Well-Being: What Matters Beyond the Mental?* Jennifer Hawkins

Most philosophers assume that more matters for well-being—intrinsically—than simply mental states.¹ Certain standard, and by now familiar, examples are thought to establish this. There is the deceived spouse whose happiness depends on belief in her partner,² or the man who is secretly despised by those he thinks of as friends,³ or the woman who dies believing falsely that she has succeeded in an important project.⁴ Probably the most famous example is provided by Nozick's experience machine.⁵ We are invited to imagine the most powerful virtual reality machine possible, one that could, simply by stimulating our brain, make us think we were having any sort of experience. Like the individuals already mentioned, a person living inside the experience machine might think that her life is going well, but the beliefs upon which this assessment would be made would all be false. In all of these examples, individuals are deluded about certain important aspects of their

^{*}

^{*} This article has been a long time in the making and I am grateful to many who offered helpful comments at the following conferences: 4th Annual Rocky Mountain Ethics Congress (RoME), University of Colorado, Boulder, August 2011; Joint Session of the Aristotelian Society and the Mind Association, Stirling, Scotland, July 2012; 4th Annual Arizona Workshop in Normative Ethics, Tucson, Arizona, January 2013; Central Division Meetings of the American Philosophical Association, New Orleans, Louisiana, February 2013. In particular, I wish to thank Stephen Campbell and Howard Nye for extremely useful commentaries, and Shelly Kagan for helpful discussion at the Normative Ethics Workshop.

¹ One recent example: In a brief discussion of Nozick's experience machine, Dan Haybron writes, "Apparently our mental states are not the only things that matter." Daniel M. Haybron, *The Pursuit of Unhappiness* (New York: Oxford University Press, 2008), 3. The focus of this article is intrinsic welfare value. If not specified otherwise, that is what is meant.

² Shelly Kagan, *Normative Ethics* (Boulder, Co.:Westview Press, 1998), 36; L.W. Sumner, *Welfare, Happiness, and Ethics* (Oxford: Oxford University Press, 1996), 157; Russ Shafer-Landau, *The Fundamentals of Ethics*, 2nd edition (New York: Oxford University Press, 2012), 34-5. Kagan builds all three of the standard types of example into one case of a man deceived about spouse, friends, and accomplishments.

³ Kagan, *Normative Ethics*, 36; T. M. Scanlon, *What We Owe To Each Other* (Cambridge, Ma.: Harvard University Press, 1998), 112; James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: Clarendon Press, 1986), 19.

⁴ Kagan, *Normative Ethics*, 36; Shafer-Landau, *Fundamentals*, 52; Griffin, *Well-Being*, 23. Griffin gives an inverse example of accomplishment unknown (as opposed to belief in accomplishment when there is none). He remarks that if Bertrand Russell's work for nuclear disarmament had actually significantly reduced the chance of nuclear war, then this would be good *for Russell* even though he would never know it.

⁵ Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974), 42-5 and *The Examined Life* (New York: Simon and Schuster, 1989), 104-8.

lives (or, if they are in the experience machine, about *most* aspects of their lives!), and we are supposed to have the intuition that this means things are going badly *for them.* Not only is there broad consensus that such examples undermine state-of-mind theories of well-being, but there appears as well to be a consensus that desire-fulfillment theories (whatever else one may think of them) handle well the extramental dimension of well-being.⁶ The problem that state-of-mind theories face is one desire theories can solve.

However, there is an important distinction that is routinely overlooked in these discussions and which may make a difference to our sense of how well desire theories handle things beyond the mental. When it is said that more matters than mental states, this could mean that certain *mind-independent events* count when it comes to assessing how good my life is, whether I know about them or not. Call this the "mind-independent events matter" thesis (MIE). MIE is central to desire theories. On such views, there is nothing more to desire satisfaction or frustration than an event in the world. The theory treats desire as an attitude towards a possible state of affairs that is satisfied if and only if that state of affairs comes to be. Thus, 'desire satisfaction' has nothing to do with personal or felt satisfaction and, more importantly, nothing to do with *knowledge that the event occurred*.

Alternatively, the idea that more matters than mental states could mean that it is important to have a connection with reality. Call this the "positive value of knowledge thesis" (PVK). Rather than place value on an event by itself, PVK places value on knowledge, i.e. on the obtaining of a certain relation between events and

⁶ Griffin, Well-Being, 9-10; Kagan, Normative Ethics, 36; Roger Crisp, "Well-Being," in Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta (Stanford, CA: Stanford University, 2012), http://plato.stanford.edu/archives/spr2012/entries/well-being/, 9.

⁷ Desire theories go beyond the mental by incorporating MIE, but they are not the only theories to do so. For example, L. W. Sumner thinks that well-being requires what he calls "authentic happiness." Happiness is a purely psychological state, one that combines an attitude of satisfaction with one's life and good affect. Happiness is authentic if it is (a) informed and (b) based on authentic values. However, both (a) and (b) are given counterfactual interpretations, such that happiness is informed as long as it would not be undermined by gaining information, and values are authentic as long as they would not be undermined by awareness of their origins. Only the information requirement is relevant here. Because the information condition is counterfactual, there is no requirement that I actually know something in order for it to make a difference to my level of welfare. This is a version of MIE. See Sumner, *Welfare*, ch.6.

the mind.⁸ In this context, I intend 'knowledge' to mean knowledge of the significant facts about one's self and one's life, not knowledge generally. Desire theories can only indirectly accommodate the thought that a connection with reality is important, namely by allowing that knowledge counts *when individuals desire it*.

