonic' in other contexts, paralleling common usage of 'hedonistic' and 'hedonic'.

One might have noticed that the two sets of appearances correspond roughly to two of the three main accounts of well-being, according to the standard Parfitian taxonomy: hedonism and objective list theories—the latter being a poor name for a loose conglomerate of views whose most popular members, by a wide margin, are Aristotelian accounts, which are not mere lists at all. What, then, of desire theories, which in recent times have been the most widely espoused? I will not argue for it, but desire theories seem mainly to be motivated, not by our everyday experience of well-being, but by theoretical considerations: a wish to respect liberal ideals of agent sovereignty, for example, or to forge an essential link between well-being and the individual's motivation. It seems to me that only a theory, or some abstract reasoning at the least, could lead one to think it plausible that the mere fact of wanting something—say, to eat a lump of gravel[5]—thereby makes it, to any extent, good for you. In fact the experience of parenting should tend to lead one to the opposite conclusion: people, especially the little ones, want all sorts of stuff, much of it in no way good for them. The attractions of desire theories are not inconsiderable—in fact the view I defend here incorporates a kind of desire component—but it is not so clear they answer to any very powerful sense of what matters in particular cases. Their attractions, like those of consequentialist morality, seem mainly to be theoretical. At any rate, I will set them aside here and focus on the pull of the hedonic and the eudaimonic.

The problem is, both appearances seem undeniable. Moreover, we're not talking moral theory here, but well-being: it is one thing to embrace a counterintuitive morality if it means, say, promoting human interests; at least people get something out of it. It is something else to embrace a counterintuitive view of well-being that effectively says, perhaps to most of humanity, that some of the deepest, most pervasive and abiding convictions about what matters in our lives, and the lives of those we love, are nonsense. And in exchange we get, well, something that sounds better in theory than in practice.

Yet what theory can encompass both appearances? They are quite different, and seem not to admit a unified treatment. Nonetheless, I will make an effort here to try to reconcile them, in a view I will call a *Millian hybrid* view. The view is so-called because, while very different from Mill's view in certain respects, it arguably mirrors the dual nature of his approach to well-being, which perhaps was motivated by a recognition of just the sorts of appearances I am concerned with here.[6] The account sketched here—and I can hardly do more than circumscribe it in this paper—builds on a partial account I developed in earlier work, but extends it most notably by making it, effectively, a kind of "list" theory of well-being, with two items: pleasure and self-fulfillment, where self-fulfillment is a form of nature-fulfillment that grounds the theory of well-being in some conception of the self—of who one is.

Like most philosophers I am not fond of lists; but perhaps they are not so bad as long as they are principled, and not ad hoc. In fact I think it is possible to give a principled basis for the list, explaining why it has just these two elements. But this basis lies in a certain story about the metaphysics of value, and moral psychology. The story is broadly Humean in form: crudely, what matters depends on what human beings take to matter; and in the realm of prudential value, human beings have two distinct sorts of concern—embodied partly in the distinction between sympathy and pity—to which correspond the two fundamental elements of well-being. Roughly, the idea is

---

[5] (Keller, 2009)

[6] To keep the discussion manageable I will not discuss the interpretation of Mill's views much here, trusting that the relevant ideas will be recognizable enough to readers (I do say a bit more in {Haybron}). [From RG: look up representative gov't 1st 5 chapters. Logic on social sciences, The science of ethology; art of living. Alan Ryan's (?) book about Mill]

that humans have one set of welfare concerns relating to pleasure and suffering, and another relating to concerns about personal development, like those arising in raising children. In short, the metaethics of well-being lends support to the Millian hybrid view in normative theory. To make good on these suggestions, however, requires an excursion into territory best saved for another occasion; and I expect that the empirical evidence needed to give them solid grounding isn't yet available, making them more speculative than would be helpful in this paper.[7]

However, it is worth pausing briefly to note one consequence of this sort of Humean approach for our project: it is liberating. Many have objected to nature-fulfillment views, for instance, on account of their teleological character: they are, allegedly, metaphysically extravagant. As many authors have pointed out, however, this need not be the case.[8] What seems less-often appreciated is how easy it is, at least in principle, to defend just about any normative position from a Humean standpoint: the norms such a view can sustain are as diverse as the responses human beings can have to things. If we project value in a desire-theoretic way, then we get a desire theory; if we do so in a hedonistic manner, then we get hedonism; and if we project value in an Aristotelian way, then we get Aristotelian norms. Pretty near anything goes. Moreover, the approach appears to leave all options on more or less equal metaphysical footing: the teleological projection of value seems no more or less mysterious than the preferentialist projection of value.

I'm not sure why more theorists of a eudaimonistic bent haven't availed themselves of a Humean metaethic. One reason may be that some sort of relativism can be hard to avoid given plausible variation in human sensibilities; this doesn't trouble me—partly because I think enough of the important bits are held enough in common—but it does others. Another reason may be that Aristotelian approaches are often thought to have a certain "naturalistic" metaethic built in. Maybe they do, but I've never understood how; Aristotle's own view seems near enough to Platonism, or Plato Lite at any rate, to me. The only thing naturalistic about his view, if I understand it correctly, is that it takes norms to supervene on facts about human nature. But nobody worries about whether values supervene on facts; the deep worries concern what values are and what determines their distribution in the world, and whether an answer can be given to those questions that doesn't look spooky from the perspective of, say, modern science. Supervenience-mongering alone, which Platonists can trade in as happily as anyone, does nothing to help with those questions.

No doubt the Millian hybrid view has problems. But I don't think a reliance on naturalistically suspect metaphysics is one of them. I emphasize that what follows will be programmatic and conjectural, with little effort to address likely objections. My hope is simply to paint a picture that, on the whole, seems plausible, and a fruitful basis for further inquiry.

2. Self-fulfillment

2.1. Introduction

In what follows I will assume a basic familiarity with philosophical theories of well-being, with typical candidates including hedonism, desire theories, (objective) list theories, and eudaimonistic or nature-fulfillment theories such as Aristotle's that see well-being as a matter fulfilling one's nature. Self-fulfillment is a species of nature-fulfillment on which the relevant nature is the character of the *self:* of who one is, which in turn is typically understood as a matter of

---

[7] The previous draft of this paper contained a lengthy discussion of these issues, which I have cut here and plan to place in a separate paper.

[8] E.g., [Annas, Nussbaum, McDowell, Russell, LeBar, Hursthouse…].

the particular characteristics of the individual, as opposed say to the generic qualities of the species to which they belong. Self-fulfillment views thus couple the theory of well-being to that of the self.

In earlier work I argued that the intuitive value of happiness, understood along the lines of an emotional state theory, is best explained in terms of self-fulfillment: well-being consists at least partly in self-fulfillment, which in turn consists at least partly in (authentic) happiness—the fulfillment of one's emotional nature. The assumes a view of the self as partly defined by one's emotional nature, understood in terms of what sorts of lives make one (authentically) happy. We can group the points mobilized in favor of this conclusion together as the *argument from happiness* (Haybron, 2008b, 2008a). A more recent argument, *the argument from authenticity*, defended the significance of ideals of authenticity for theories of well-being, suggesting that this in turn supports a self-fulfillment approach (Haybron, 2019). Still more recently, I marshalled some general considerations for thinking that theories of well-being need to self-aware, as it were, clarifying the relevant concepts of self, arguing that the notion of self is indispensable, and neither obscure nor metaphysically problematic, and giving two further arguments for a self-fulfillment approach.[9] First, *the convergence argument:* most major philosophical accounts of well-being—list views excepted—face internal pressure to converge on one or another form of self-fulfillment theory. Second, *the piggybacking argument*: for many philosophers, other theoretical commitments regarding the self exert give them strong reasons to accept some ideal of self-fulfillment in the domain of well-being. In a manner of speaking, all roads lead to self-fulfillment. Here I offer a further argument, *flourishing in other lifeforms*: that is, the self-fulfillment view is a plausible development of the approach to nature-fulfillment that makes most sense when looking across the living world. In total, the case for self-fulfillment as at least one major element of well-being rests on five arguments (though one might add to these, as background for the rest, the above-mentioned argument for the general need to embed notions of the self in ethical theorizing):

1. The argument from happiness
2. The argument from authenticity
3. The convergence argument
4. The piggybacking argument
5. Flourishing in other lifeforms

I offer a summary or restatement of the four extant arguments in the following section. While I add somewhat to prior discussions, readers already familiar with these arguments can likely skip to section 2.3. In that section I elaborate the self-fulfillment account I favor, drawing on a dual-aspect model of the self; the overall plausibility of that model, and of the resulting theory of well-being, helps bolster the case for the present view. Section 2.4 develops the final argument for a self-fulfillment approach.

2.2. Well-being as self-fulfillment: the first four arguments summarized

1.1.1. The argument from happiness

While ideals of self-fulfillment have been around for some time, they have not often been explicitly distinguished as a species of nature-fulfillment theory, nor hitched to theorizing about

[9] [Haybron forthcoming RRR]. This paragraph and the one preceding are adapted from a passage in that paper.

the self. The present inquiry has its roots in puzzles about the value of happiness, specifically as conceived along the lines of an emotional state theory (Haybron, 2008b, 2008a). This view seems not to have been distinguished from hedonism previously, which may be why the issues only came to light recently. Collectively, they give rise to the *argument from happiness.*

The relevant view of happiness trades on a distinction between what I've called "central" and "peripheral" affects, which intuitively corresponds to the difference between affects that do, and do not, implicate one's emotional condition—for instance, being depressed versus experiencing the sensory pain of a backache (Haybron, 2005, 2008b). Roughly, central affective states include moods and emotions—the latter counting insofar as they themselves are mood-constituting—while peripheral affects include at least sensory pleasures and pains (as such), as well as "notional" affects like being mildly pleased or annoyed, or the constant stream of pleasant/unpleasant experiences one has in response to ordinary stimuli, as when passing by an attractive versus unattractive house. "Mere pleasures" are good in their way, namely hedonically, but they seem to matter for quite different reasons than emotional fulfillments like, well, feeling fulfilled or being in good spirits or being at peace. As these examples suggest, common sense and philosophy have long charted the distinction in various ways, also in such injunctions as not to let minor pains or annoyances "get to you," "bring you down" and so forth. The distinction has been crucial for many philosophical and spiritual disciplines, notably Stoicism and Buddhism, both of which would have us strive to eliminate emotional distress, but not of course pain, from our lives, for instance distinguishing the one "arrow" or "dart" of pain—say a toothache—from the second, say of being aggrieved about it. The one is unavoidable; the other is at least thought to be optional. In fact perhaps every extant ethos regarding virtue accords some importance to managing our emotional conditions—it is among other things a hallmark of maturity—whereas I suspect no school of thought has ever suggested we are to blame if we feel an unpleasant sensation of pain when a wasp stings us. By wide if not universal agreement, what makes us happy or unhappy is a matter of critical importance ethically speaking[10]; what feels pleasant or unpleasant to us is far more a matter of debate, and insofar as it matters, it isn't as a sign of our characters.

As these considerations suggest, the central/peripheral distinction—the distinction between happiness and mere pleasure—relates to matters of the self: one's mere pleasures seem not to reflect on who one is, whereas what makes one happy or otherwise—what impacts one's emotional condition—does. If an insult does not affect your happiness, there appears to be a literal sense in which it does not get to you or bring you down; *you* are unshaken, unbowed, and just shrug it off. The self remains as it was. (We should not let the ease of making these points without deploying the expression 'the self' fool us into thinking they are about anything other than the self: if you think a person's character matters and is evinced by emotional but not sensory responses in anything like the ways just noted, then you are committed to some conception of the self, as surely as any Cartesian dualist.[11])

If these reflections are even roughly on the mark, it would be quite odd if they did not ramify for our understanding of well-being. How could the significance of happiness and pleasure differ in these ways, for instance only one being crucial for virtue, and not also in how they impact our well-being? (Note also that a tempting hedonic explanation of happiness' importance—that is

[10] Apparent exceptions like Kant may often owe to dubious readings of their views. While Kant was no Aristotelian about the role of (emotional state) happiness in virtue, he would not likely be taken seriously if he thought it unimportant to cultivate appropriate emotional responses. An adult given to the emotional regime of a toddler would not pass muster on any serious ethical framework I'm aware of.

[11] This example should be added to the piggybacking argument described below, but was not in the original paper.

is especially pleasant, while merely liking the appearance of a house is not—won't work at all here: orgasms, among other things, are *intensely* pleasant, but for all that, they can be mere sensory pleasures. But they can still leave you cold, or worse.) At any rate, this should be sufficient background to grasp the argument from happiness, which is really a cluster of arguments.

To begin with, the case is intuitive (and was introduced with two examples, of Henry and Claudia). Individuals whose occupation or way of life leaves them unhappy—anxious or depressed or just spiritually deflated, say—tend to give rise to distinctive worries about their well-being: not that their way of life is unpleasant, but that it doesn't *suit their nature*. It doesn't fit with who they are. And in fact if they remedy the situation by finding a new occupation that makes them happy, fulfilled, friends might remark that "she's really come alive and is truly herself now." This is very different from how we think about, say, construction work that results in physical discomforts. If taking analgesics relieved the pain but made the (fairly Stoic) worker no happier, it would be surpassingly weird for family to offer remarks along the lines of, "he's finally himself now." These sorts of cases suggest that happiness' value is not merely hedonic, but has something to do with the self. The best explanation, I have argued, is that well-being is at least partly a matter of self-fulfillment—very crudely, a life in accordance with who you are—where the self is at least partly defined by one's emotional nature: what characteristically tends to make one happy, versus unhappy. While much of the argument focuses on the intuitive value of happiness, a distinct part of the case centers on the character of hedonic value, which appears to be quite different from the way happiness matters in relation to the self. For instance, Henry who languishes in an unsuitable line of work might feel better if given a really effective pleasure pill, essentially a spiritual analgesic; but while that would have some benefit in making his experience more pleasant, it hardly solves the more salient problem, namely that his way of life doesn't fit his nature. Perhaps it makes him worse off.

An important qualification is needed here, namely that happiness yields self-fulfillment only insofar as it is *authentic*, in roughly the manner elaborated by L.W. Sumner in his theory of well-being as authentic happiness (with life satisfaction in the place of emotional state happiness). That is, one's happiness must be a response to *your life* that is genuinely *yours*—i.e., it must be informed and autonomous, and not the result of deception, manipulation, brainwashing, etc. While the details of our conceptions of authenticity differ, the main departure my view makes is adding a *richness* constraint. Intuitively, one attains a higher degree of self-fulfillment to the extent that one's way of living more fully expresses who one is, one's personality or character, and this seems to favor richer ways of living than, say, counting blades of grass all day long. This is a major theme of self-fulfillment ideals in humanistic psychology, for instance, which focus on actualizing one's potential, and allows the theory to make better sense of a core eudaimonic intuition. (While it seems plausible that a richer life is more authentic, it may be that richness is an independent constraint on well-being.)

1.1.2. The argument from authenticity

The inclusion of authenticity yields a further argument in favor of the self-fulfillment view, namely that it better explains the normativity of authenticity than the subjectivism that motivates Sumner's account. As multiple commentators have noted, authenticity is not easily reconciled with subjectivism, as it seems to have the effect of reducing individuals' authority about their lives, imposing an objective constraint on their happiness. To the inauthentic person, it may certainly seem as if the theory is doing that: "who are you to say that these are not my authentic values?" It is not easy to explain the significance of authenticity in subjectivist terms, but quite natural to do

so in terms of self-fulfillment, which readily lends itself to the idea that one's true self may differ from one's manifest self.[12] The distinctive value of happiness seems dependent on its being authentic, then, the value of which in turn is best explained within a self-fulfillment framework. And, independently of authenticity, happiness' value seems best explained in terms of self-fulfillment.

The *argument from authenticity*, elaborated further in [RRR], draws on various intuitive considerations to make the case that authenticity is tacitly regarded as an important aspect of human well-being—and of other goods such as art and nature—perhaps across all cultures, as the Granny example above illustrates. A concern for authenticity appears to underlie a wide range of concerns, for instance widespread revulsion at the use of lobotomies or other extreme manipulations to alter a person's personality. See, for instance, *One Flew Over the Cuckoo's Nest*, where viewers are invited to endorse Chief's euthanizing of McMurphy post-lobotomy, and most seem readily to accept it. The horror of what was done to him does not hinge on any thought that life would be unpleasant for him—perhaps the reverse—or that he would not have wanted it (which might well have made it even more terrible, like Winston in *1984*). And, as hinted in the example of the construction worker, we feel differently about the use of analgesics for physical pain than about mood-altering medications like antidepressants, and the range of attitudes people evince about these drugs is revealing. On the one hand, whatever relief from pain or suffering they bring is a positive, hedonically; but on the other, drugs like antidepressants are seen variously as either reducing ("uplift anxiety") or—more often, I think—enhancing ("I'm finally myself") authenticity, this playing an important role in the perceived desirability of taking them. No one frets about whether they'll still be themselves if they take an aspirin.[13]

1.1.3. The convergence argument

The *convergence argument* shows that each of three major alternative approaches to well-being—hedonism, desire theory, and Aristotelian nature-fulfillment—all profit from a tacit association of their core ideals with the notion of self-fulfillment, so that the appeal of self-fulfillment underwrites a substantial part of their allures. Indeed, it appears that all either rest on some understanding of self-fulfillment or face internal pressures to collapse, at least partly, into a self-fulfillment theory. Apart from list theories, which infamously lack any core ideal, all roads seem to lead to self-fulfillment.

The shape of the argument regarding hedonism should not be hard to guess. But in brief, the pleasures that make hedonism seem most compelling as a conception of welfare are the pleasures of happiness—emotional, and not merely sensory or notional pleasures. And it does not seem possible to explain this merely in terms of quantity of pleasure, as if hedonism would still seem persuasive if a person could be wired to enjoy a never-ending stint of sensory ecstasy on an Orgasmatron (see *Sleeper*). Rather, there seems something distinctly important, beyond the hedonic,

---

[12] This is not to say subjectivists can't adopt self-fulfillment views, and perhaps the value-fulfillment component of my own view is subjectivist, though if so its authenticity requirement raises the same issues that came up for Sumner. But subjectivism takes a range of forms, some like full-information theories notoriously raising similar difficulties about their subjectivist credentials, and plausible subjectivisms will somehow need to place limits on agent sovereignty given the obvious phenomenon of mistakes. I am doubtful that any kind of subjectivism can reconcile the various demands on such a theory of well-being, some sort of subject-dependence being the closest we can reasonably hope for. For that matter, I'm not sure subjectivism about almost anything is true.

[13] My dentist reports that some patients refuse anesthetic for dental procedures, and not apparently for fear of needles (versus drills?). Whatever the rationale—I suspect not a rational one—it seems unlikely to be a fear of identity crisis either, nor a surfeit of consideration for the poor dentist, who appears to be somewhat traumatized by doing this to people.

about our emotional conditions—and this leads us right back to the argument from happiness. Reflection on the allures of hedonism leads us to a self-fulfillment account (and further, we shall see, a Millian hybrid account, as there remains a considerable hedonic residue that cannot be explained in terms of self-fulfillment).

In the case of desire theories, there is reason to think that the strongest motivation for such views is a concern for non-alienation, or what's often called the "resonance constraint" (Railton, 1986). What's good for us should not be alien to us, and a natural understanding of this idea is that one's good should connect with one's cares or wants. The problem is that we can be alienated even from many of our desires, as in the case of a smoker's cravings: many of the things we want, we don't value, and even wish we did not want. To insist that smoking is nonetheless good for the person, despite going against some of his deepest cares, seems to make a mockery of the subjectivist ideal, essentially railroading the person's evaluative perspective on his life. Even to formulate this thought seems to embroil us in a distinction between desires that do, and do not, issue from the "true self"—again illustrating just how inescapable notions of self are in our practical reasoning. And indeed it is no coincidence that popular theories of the self often focus precisely on the sorts of desires that a desire theorist would naturally advert to in addressing problems of alien desires: higher-order desires or values, for instance.

It seems, then, that desire theories face considerable pressure in the direction of a self-fulfillment view, on which the self is defined by certain of its desires, which among other things embody one's sense of what is *worth* wanting. I think these are most plausibly understood as the person's values: who we are is defined partly by what we care about, or value. And if this is right, then self-fulfillment consists at least partly in value-fulfillment.[14] Such a move readily accommodates important cases not obviously explained otherwise, such as the lasting harm of losing a spouse, which may persist well after the survivor's emotional recovery: still, there may be a sense that she is no longer whole, that part of her is missing, because the one she cared so much about is longer with her. Her spouse had become part of who she was, and that in itself is part of the explanation of her loss.

For the same reasons as before, an authenticity constraint applies: the fulfillment of machine-manufactured values, say, does not benefit us in this way. This line of reasoning is so natural that one might wonder why it hasn't been made explicit more often. Most likely, because there are other ways to make the case, and philosophers have been unfortunately averse to invoke the troubled concept of the self if they don't have to. (As well, the restriction to values raises difficulties about accounting for things that might seem good for us even if they don't quite speak to our values. If this is a problem, it will be less so in the context of the Millian hybrid defended here, which has other ways of accommodating such cases.)

Finally, nature-fulfillment theories in the Aristotelian vein ground the individual's well-being in species norms—in species nature—and this in itself may be cause for worry, for instance because of alienation problems, and perhaps motivate a refocusing of the theory on the particulars of the individual's makeup, yielding a self-fulfillment theory. I think there are good reasons for Aristotelians to take seriously such a reorientation of the approach; that perhaps they have misunderstood the options available within the basic eudaimonist framework.

But another, less revisionary option is available to them, and indeed at least this seems hard to escape: namely, to take nature-fulfillment to consist at least *partly* in self-fulfillment; not only in exemplifying goodness in human living, but also in living as uniquely suits *your* nature. To deny this may in fact be incoherent, at least if one has already adopted some view of the self. For

[14] See also (Raibley, 2010, 2013; Tiberius, 2018).

example: an Aristotelian who addresses problems of freedom and responsibility using something like Frankfurt's higher-order desire model of the self, but who then deemed it irrelevant to the individual's flourishing whether his higher-order desires were fulfilled, would be saddled with a combination of views that, if not incoherent, is at least bizarre. To wit: what's good for you is fulfilling your nature; but the nature of your self—the characteristics that define who you are—has no bearing on the question of what's good for you. In essence, your self is no part of your nature, in the relevant sense. As an understanding of the nature of a human being, presumably a person, this would be an eccentric position, to say the least. If not that, then what *could* be part of a human being's nature? One might have thought that, somewhere in the story of the distinctively human form of life, would be that is the life of a person, with a personality—that is, a self. If, as seems abundantly plausible, a person's self is part of her nature, then nature-fulfillment ought to incorporate self-fulfillment as at least one aspect of it. Aristotelian readers not persuaded of the Millian hybrid, then, might still have good reason to incorporate the conception of self-fulfillment offered here as part of their theory.

1.1.4. The piggybacking argument

Finally, part of the *piggybacking argument* has already arisen here: insofar as you are already committed to some conception of the self, then your theory of well-being needs to cohere with it; the theory of well-being rides piggyback on one's view of the self. Perhaps you have a view of autonomy, or responsible agency, or the political implications of identity, or character, or of alien versus non-alien desires, or some other view involving commitments about the nature of the self. If so, then it would be very odd to adopt a theory of well-being that clashed with those commitments, so that what's good for you could be entirely at odds with who you are. If you think selves are defined by people's values or higher-order desires, for instance, then it would be strange indeed to take up Benthamite hedonism in the domain of well-being. In that event, a vigorous anti-hedonist like Nietzsche's good, the life of pleasure, would be utterly contrary to who he is. This is an unseemly pairing of ideas, at best. Note that a congenial side benefit of the piggybacking argument is that it highlights the discipline that can be applied in theorizing about self and self-fulfillment: we have diverse concerns about the self, and it may be that different aspects are relevant in different domains—e.g., well-being versus morality—but it should at least be the case that our views of the self in one domain hang together reasonably well with our views in others. It is not just "anything goes" when it comes to theories of self.

2.3. A dual-aspect model of self, and self-fulfillment

I have argued that well-being consists, at least partly, in self-fulfillment, but I have left it somewhat open what exactly that entails, suggesting that both emotional nature-fulfillment and value-fulfillment views have merit. So, which is it? Arguably, both. In earlier work I gestured at a dual-aspect model of the self, and accordingly of self-fulfillment, with emotional nature-fulfillment as one component, the other being something to do with "identity," or what we care about.[15] The basic idea draws on an ancient model of the self as having two sides: crudely, a rational side,

[15] (Haybron, 2008a; 2008b). There I raised the possibility that the relevant notion of identity was something like a narrative self, with success perhaps consisting in "narrative role fulfillment." It now seems to me that narrative has more to do with a person's self-conception—roughly, a description of oneself—than with defining the self proper, where the latter notion corresponds to the traditional idea of the "soul" and thus is a psychological entity rather than something like a description. Narrative and self-conception play important roles, perhaps, in influencing our choices, but they are (at least for present purposes) products of the self, not the self itself.

and an emotional side. As many commentators have observed, this dichotomy is a little oversimplified: the emotional has rational dimensions, and vice-versa. The precise character of this divide is a bit obscure, but it is possible that what I have called central affective states, such as moods and emotions, are distinguished from peripheral affects like mere sensory pleasures in part by their aptness for engagement or enmeshment with rational processes. While moods like anxiety are sometimes free-floating and at least somewhat independent from any sort of conscious judgment, states of those types are nonetheless apt to be influenced by our judgments, and even to incorporate them. Whereas physical pains, say, do not seem like that, which is one reason we judge people's characters by their tendencies to have negative emotions, but not by their tendencies to feel pain.

Our values, by contrast, relate more strongly to rational processes, the popular paradigm of a value being what you reflectively affirm. But values seem not merely to involve rational processes, and it is standard to include an emotional component in accounts of valuing: wholeheartedly affirming something, for instance, or tending not just to make certain judgments about it but also to feel in certain ways about it—feeling ashamed or guilty when you fail to measure up, for instance. In fact, it is arguable that "emotional" processes count for more than "rational" processes in determining what we value, as for instance in Huck Finn-type cases where our best judgment (justice demands turning in the slave) is out of sync with our emotional responses (that judgment rings hollow and "feels wrong"), which seem better to represent what we really care about.[16] If so, then even the "rational" self may be more emotional than rational. Perhaps we should conceive of the self as, in two different ways, relating to matters of what we colloquially call "the heart": crudely, and at the risk of sounding a saccharine note, who you are is where your heart is. This, in turn has to do partly with what makes you happy, and partly with what you care about.

We can get the same result, I think, from a different direction: by thinking about the different functional roles that our values and emotional conditions play in our lives. On the one hand we are planners, or *agents*: prospectively, we must choose what to do. On the other hand we are *evaluators*: retrospectively, we must assess how things have gone. Arguably, the chief role of our values is to govern our deliberations; when all goes well, at least, what you care about defines what manner of agent you are, and how you choose to live. One might think that the retrospective task would also hinge on your values: whether you respond to your life as going well for you or not, and whether you are happy, will depend on how well it corresponds with your values. This is roughly the picture many economists take for granted, so that they shift automatically from talk of preference satisfaction to talk of happiness, as if it is obvious that getting what you want invariably makes you happy.

In fact this is obviously wrong, and I've argued for it at great length elsewhere, so I won't rehearse the arguments here.[17] From a biological standpoint, a person defined purely by her values will face a serious problem: she could develop all manner of disastrously nutty values; becoming a Shaker, say. We need a feedback mechanism to let us know when we're getting things wrong, despite whatever crazy ideas might have come into our heads. Our animal ancestors developed such mechanisms to look after physical needs, and accordingly feel physical discomfort when those needs aren't met. As rational social creatures, we need feedback not just about immediate physical demands, but about the bigger picture: are we socially well-situated, and otherwise arranging our lives in ways that make sense? If not, our psyches have ways of making us change our ways: the person with a penchant for abusive relationships gets depressed, for instance. In earlier work I've argued that central affective states function to govern our responses to the general

[16] E.g., (Arpaly, 2004).
[17] (Haybron, 2008b).

character of our circumstances, as opposed to immediate stimuli, but nothing here hangs on any particular evolutionary or biological story; these remarks are meant only to illustrate one reason it might make sense for human beings to have emotional natures that do not align perfectly with their values.[18]

The agent/evaluator distinction is not meant to imply that retrospective evaluation only involves our emotional conditions. Clearly, people also evaluate their lives by thinking about them; human agency isn't purely prospective, but is also involved in assessing the past and present. Indeed, reflection on how things have gone and are going is what we most commonly associate with "evaluation," and might be essential for living well.[19] So there is not a neat division of labor here. But valuing, as a species of desire or conation, seems primarily to have a prospective function, being most salient in governing our deliberations. Moreover, reflective judgments about how one's life is going may have less impact on how one proceeds than the verdicts of one's emotional self—how happy or unhappy one is. Forming the opinion that you're living badly may not have much effect if you're nonetheless quite happy. And, as noted above, such judgments may well fail to reflect one's values in any event; sometimes, your values may play out in your emotional condition more than in your reflective pronouncements about your life.

On the present view, then, self-fulfillment has two aspects: authentic happiness, and authentic value-fulfillment. Our flourishing involves living in accordance with both our values, and our emotional natures. Sometimes these aspects of the self will conflict, as with someone like Nietzsche who doesn't much value happiness. In such cases we get a mixed verdict: the individual's life goes well for him insofar as he succeeds in relation to his values, say; but it goes badly for him insofar as he is unhappy. This sort of ambivalence seems to me exactly right: there is a tradeoff in such cases, and ideally we should want the different aspects of the self to be consonant. If you have values that fit ill with your emotional makeup, your prospects for flourishing will be diminished. In some cases, you will be better off seeking the best fit with your values (e.g., a social worker who isn't particularly happy, but also not miserable and finds the work meaningful); in other cases, you will fare better doing what makes you happy (e.g., someone who values farming, which makes him miserable, and he would be much happier doing other work).[20] Sometimes a change in values may be called for—perhaps you simply value the wrong things, for instance giving too little weight to your own happiness. But not always.[21] It may be, for instance, that you should sacrifice your own well-being, say to care for your children by working at a grueling job.

Given the strongly intertwined nature of the two aspects of the self, it is not implausible that human flourishing will tend, in the ideal case, to involve the fulfillment of both aspects in a harmonious way: (authentic) *happiness in value-fulfillment*. Insofar as value-fulfillment is a subjective ideal—though I am not sure it is, given the possibility of error about one's values—then we might say "(authentic) happiness in a subjectively worthwhile life." This would, interestingly, yield a more individualistic, non-perfectionist counterpart to the ideal of well-being recently defended by Neera Badhwar: *happiness in an objectively worthwhile life*.[22] It may indeed be that the crucial differences between my own view and its Aristotelian relations—at least on the nature-

---

[18] .(Haybron, 2008b; Haybron & Tiberius, 2015)
[19] (Tiberius, 2008).
[20] See the cases of Henry and Claudia in (Haybron, 2008a; 2008b).
[21] An important question, which I set aside here, concerns whether all values count toward well-being, or just one's "personal welfare values," or some other subset of one's values. Perhaps values having no bearing on one's own life, or aimed at one's own ill-being, don't contribute to one's welfare when fulfilled.
[22] (Badhwar, 2014)

fulfillment side of things—can be roughly understood in terms of the differences between these two ideals.[23]

2.4. Flourishing in other lifeforms, and the human case

A fifth type of argument for the self-fulfillment approach is already familiar from the Aristotelian literature: to get a fix on human flourishing, start by looking at flourishing in other life forms, and extend the framework accordingly.[24] Aside from bolstering the case for the present theory, it is independently important that we can account for well-being in other species, and not just human beings. Plants make a natural starting point given the botanical shades of 'flourishing' itself, and ubiquitous metaphors regarding the development of seeds, acorns and the like into thriving mature plants. What does it mean for a plant to thrive? Plant flourishing, if it is anything at all, must be *physical flourishing*: fully realizing the organism's potential, or tendencies, to survive, grow, and reproduce. Its functioning proceeds unimpeded, not being thwarted or stunted for want of nourishment, space, or other needs. It is healthy and vital, at maturity more like the Kennedy Center Christmas tree than the Charlie Brown variety.

Animal life brings at least one major change: the introduction of agency, or least a fuller kind of agency than one finds in plants.[25] Accordingly, animal flourishing will plausibly involve the exercise of agency, and this indeed is how many of us think about well-being in other animals. For wolves, flourishing involves hunting, and doing so in packs. A wolf held captive, alone, in a zoo, strikes very many of us as a sad spectacle, even if it leads a comfortable and happy life. For the same reason, many of us prefer to let our pet cats roam outside, expressing their nature more fully than a housebound cat; or, if we keep them indoors, say to spare our avian friends from a buzz saw of feline death-dealing, it is only with regret. The standard metaphors of animal flourishing center on carnivores, interestingly, but even the seeming dull lifeways of herbivores can elicit these intuitions. Sea turtle agency doesn't strike me as terribly interesting, mostly just paddling around munching on turtle grass, coming up occasionally for air. But even if I haven't underestimated the beasts, it would be a sad sight to find a zoo exhibit of disembodied turtle heads, kept comfortable through drugs and sustained by tubes and wires, doing nothing but sitting there, perhaps in front of a screen depicting pleasingly grassy ocean bottoms. Their agency has been almost wholly thwarted, and their lives accordingly seem impoverished—reptilian equivalents of Homer's lotus-eaters.[26]

Now there are several different philosophical stories we could offer to explain such intuitions. Aristotelians, for instance, would spin these cases in terms of failures to exhibit the excellent activities characteristic of the species. I have argued elsewhere against the Aristotelian focus on excellence, and on species-based norms, and will not rehearse those arguments here.[27] For now I want to offer a different story—a variety of eudaimonism centering not on capacity-fulfillment,

---

[23] [This last paragraph is a last-minute addition, so I'm not certain what to make of it. But I'm inclined to resist the option of turning the slogan into a part of the account of well-being, as it seems unlikely that happiness in fulfilling one's values has distinctive value vs. the sum of happiness and value-fulfillment, taken as distinct goods.]

[24] E.g., (Foot, 2001; Kraut, 2007).

[25] In many cases it also adds hedonic value to the picture, which I discuss later.

[26] Some might doubt the intuitions canvassed in this paragraph are widely shared. It's an empirical question, but absent data I propose the "Pixar test": could Pixar reliably elicit such intuitions in a film aimed at a global audience? (Think *Wall-E.* Actually, the fact that such filmmakers' living depends on their ability to predict the reactions of diverse audiences suggests they might be uniquely qualified to judge what a genuinely inclusive "we" would find intuitive.) I would wager a few bucks that the turtle head scenario would play out as I suggest, anywhere you show it.

[27] (Haybron, 2007; 2008b).

but on *goal-fulfillment*: creatures flourish by succeeding in the goals that structure their makeup and functioning.[28]

In the case of animal agency, we note that animals have desires, or drives, that structure their activity, and they fulfill their natures, at least in part, by fulfilling those drives. This is essentially a variety of desire fulfillment, though I will leave open whether it is some subset of the animal's desires that sets the conditions for its flourishing. This picture differs from the Aristotelian schema, first, in eschewing reference to virtue or excellence: what matters in flourishing is success, not excellence—at least, excellence matters only insofar as it bears on success—a supposition that perhaps seems more plausible when regarding herbivorous animals. When I see a turtle happily munching on seagrass, "excellence" and admiration are not the first things that come to mind, or if they do, it is purely aesthetic: they are beautiful. If lupine excellence seems important for wolf flourishing, it may simply be because a bad hunter is not likely to be a very successful hunter.

It is interesting that Aristotelian theories are widely prized for their application to diverse species, not just humans. Yet, even as Aristotle takes a swipe at cattle in illustrating one kind of human unflourishing, he seems unable to offer a credible story about *cattle* flourishing. One can understand the attractions of thinking that cattle flourish by leading fully bovine lives, doing what cattle do. In some sense, they're succeeding in their goals as animals. Perhaps I am just blind to ruminant talent, but it is harder to see how a cow chewing its cud amounts to a kind of *excellence*. And harder still to grasp why we should explain what's good about the grassy life for a cow in terms of the putative excellence of activities like cud-munching.

A second difference is that it matters not a whit what is normal or characteristic of the species. If my cat lacked any interest whatsoever in venturing outdoors and hunting, I would not think it sad to keep it indoors, at least if I were convinced that it was still a healthy cat. Ferdinand the bull, again, seems to thrive insofar as he follows his impulses; that other bulls prefer to fight is neither here nor there. Now Aristotelians observe that we might think there's something wrong with a wolf that won't hunt, or isn't social, or free rides off of his pack mates. That may be true; but it's a different question whether it is bad *for the wolf* to be a bad wolf. Just as a bad person might seem to flourish, perhaps a bad wolf can flourish too.[29]

One concern about the present approach is that it seems to yield a passive view of well-being—a kind of desire fulfillment—that fits poorly with our stereotypes of flourishing, which so often seem to center on *activity*. Aristotelians and other eudaimonists have often claimed that well-being is more a matter of what you do than what happens to you or what state you're in, and such claims have significant appeal. Consider that eudaimonistic flourishing is often characterized as "well-functioning." I think, however, that we can accommodate the appearance by noting that animal drives very often are *for* certain sorts of activity. The cat doesn't just want food; it wants to *hunt*. (Indeed I'm told that most of the fruits of my cat's hunting won't be eaten by him—he's just in it for the sport. This has me thinking that the next cat will stay indoors.) Similarly, the wolf isn't

[28] This distinction, with slightly different terminology, frames Gewirth's helpful discussion of nature-fulfillment theories (Gewirth, 1998).

[29] There remains a hard question about whether individuals who "miss out" on characteristic elements in the lives of their species are thereby worse off. A psychopath who never knows love can seem to have a sad life, missing out on a major element of human life; whereas a tribal warrior who enjoys good relationships but is ruthless to outsiders does not seem, to some of us at least, to be missing out on anything, or in any worse off for it. Such worries, regarding *deprivation*, strike me as the chief motivation for being an Aristotelian, and more broadly for rejecting what I've called "internalism" or subject-dependence for an "externalist" or subject-transcendent approach to well-being. I think we can resist the pull of subject-transcendence, but will not argue it here.

driven simply to obtain a certain result; its agency is structured around *social functioning*—playing with other wolves, coordinating the hunt, raising the young, and generally doing whatever wolves do together.

In sum, for (most of) the other animals, nature-fulfillment has two components: physical flourishing (health and vitality), which they share with the plants, and *agential flourishing*.[30] Turning now to the human case—and perhaps, at least to some degree, some of the other animals—we introduce person-level functioning: we are persons, with selves. Accordingly, our agency is structured differently: our activity, at least in healthy cases, is self-governed, in a literal sense: the self is in charge, not our animal drives. And our flourishing centrally or perhaps wholly involves the fulfillment of our natures as persons or selves: self-fulfillment. If we insert here the dual-aspect model of the self, we get the result that human flourishing centrally involves, on the one hand, authentic value-fulfillment (the fulfillment of our natures qua agents), and authentic happiness (the fulfillment of our natures qua evaluators). Substantively, this picture departs from that for other animals in two ways: in place of drives we have values; and, further, we have complex emotional mechanisms for evaluating our lives and inducing corrections where needed. Most other animals seem only to have simpler and more narrowly targeted affective mechanisms, like the fear response, or physical pleasure and pain, that do not add up to any sort of generalized evaluative or governance system. It is not clear that the notion of an "emotional condition" really applies to lizards, or to rabbits or birds, so that we can sensibly speak of them as being "happy" or "unhappy." As best I can tell, lizards don't get depressed. With dogs and other social hunting mammals, which seem to have much richer emotional lives, the picture is fuzzier; and with other primates, it is not obvious that we can't sensibly speak of them as persons, with selves, who can be happy or unhappy.

A natural proposal is that nature-fulfillment in the human case has two basic components: not just self-fulfillment, but physical flourishing, in light of the nature we share with all living things. However, there are reasons to think that physical flourishing drops out of the picture in the human case: when dealing with persons, our welfarist concerns center on how the *person* is doing, and not the organism or creature.[31] When a friend ails, one wants to know what to do for *her* sake, and it isn't clear that sympathetic or benevolent concern for her involves any fundamental concern for her body, health perhaps mattering only as it bears on the concerns of her self.[32] When severe brain damage leaves someone in a permanent vegetative state, it is not clear that anything can benefit or harm him: he's gone.[33] Plausibly, there is no longer any reason of welfare to keep him alive, if indeed we can even say he lives.[34] Note also that it is commonly supposed that one might survive death and go on to heaven in a completely disembodied form—a condition, notice, that is generally deemed a very good thing for the person. Finally, it is significant that physical health or flourishing typically does not appear among the list items in objective list theories, suggesting that

---

[30] I borrow the latter term from (Raibley, 2010; 2012; 2013).

[31] [RRR Darwall's rational care theory of welfare.]

[32] As I note below, pleasure seems not to matter solely for reasons of nature-fulfillment, as its value does not seem dependent on our conception of the self but rather appears to be brutely phenomenological. But while pleasure's value may not be explicable in terms of self-fulfillment, its value for a person may still be closely if not essentially tied to its bearing on the self: it is the experiencing self that benefits from pleasure.

[33] [rrr Kraut 2018 makes this point]

[34] Perhaps other reasons for sustaining the body's life exist, such as reasons of dignity.

there is not a good deal of intuitive support for counting physical health or flourishing as a basic element of well-being.[35]

Perhaps physical flourishing drops out of the scheme even for the other animals, at least the sentient ones. As with the human case, there seems no benefit to the animal in keeping one's dog or cat alive if all consciousness has been irretrievably lost. Indeed it would seem bizarre to do so without special cause, whereas one might well have reason to tend to the needs of a tree, just for its sake.

These are not decisive arguments. Physical functioning could have fundamental prudential significance, for instance, even if its value is conditional on the animal's being conscious. But even if nature-fulfillment in human beings does include physical flourishing, its contribution seems to be minimal—at best, a small enough part of the picture that little of practical significance will be lost if we exclude it from our account of well-being. Of course, even this position might be challenged, for instance on the grounds that the self is essentially embodied, so that self-fulfillment includes the fulfillment of our bodily natures. If the objector finds plausible a self-fulfillment theory that departs from mine on account of its different understanding of the self, then at least that point of disagreement signals a deeper and more important convergence on the basic structure of well-being. Given the extensive dissensus about the nature of the self, one should expect divergence about the character of self-fulfillment. It is more important for current purposes just to get the self-fulfillment framework on the table.

3. Pleasure

Just about everyone who has articulated a nature-fulfillment ideal of human well-being has stopped there: that's all there is to well-being. And why not? One of the beauties of a eudaimonistic approach is that it can sustain a complex view of well-being, with multiple components, within a unified, elegant theoretical framework. We don't need to resort to the Frankenstein's monster of a "brute list" theory, nor must we try to wedge everything into a monistic theory, like hedonism or the desire theory.

Be that as it may, a nature-fulfillment theory can't encompass everything, and it fails to offer a remotely plausible story about the one good that a theory of well-being has to get right, if it is going to get anything right at all: pleasure, including suffering and other forms of unpleasure. That is, the account given thus far runs afoul of what I called the phenomenological intuition, which I will again frame mainly in its most compelling form, suffering: suffering matters for well-being, and it matters chiefly because of what it is like to suffer. The disvalue of suffering appears to lie in the phenomenal quality of suffering. Conversely for the value of pleasure. And the badness of suffering does not, by all appearances, lie in its being an impediment to functioning, or its being a sign or component of vice, or its in any other way being an instance of frustrated nature-fulfillment. In fact a eudaimonistic explanation of suffering's badness, at least of any conventional form, just seems to me bizarre, to be tolerated, if at all, then only for want of a better alternative.[36] The value strongly appears to reside in the quality of the experience itself.

---

[35] See Hurka and Nussbaum [rrr] for possible exceptions. In the latter case, bodily health is counted among the central human capabilities, and hence plausibly as among the functionings that constitute well-being. However, its inclusion may reflect the demands of her theory of social justice, where health-related freedoms be might have greater significance than health does in constituting well-being.

[36] In fact I think this point alone is fatal to the Aristotelian enterprise, and have pressed hard on it in earlier work (Haybron, 2007; 2008b). Probably the best hope for a pure nature-fulfillment theory is to claim that pleasure is a kind of subjective or experiential nature-fulfillment—for instance, the fulfillment of the "phenomenal self"

It is only slightly less implausible to claim, as most philosophers seem to, that the disvalue of suffering lies in the fact that it is unwanted: what makes it bad is that we desire to be rid of it, or alternatively that suffering itself involves aversive desires. (A closely related claim is also very popular, namely pleasure and suffering can't be defined purely in phenomenological terms, because the experiences are so diverse that they seem to share nothing in common; we seem unable to analyze the experiences involved in terms of some shared components. This sort of argument is underwhelming: those impressed by the phenomenological intuition aren't likely to suppose that pleasure and suffering could be analyzed in other terms, and have a ready answer to the question what unifies all the different experiences: *they're all pleasant*, or unpleasant, as the case may be.[37] That conceptual analysis can't decompose them into other, more basic, concepts is sort of the point.)

No doubt there is, for obvious reasons, a strong connection between suffering and desire. But it certainly does not *seem* as if suffering, or the bad experience involved in suffering, is bad only because we want not to have it, or otherwise it involves something contrary to our desires. It does not seem as if the experience in question is itself neutral, so that it should be a matter of indifference, well-being-wise, whether we are constructed to desire suffering instead of pleasure. Rather it seems that God would have played a sick joke on us to make us like that.[38]

Note also that whatever desires might necessarily be implicated in suffering, those must be very primitive, low-level conative states bearing only a distant relationship to those values that, we saw earlier, are a much more plausible candidate for conferring normative force. If suffering and pleasure essentially involve desire,[39] the desire in question is more like that of a smoker's craving than of a valuing, and so the suffering of a waterboarding is bad in the way that a frustrating someone's urge to smoke a cigarette is bad: both involve the frustration of desire in the form of primitive inclinations. Even those inclined to cash out well-being in terms of desire should be hesitant to run this line: one probably does not want to give low-level desires like the smoker's cravings, which if anything seem to be *less* primitive than whatever conative states might be constitutive of unpleasantness and suffering, such a high degree of normative force. In fact this result seems to me a reductio of that sort of desire theory, among other things abandoning any notion that it embodies some ideal of agent sovereignty, making individuals authoritative about their well-being. A more plausible sort of desire-based approach would say: suffering is bad because it is contrary to your values. But then, if you valued suffering, it would be good for you. This result seems to me a reductio of a pure value-fulfillment theory of well-being.

The disvalue of suffering, and the value of pleasure, seem to be sui generis, unlike anything else. This seeming seems to me pretty bedrock, and pretty plain and unavoidable, and hence to pose a monumental problem for nearly all theories of well-being. To my mind, failing to give a credible account of hedonic value is, for a theory of well-being, rather like a physics that can't account for material objects. Short of incoherence, it's about the worst failure a theory could have. Only hedonists, who take that to be the whole story, and brute list theorists, who can say pretty much anything they want to, seem able to accommodate this appearance. But hedonism does untold violence to many other appearances, experience machines being just the tip of a rather large

{Klausen:2015ki}. This may not be bizarre, and perhaps it can be made to work, but it still seems false to the phenomena.

[37] E.g., (Crisp, 2006a; 2006b). [RRR add refs in this section to recent lit on pleasure]

[38] [RRR]

[39] Where the good/bad-making property is the desire, and not what it feels like to have the desire, which just brings us back to the view I'm offering.

iceberg. So it would be desirable to find some other way to deal with pleasure. All theories in ethics have problems, to be sure, but this one is a doozy.

I see only two avenues of escape: a list theory, or the fragmentation of well-being—that is, positing multiple concepts corresponding to the traditional philosophical category of well-being, and assigning hedonic value to one of them. In either case, the effect on the present account is to yield a 2-part theory of "well-being": nature-fulfillment, specifically self-fulfillment, and pleasure. The question is whether these are different sides of the same value—well-being—or two distinct values, each meriting a different theory. Which avenue to take is not entirely obvious. The fragmentation option has already been raised a number of times in the literature regarding various issues, and has considerable attractions (though, as Shelly Kagan notes, it also creates hard puzzles about why those values should be posited and not others, and how the different values are related[40]).

I am inclined, however, to go the list route, despite its having a glaring problem: list theories don't really look like theories at all, just ad hoc assemblages concocted to deal with various issues as they come. One reason to go this route is simply that it seems wise to avoid proliferating values more than necessary. Another is that nature-fulfillment and pleasure seem genuinely to belong within a single domain of value, in which we regularly make tradeoffs. Take, again, the case of antidepressant use. Suppose you are depressed but don't regard it as a disorder: things are just bad in your life, and you are miserable. Should you take an antidepressant? Many people in this scenario will be ambivalent, even setting aside concerns about dependency or learning to cope with adversity: on the one hand, it will make life more pleasant, and hence benefit you in that way. On the other hand, you might feel that it makes your emotional responses to your life less authentic, and see that as diminishing its benefits. Even if you end up happy, you might see that happiness as less than fully authentic, and hence as less valuable than an unmedicated happiness. There seems to be a tradeoff between two aspects of a single value, well-being: hedonically, it's good for you; eudaimonically, not so much.

The hedonic element also promises to address an oft-noted problem for nature-fulfillment theories generally: accounting for ill-being. It is not clear whether to regard unhappiness and the frustration of one's values as positive evils, but generally the nature-fulfillment framework invites a reading of welfarist bads as mere privations: a lack of nature-fulfillment—languishing, for instance. Pain and suffering, by contrast, strongly appear to be positive evils: something downright bad is happening, and not just a very regrettable absence of good. And there is accordingly a special moral urgency to suffering: people who have languished for years in a refugee camp do arouse our sympathies, partly over wasted lives. But the intense suffering of a struggling COVID patient or bombing victim can prompt an entirely different order of distress, even trauma; one fervently wants something to be done to stop this horror—and it is a horror—*now*, and to comfort the victim. There is no easy way to make good on such appearances in terms of nature-fulfillment. (One suspects that approach even gets the time scale wrong: it takes a while to establish whether someone is flourishing or languishing, and perhaps these values aren't even defined at shorter time scales. The onset or cessation of suffering, and our discernment of it, can be essentially instantaneous.)

The resort to a list theory may be less worrisome if, as I suggested in the introduction, we can at least find a rationale for the list at the metaethical level. At any rate, it does seem plausible that our concerns regarding well-being cluster in these two broad areas, eudaimonic and hedonic. There may be good reason why empirical research on well-being has been dominated almost entirely by these two approaches, and increasingly combines both given the broad appeal of both

[40] (KAGAN, 1992; 1994).

sorts of ideal. The stability of this state of affairs in recent decades, and the dim prospects for either approach giving way entirely to the other, suggests that some sort of eudaimonic-hedonic hybrid, Millian or otherwise, may be the only way for philosophical theories to give our diverse concerns about well-being their due. And going this route does not seem to invite an unwholesome proliferation of list items, each list according to the intuitions of the intuiter, with no clear end in sight.

4. Summing up: the Millian hybrid view

So we have, at least in outline, a complete theory of well-being on offer, a list theory with two basic components. I will not say much here about the "Millian" credentials of the view, but note briefly that it shares with his writings a dual focus on both pleasure and self-fulfillment. In another paper I argued that the self-fulfillment aspect of this view is a species of "Millian eudaimonism," contrasted with "Aristotelian eudaimonism" chiefly in that it centers on the fulfillment of one's nature *qua* individual, rather than on species norms.[41] That is, it is "internalist," grounding well-being entirely in the arbitrarily idiosyncratic particular's of the individual's makeup, whereas Aristotelian and many other objectivist views are "externalist," allowing that what's good for us may not depend entirely on the individual's particulars. However, perhaps it would be clearer to refer to internalism, of which subjectivism is a prominent variety, as "subject-dependence" about well-being, while externalist views are "subject-transcendence."[42] I will leave the terminology open here.

<u>The Millian hybrid theory</u>:

1. Self-fulfillment (eudaimonic well-being: flourishing)
    a. Authentic happiness (psychic flourishing)
    b. Authentic value-fulfillment (agential flourishing)
2. Pleasure (hedonic well-being)

The present view of well-being seems to me to fit well with the appearances about well-being, though of course we should expect some clashes here and there with intuition, some of which I've already noted. One is that it may not adequately address widespread concerns about *deprivation*—intuitively, "missing out" on what are seen as important elements of a full human life—since we might imagine people being happy and fulfilling their values in lives that do not include goods like sight or hearing, or certain goods typically enjoyed in childhood, adulthood, or old age. Even if the circumstances do not preclude flourishing—and I think most disabilities for instance do not—it is not clear that any subject-dependent view can make sense of the thought there is something to regret, a deprivation, in such cases. While I share these concerns to some extent, there is certainly no consensus about whether theories of well-being should cater to such intuitions, and I suspect they can be better addressed otherwise, for instance via moral norms of inclusion and equity.

In this sketch I left out some details of relevance, for instance what is meant by authenticity; but in earlier work I incorporated a *richness* constraint, such that we live more authentically, more fully expressing who we are, in richer, more complex ways of living. This goes some way toward diminishing worries about impoverished lives, but it is not obvious that significant problem cases won't remain. My treatment of *value-fulfillment* has been cursory, largely deferring to intuitive

[41] {Haybron:wn}.
[42] I borrow these terms from [Tiberius and Hall 2015].

understandings of the notion and related discussions in the literature. But much remains to be said, not just about what exactly one's values are, but about the scope of the values that count; what if you have values that have no bearing on how your life goes? One possibility is that value-fulfillment should be understood as a kind of *success*, hinging on the values that structure the way you lead your life, and not on values concerning only remote states of affairs in which you take no particular interest.[43] Such questions I set aside for later.

On balance, however, this strikes me as an intuitively attractive position: it seems to get the right answers about what's good for us. Just as importantly, it seems to do so for the right reasons.[44] And there is an advantage to leaving the view somewhat open to further development: if the basic idea has legs, then we should expect different readers to want to fill it in differently, with different understandings of the self, of authenticity, of whether to go the hybrid route or adopt a pure self-fulfillment account, or of whether to incorporate the present account into a nature-fulfillment framework like Aristotle's. If there is ever a time to seek the last word in normative ethical theorizing—and I doubt there often is—this is not it.

5. Millian hybrid and Aristotelian approaches compared

5.1. Hybrids, Millian and Aristotelian

I have argued for a "Millian hybrid" theory of well-being, essentially a two-item list theory with pleasure as one component and nature-fulfillment as the other. Nature-fulfillment in turn centers on self-fulfillment: authentic happiness and authentic value-fulfillment. But, as embodied beings, our well-being consists partly also in physical flourishing, though this has only secondary importance. The view is somewhat complex, and this might seem a problem: when it comes to well-being, simple and compelling ideals seem more likely to be true (or at least to be remembered). But it is significant that we were able to arrive at it through a number of converging lines of argument.[45] As noted in the introduction, I suspect that evidence from moral psychology, particularly regarding the structure of human valuing, would lend further support.

As well, the view is not that complex: at bottom it only has two elements and might crudely be summarized, pace Badhwar, as identifying well-being with *authentic happiness in a subjectively worthwhile and pleasant life.*[46] Moreover, the basic structure of the view embodies two major threads of thinking about human welfare, which recur in a variety of cultures, and indeed are taken to be the primary options in psychological research on well-being: the eudaimonic and

---

[43] [RRR Keller, Raz, Scanlon]

[44] Very often, philosophical theorizing seems focused mainly on extensional adequacy: if your theory gets the extension of a concept right—gets the "right answers"—that's good enough. Accordingly, a good deal of verbiage is spent playing some variant of "counterexample whack-a-mole," showing how one way or another, the theory can be construed to avoid giving implausible-seeming answers in various cases. Yet it seems just as important that the reasoning by which those answers are gotten be plausible. Indeed, there are indefinitely many crazy theories that can generate whatever answers you want; isomorphisms are easy to come by, if you don't really care what's behind them. So yes, it's good if your theory implies that happiness is good for us; but not so good if it does so only because, say, happiness makes you more successful at work, or for still weirder reasons, such as that three is a prime number. Showing that your theory can get the right answers, then, does not even come close to settling whether the view should actually be taken seriously.

[45] Some of which I did not even mention here, having discussed them in earlier work RRR.

[46] At least, if value-fulfillment can properly be identified with a subjectively worthwhile life. Since I take happiness and pleasure to be objective goods, the view overall is not aptly deemed subjectivist, though "quasi-subjectivist" would not be an unfair description; it is not strongly objectivist in the manner of Aristotelian and list theories.

hedonic approaches. The persistent and widely perceived division between these schools of thought even among non-philosophers raises the question whether each approach gets a major part of the story right, and might even be complementary. So compelling are these two elements that Mill, with his admirable commitment to getting things right rather than keeping them simple, seemed unable to sustain the pure hedonism for which he is best known. His views also took a eudaimonistic turn, marrying hedonistic ideals with ideals of self-fulfillment. Exactly what form that marriage took has never been entirely clear, but perhaps that unclarity owes to the difficulty, if not the impossibility, of genuinely combining the two strains of thought in a unified framework. He might have met with more success, perhaps, had he grabbed the bull by the horns and gone the hybrid route.

Interestingly, recent scholarship suggests that Aristotle might have confronted a similar tension in his views. Whereas Richard Kraut's past work embraced a more or less orthodox Aristotelian "developmentalist" account of flourishing grounded in the early books of the *Nicomachean Ethics*, he recently undertook a radical shift to an "experientialist" (or experiential developmentalist) theory of well-being that departs from hedonism only in counting a much wider range of experiences than pleasure as basic elements of well-being.[47] Strikingly, Kraut's experientialist position also is rooted in Aristotle's work, this time especially the later account of *eudaimonia* as the life of contemplation from Book X, which notoriously looks quite different from the picture of well-being charted earlier in this work. And it is not clear, in fact, that the good of the contemplative life need involve anything other than certain sorts of experiences. Whatever the basis of the tensions in Aristotle's views of well-being, it is clear that Kraut is impressed by similar considerations to those raised above in favor of a hedonic element to well-being (and I am not sure Kraut isn't right to expand the relevant goods beyond the hedonic to experiential generally).

There simply is no remotely credible way to explain the significance of pleasure and suffering within a conventional Aristotelian framework centered on the life of virtuous activity, a shortcoming I have argued to be fatal to the approach. Indeed, it is plain that Aristotle's discussion of pleasure's role in eudaimonia is not even about pleasure at all, but rather happiness: virtue does not depend on whether one feels *sensory* pleasure, like experiencing a pleasant odor or sexual tingle, when doing the right thing. Rather, virtue requires having the right emotional response—gaining happiness from doing so, rather than feeling distress. About the pleasantness of experience *per se*, the Aristotelian framework may have little or nothing to say; it isn't even the icing the cake, it seems, but may not matter at all. This makes the standing of orthodox Aristotelian views even more precarious, and suggests that pleasure (or perhaps experiential goods generally) calls for substantial revision of the theory. But there is an alternative to going full experientialist: an Aristotelian hybrid account marrying eudaimonic with hedonic (or perhaps experiential) well-being, differing from the Millian hybrid only in how the ideal of nature-fulfillment is spelled out. It seems to me that Aristotelians can accommodate hedonic goods just as well as self-fulfillment theorists—which is to say, by going the hybrid route.

5.2. The distinctive importance of happiness and pleasure

It may seem odd if not redundant for the theory to center on two closely related goods, happiness and pleasure. But they in fact matter for quite different reasons, and it may say more about the primitive state of this part of our culture than about the present theory that so many have failed to distinguish them. Unemployed people tend to be less happy than the employed, for

[47] [rrr]

example, yet hedonically it is possible—and may often be the case—for them to fare no worse, simply because their time is allocated differently.[48] While emotionally unfulfilled and getting less pleasure from the same activities, for instance, they might spend less time than the employed in unpleasant situations like commuting and meetings with the boss, allotting more of their time toward unrewarding but more pleasant pastimes like television and video games. They've still had the stuffing taken out of them, and this spiritual deflation is a central element of their misfortune—and arguably the more important fact—even if their experience isn't less pleasant.

Usually, of course, the stories of happiness and pleasure more or less converge. But that does not mean their contribution to well-being is identical. (Indeed, one of the morals of empirical research on well-being is just about everything you might care about tends to correlate with everything else: the good things in life tend to travel together, as do the bad things. But there remain many different varieties each of good and bad things.) Depression, for instance, is primarily bad in two very distinct ways: it is of course extremely unpleasant, and in this respect it resembles the bad of chronic pain. But the evil of depression goes well beyond this, as a *languishing* of the self, and languishing is hardly reducible to unpleasantness of experience. Some people manage to maintain decent spirits, to be fully themselves, despite severe pain (which is not to deny that some pains might be beyond any human's powers to endure). Unpleasantness is one dart, as the Buddhists say; but unhappiness adds to this a second dart: not just further unpleasantness, but a kind of smothering if not extinction of the self. When depressed, one is like an parched shrub, withered and shriveled, and sharply diminished in functioning. A shadow of oneself, we sometimes say. One's spirit may be crushed or broken, so that a kind of apathy, even anhedonia or avolition, may ensue. But this cannot be the liberation Buddhists or Stoics are after. It is in fact one of the worst things that can happen to a person. Similarly, to break a person's spirit is one of the greatest evils human beings can perpetrate, far worse even than the related harm of breaking a person's heart, which may at least leave their personality more or less intact.

It is an important question what exactly these evils amount to, but here I simply want to note that they are quite obviously not merely hedonic evils. Indeed, one way to break a person's spirit is to get them to accept the yoke, even not to mind it at all, so that hedonically little or nothing is lost; perhaps the contrary. (See, e.g., *Brave New World* and *1984*.[49]) I suspect there are many things wrong in such cases, but the diminished authenticity of one's emotional responses, and of one's valuing and willing, appears to be a large part of the story. In slightly plainer terms, what makes a broken spirit so bad is, in great part, the loss of self, or of its expression. Authentic happiness does not seem possible in this condition, and one's valuing may be impaired as well. Note that even such happiness as a broken person might enjoy is bound to be, barring extreme cases that may reside only in science fiction, of a very wan, shallow sort. Even if the individual is not quite unhappy, psychic flourishing is out of the question. This thought is not available to the hedonist or, for that matter, any theorist who lacks some robust notion of self-fulfillment. The pleasantness of our experience certainly matters, but so too does the full, positive expression of the self.

5.3. Individuality and broken spirits in Aristotelian theory

What of the Aristotelian? It is possible to say all of this within a broadly Aristotelian framework—at least of a hybrid sort—if only because the present conception of self-fulfillment can be

---

[48] [Knabe RRR. It is possible that limitations of the measures skewed results to underestimate the hedonic penalty of unemployment, but the basic phenomenon of "hedonic compensation" is clearly possible, and plainly sometimes actual.]

[49] [RRR Stump et al.; explain]

embedded within an Aristotelian theory: one might say, with some plausibility, that a central aspect of a characteristically and fully human life rests on our status as persons, with selves, who thus flourish partly through authentic happiness and value-fulfillment. The Aristotelian would add to this that our species nature contributes various other elements to the picture, such as virtuous activity and otherwise functioning well *qua* human being. But while such an interpretation is possible and perhaps advisable, it is hardly orthodox.

Barring this sort of move, the Aristotelian needs somehow to explain, in an intuitive manner, how someone with a broken spirit fails to flourish. I am not sure how exactly this story would go. Aristotle himself might have observed that the *megalopsuchoi*—the great-souled men—could hardly have broken spirits, and one it may not be hard to identify various ways in which the expression of the virtues would be impaired in such a case (wit, for one, seems improbable). But all this would miss the point if there were not also some reference to the importance of expressing your unique characteristics as a person, and that seems not to be a characteristically Aristotelian concern.

The main piece of parenting advice I got from my parents was: "listen to the child." Pay attention to the distinctive needs of each, and don't impose your agenda on them. This seems to me sound advice. It also seems to be putting the cart before the horse to say, *it's because of certain facts about the species that a parent should listen to the child*. There's at least some pull to the notion that the goods of individuality should not be wholly subservient to species norms.

Arguably, Aristotle's framework is less at home grounding these sorts of homespun truths than in providing underpinnings for the sorts of social institutions in which not listening to the child, but something closer to breaking them, is sort of the point. Perhaps people are generally prone to stray unless firmly habituated to realize the goods appropriate to the human species, which local institutions in various times and locales might reasonably deem to require conformity to one or another mold that is, to varying degrees according to the individual, Procrustean. For whatever reason, it is fairly standard in a great many societies for social institutions more or less purposely to engage in various forms of pounding down the nails that stick out, rather as show dogs are bred to weed out idiosyncrasies that make a specimen less representative of the breed's essence. The infamous model of British schooling comes to mind, and doubtless helped inspire Mill's musings on individuality, but there are certainly worse offenders. We don't ordinarily speak of such practices exactly as breaking people's spirits, as they are usually not so extreme as that term suggests, and a truly broken person is unlikely to make a productive corporate widget. But they are plainly somewhere on the spectrum, and raise the same basic worries. It is perfectly ordinary even in liberal democracies to worry that our regimes sometimes function to break people to the lash. This is the stuff of children's stories like *Ferdinand the Bull*, whose protagonist would rather smell the flowers than fight like a normal bull, and which would never have gotten off the ground were there not some discernible tendency of society to penalize similarly atypical humans.

Such penalties make very little sense if your goal is to foster self-fulfillment, but very much sense if your goal is for people to instantiate some characteristic norms of the species, or whatever collective one fancies. Individuality may well be a key norm in the human case, but evidently this has not been apparent to everyone, and ideas about what human beings are characteristically or even ideally like notoriously vary a good deal, even within the Aristotelian camp. Even for us today it is not crazy to wonder if human beings are really suited to democratic self-governance. (Many have thought the answer is obviously no.) Or perhaps we are basically sheep-like herd animals, most of us anyway, who need to be kept in line. Or maybe we are fundamentally vile, or at least violently predisposed toward outsiders. Perhaps heterosexuality is the characteristically

human form of sexuality, so that homosexuality is neither virtuous nor compatible with flourishing. It is even feasible, within a fairly orthodox Aristotelian perspective, to defend slavery and the subjugation of women, and in fact there is some precedent for this interpretation.

Again, it is perfectly possible for Aristotelians to avoid such conclusions, and contemporary philosophers in this tradition seem to do so about as reliably as non-Aristotelians. Even looking well back into the historical record, I find it hard to fathom that Aquinas, or for that matter Aristotle, would have smiled on Torquemada's enthusiasm for spirit-breaking in the name of an Aristotelian church.

5.4. Theoretical virtues and interpretation

But a theory should be judged partly by the interpretive disputes it invites. Normative ethical theories, and perhaps philosophical theories generally, tend to entail a range of subjunctives—that is, alternative interpretations—depending on how key terms and claims are understood. Often if not usually there is more than one reasonable interpretation of the theory's core commitments, and it is desirable if the reasonable interpretations also yield similarly reasonable conclusions.[50] To the extent that they do, this suggests that whatever plausibility the theory has is *robust*, and not fragile, surviving plausible variation among the theory's core themes and not depending wholly on specific readings that might reasonably be disputed. This robustness is enhanced to the extent that the theory's natural inertia tends toward more rather than less reasonable results—that is, if it is not only possible to read the theory in a way that seems to get the right answers, but also natural to do so. This is especially important if, as it seems to me, a normative ethical theory can rightly be evaluated to some extent as a *heuristic* for thinking about how to live, a theory that tends to improve practice being preferable to one that tends to worsen it, given how people are likely to employ it. (How much weight to give this role is an open question, but presumably we should continue to wish for our theories to have the appearance of truth, and not just usefulness.)

To illustrate: the Kantian ideal of respect for persons yields very different interpretations, for instance egalitarian or libertarian, according to the interpreter's sensibilities. This is a strength of the approach, as it suggests that the core ideal captures a deep insight common among reasonable viewpoints, even where those have very different practical implications; and whatever the weaknesses of Kantian ethics, being radically counterintuitive tends not to be among them. (The example also illustrates that more robust is not always better than less, as it can come at the expense of substantive content: a theory that seems immediately plausible to everyone might be too trivial to be interesting or useful. While the enormous flexibility of the Kantian framework hardly renders it trivial, it notoriously raises questions about whether the approach offers enough guidance.)

Relatedly, one way of putting the usual worries about consequentialist moralities is that, even if there exist plausible interpretations of such views that don't license judicial killings of the innocent and other moral grotesqueries, there's also no shortage of reasonable interpretations that do. And certain things shouldn't really be up for debate; there's something wrong with one's moral theory, for instance, if it puts genocide or slavery on the table, even if the very best interpretation of it thankfully spares us that result. Fortunately, in the world as we know it there doesn't seem to be a reasonable reading of consequentialism that would green-light genocide. But that leaves no shortage of ostensible counterexamples on matters that shouldn't even be up for discussion. Anscombe's infamous complaint about corrupt minds may have been objectionably uncharitable,

[50] I leave the interpretation of 'reasonable' open to readers to understand as seems plausible to them, but it should track one's intuitive sense of the span of reasonable disagreement.

but she was right to be concerned about a theory that promises to make the unthinkable abundantly thinkable. (The moral record of economics, where consequentialism has effectively enjoyed official standing, does little to alleviate such worries; see, among others, the Larry Summers memo.)

Returning to theories of well-being: a self-fulfillment view invites disagreement about the nature of the self—perhaps a narrative, relational or embodied view is preferred—what counts as authentic, what the right tradeoffs are between various aspects of self-fulfillment and other ideals, and so forth. It is not yet clear how such debates will play out, and of course there will be objections to the approach as to any theory, but at this juncture it is not clear that we should expect particularly worrisome views to arise as reasonable options. There is an ugly history of "true self" ideology deployed to rationalize brutality, as Berlin observed,[51] but it is hard to discern any serious basis for interpreting a self-fulfillment theory of well-being in such a manner, which has an Orwellian "we had to destroy the village in order to save it" flavor to it. A fairer concern is that talk of "self-fulfillment" conjures images of self-indulgence, New Age hokum, and individualism run amok, which suggests it may be better where possible to employ other terms like 'nature-fulfillment'. More to the point: while I hope to have made clear that ideals of self-fulfillment could be at home as much in collectivist cultures as anywhere else, the framework may well tend to skew thinking toward undesirable forms of individualism, like those characteristic of consumer culture. This is a legitimate concern, though not I think terribly worrisome, and it is arguably shared by any subject-dependent approach, including subjectivism, which is to say most current views of well-being. Recognizing other values beyond well-being, such as virtue, excellence or beauty, seems to me solution enough.

The Aristotelian framework, by contrast, invites disputes about whether it's possible for gay people to flourish or be virtuous, whether it is really good for some of us to be governed by our superiors, or pressured into conformity like the sheep we might in fact be,[52] whether we ought to be provincial and hostile to outsiders, etc. All of these seem to be implications of one or another reasonable view of human nature. (And for my money the happy Comanche warrior who leaves captives to the vultures buried in the desert up to their necks, *sans* eyelids, may well be a paradigm of distinctively human flourishing, perhaps more so than many a chair-bound scholar whose moral ledger bears a good deal less red ink. It certainly is not obvious that fulfilling human nature demands the acquisition of cosmopolitan virtues. Indeed, on this count Aristotelians seem clearly to be paddling upstream.)

It is a problem if reasonable interpretations of a theory yield unreasonable results, and one does not escape it simply by pointing out that some reasonable interpretations do not. Compare: "The superior race should rule. But all races are equally good, so all people are equal." This is not a reassuring line of thought, even if all ends well, in some fashion. When a theory invites serious debate about matters that do not merit debate, this suggests that the underlying ideal may not be the right one, since it steers some of us in quite the wrong direction. It suggests also that the theory may not serve well as a heuristic, which is arguably one role for normative ethical theories: we should like such theories to help us improve our practice, but in practice such an account may worsen people's ethical outlooks (see, again, economics). However different it may be in content, in style the Aristotelian framework shares with consequentialism a propensity to breed monsters, in theory if not in practice. The Inquisitor could perhaps be forgiven if he reckoned heresy, in light of the ostensibly reasonable doctrines of his day, to be a greater threat to the fulfillment of human nature than a shortage of individuality.

---

[51] [RRR]

[52] [RRR Frith, herding in humans; Haidt, hive psychology]

I do not deny the intuitive pull of the Aristotelian framework, but the question arises how much its appeal owes to atavistic impulses—tribe- and kind-obsessed perspectives on life—best put behind us. It is surprising, in this regard, how quickly we pass over familiar worries about the sizeable role allotted to an arbitrary concept of species in Aristotelian thinking, a conceit that would not likely be available to us were humans no easier to distinguish from the other apes than *Empidonax* flycatcher species are from each other. One suspects bird philosophers would not make heavy weather over the essence of a distinctively Pacific-slope flycatcher life, seeing as that one can barely distinguish them from other flycatchers at all. And one wonders how much purchase the Aristotelian view would have on our imaginations these days if our poor hominin brethren hadn't all done us the favor of going extinct.

One does not need to adopt a self-fulfillment theory of well-being to make sense of the role of happiness in flourishing; an Aristotelian view, for instance, could do so by incorporating self-fulfillment in a broader theory of well-being as nature-fulfillment, and arguably some have implicitly done so.[53] But a species-based approach like this does not obviously lend itself to a plausible treatment of concerns regarding individuality—it can, for instance, make sense of the evils of having one's spirit broken, but only insofar as it approximates a self-fulfillment theory. And there is no natural inertia to the framework to yield that sort of result; if anything, inertia points in the opposite direction, toward conformity to group norms.

At any rate, it should be clear that the happiness element of the present theory is not redundant with the pleasure element, and that its full significance can only be understood with some reference to self-fulfillment.

5.5. Final remarks: well-being and objective worthwhileness

In short, it is plausible that we flourish through being authentically happy in lives that fulfill our values: in lives that fully realize and express who we are. And it is plausible that we have, in addition, a fundamental interest in the quality of our experience, just for what it is like. While these points could be incorporated as part of an Aristotelian theory of well-being, they appear to be in tension with the fundamental thrust of that approach.

A clearer point of disagreement with the Aristotelian approach has to do with the role of objective goodness in well-being. Aristotelians such as Badhwar disagree with the present view on the necessity of objective worthwhileness, including virtue, in well-being. But we may well agree on what I take to be the more fundamental and important question, namely the character of the *good life*—the sort of life one should want, whether for one's own sake or for whatever reason. On this question, nearly all ethical theorists would agree that there are objective standards at least of morality, and that a life of immorality is not choiceworthy, and hence not a good life, whether the individual manages to flourish or not. To well-being, then, I would add that a good life must also be objectively worthwhile, including at least moral virtue and other forms of excellence. Our slogan for the good life, then, comes more closely to resemble Badhwar's formulation of well-being: happiness in an objectively worthwhile life. This is a different and more specific way of framing the conception of the good life that I have dubbed *a life well-lived, and well worth living*.[54] If all this is right, then the divide between the present view of well-being and Aristotle's may not be so great, once the totality of our ethical frameworks is taken into account. But it seems to me that in the domain of well-being, a Millian hybrid holds more promise than its Aristotelian cousins.

---

[53] [RRR Badhwar, Russell…]

[54] [rrr 2013]

6. References

Annas, J. (1993). The Morality of Happiness. New York: Oxford.

Arpaly, N. (2004). Unprincipled Virtue. Oxford University Press.

Badhwar, N. (2014). Well-Being: Happiness in a Worthwhile Life. New York: Oxford University Press.

Crisp, R. (2006a). Hedonism reconsidered. *Philosophy and Phenomenological Research*, *73*(3), 619–645.

Crisp, R. (2006b). Reasons and the Good. New York: Oxford.

Doris, J. (n.d.). Talking to Our Selves: Reflection, Skepticism, and Agency. New York: Oxford University Press.

Foot, P. (2001). Natural Goodness. New York: Oxford University Press.

Gewirth, A. (1998). Self-Fulfillment. Princeton: Princeton University Press.

Haybron, D. M. (2001). Happiness and Pleasure. *Philosophy and Phenomenological Research*, *62*(3), 501–528.

Haybron, D. M. (2005). On Being Happy or Unhappy. *Philosophy and Phenomenological Research*, *71*(2), 287–317.

Haybron, D. M. (2007). Well-Being and Virtue. *Journal of Ethics & Social Philosophy*, *II*(2).

Haybron, D. M. (2008a). Happiness, the Self, and Human Flourishing. *Utilitas*, *20*(1), 21–49.

Haybron, D. M. (2008b). The Pursuit of Unhappiness: The Elusive Psychology of Well-Being. New York: Oxford University Press.

Haybron, D. M. (2013). Happiness: A Very Short Introduction. New York: Oxford University Press.

Haybron, D. M. (2016a). MENTAL STATE APPROACHES TO WELL-BEING, 1–32.

Haybron, D. M. (2016b). The Philosophical Basis of Eudaimonic Psychology. In J. Vittersø (Ed.), *The Handbook of Eudaimonic Wellbeing* (pp. 1–28).

Haybron, D. M. (n.d.). Eudaimonism, Ancient and Modern.

Haybron, D. M., & Tiberius, V. (2015). Well-Being Policy: What Standard of Well-Being? *Journal of the American Philosophical Association*, *1*(04), 712–733. http://doi.org/10.1017/apa.2015.23

KAGAN, S. (1992). The Limits of Well-Being. *Social Philosophy and Policy*, *9*(2), 169–189.

KAGAN, S. (1994). Me and my life, *94*, 309–324.

Keller, S. (2009). Welfare as Success. *Noûs*, *43*(4), 656–683.

Klausen, S. H. (2015). Happiness, Dispositions and the Self. *Journal of Happiness Studies*, 1–19. http://doi.org/10.1007/s10902-015-9628-6

Kraut, R. (2007). What is Good and Why. Cambridge, Mass.: Harvard University Press.

Raibley, J. (2010). Well-being and the priority of values, *36*(4), 593–620.

Raibley, J. (2012). Health and well-being. *Philosophical Studies*, *165*(2), 469–489. http://doi.org/10.1007/s11098-012-9951-2

Raibley, J. R. (2013). Values, Agency, and Welfare. *Philosophical Topics*, *41*(1), 187–214. http://doi.org/10.5840/philtopics20134119

Railton, P. (1986). Facts and values, *14*(2), 5–31.

Scanlon, T. M. (1999). What We Owe to Each Other. Cambridge, Mass.: Harvard University Press.

Sripada, C. (2014). *At the Center of Agency, the Deep Self* (pp. 1–19).

Sumner, L. W. (1996). Welfare, Happiness, and Ethics. New York: Oxford.

Tiberius, V. (2008). The Reflective Life: Living Wisely With Our Limits. New York: Oxford.

White, M. P., & Dolan, P. (2009). Accounting for the Richness of Daily Activities. *Psychological Science*, *20*(8), 1000–1008. http://doi.org/10.1111/j.1467-9280.2009.02392.x

Williams, B. (1973). A Critique of Utilitarianism. In J. J. C. Smart & B. Williams (Eds.), *Utilitarianism: For and Against*. New York: Cambridge University Press.