Interestingly, no doubt because the distinction between MIE and PVK is not usually made, no one seems to have noticed that the standard examples are ambiguous in the sense that the intuitions called forth could just as easily support the value of knowedge (PVK) or the value of events (MIE). Consider the deceived individuals mentioned above. We might feel that the lives of these individuals are bad because their lives are so very different from what they objectively want them to be. Or we might think their lives are bad because they are so ignorant about their lives. Or our intuitions might be tracking both. The same holds for Nozick's experience machine case. Nozick's own presentation seems to emphasize the idea that subjects in the machine lack all connection with reality. But it is also true that their lives are objectively very different from what they probably want them to be. In desire satisfaction terms, most desires of a person in the experience machine would be frustrated, though the person would not know this. If we are troubled by the experience machine, are we troubled by ignorance, desire frustration, or both?

The easiest way to see just how different MIE and PVK are is to think about how they might come apart. And the easiest way to do this is to consider how—in certain types of cases—they would rank possible lives.

Consider the following four cases and assume they occur within lives that are otherwise identical to one another (so that any differences in the ranking of lives must be attributable to the differences recorded here):

Case 1: My desire for X is satisfied and I know this.

Case 2: My desire for X is frustrated and I know this.

⁸ It is actually unclear whether knowledge is valuable or one of its epistemic relatives, such as justified true belief. I consider these matters later in the article. For simplicity now I will just talk about knowledge.

Case 3: My desire for X is frustrated and I don't know this.

Case 4: My desire for X is satisfied and I don't know this.

Desire theory always places value on the satisfaction of desire. But it only places value on knowledge if the individual (actually or hypothetically) desires knowledge. To make the contrast as clear as possible, consider a person who lacks any desire (actual or hypothetical) to know how things stand. In such a case, desire theory treats 1 and 4 as equivalent and as superior to either 2 or 3. However, a theory that incorporated only PVK—the idea that *knowledge* matters—would rank the same lives quite differently. It would treat 1 and 2 as equivalent and as superior to either 3 or 4. Clearly there is more than one way to go beyond the mental. 10

I have three goals for this paper. First, I aim to convince people of the importance of the distinction between MIE and PVK.¹¹ Whatever else one concludes about things beyond the mental, I hope in the future that theorists of well-being will no longer consider it sufficient to say simply, "I accept the conclusions of Nozick's experience machine: there are non-mental components of well-being." One must also say which extra-mental things or relations have intrinsic welfare value, and whether they *always* have it or only *sometimes* do.

Second, I want to raise doubts about the way current desire theories handle this issue. There are several worries one might have in this regard. First, one might think that events have no intrinsic welfare value apart from our knowledge of them.

⁹ An actual desire theory equates a person's good with what she actually desires. There are relatively few defenders of actual desire theories, though an exception is Mark Murphy, "The Simple Desire-Fulfillment Theory," *Nous* 33: 2: (1999): 247-72. Informed desire theories, on the other hand, equate a person's good with what she would desire if she were fully informed and rational. See for example, John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 407-416 and Peter Railton, "Moral Realism," *Philosophical Review* 95: 2: (1986):163-207 and "Facts and Values," *Philosophical Topics* 14: 2: (1986): 5-31. For my purposes here, the differences between actual and informed theories are of little relevance. So I shall simply discuss desire theories and assume that readers can work out for themselves how particular points would apply to each kind of desire theorist.

¹⁰ Indeed, there are really many ways to go beyond the mental. As we shall see, there are numerous variations on MIE and PVK. I begin with MIE and PVK because they are the two most obvious theses to suggest themselves on the basis of the traditional examples.

¹¹ And between their relatives. See previous note.

If so, then MIE is false and desire theory runs into trouble because of its incorporation of MIE. Second, one might worry (or worry instead) about desire theory's account of the value of personal life knowledge. There are, in turn, two ways such worries might go. There might be cases where, intuitively, personal life knowledge seems like a great benefit but where the individual in question does not desire it, and so desire theory claims it is not good. Less often recognized is the alternative problem, where, intuitively, personal life knowledge seems non-beneficial but where the person in question *does* desire it and so the theory claims it *is* good.

Finally, I want to begin to work out the implications for the theory of well-being of adopting or rejecting various different claims (or combinations of claims) about what matters beyond the mental. If one is bothered by any one of the points raised above, then the question naturally arises: what should a theorist of well-being say about these matters? Can a desire theorist revise the theory in some way to meet these concerns, or is there another subjective theory—a near cousin of desire theory perhaps—that might do so and might be preferrable for that reason? Or shall we discover that the intuitions appealed to here ultimately support some kind of objective theory? I shall first consider these questions in relation to MIE, and then in the second half of the paper turn to considering PVK and its relatives. However, given that this is largely uncharted territory and given the complexity of the issues, I shall not try to reach definitive answers here, but merely try to suggest which are the most promising avenues for further reflection.

§1.0 Are the Limits of Well-Being the Limits of the Person?

I wish to begin by looking briefly at an argument of Shelly Kagan's. There are several reasons why it makes sense to begin with Kagan. First, *if* his argument succeeds it would entail either that very little beyond the mental affects well-being, or that nothing beyond the mental does. Before going further then, we need to assure

ourselves that it is possible to escape his conclusion. Second, it is useful to see how philosophical thinking about the extra-mental aspects of well-being has been shaped by exclusive focus on MIE. Kagan's discussion provides an illustration.

Kagan states the *conclusion* of his argument as "the limits of well-being are the limits of the individual." However, this needs some explanation. What Kagan means is that facts about well-being are limited to fact about an individual's mind *or body*. Now it might seem odd to include facts about the body, but this flows from Kagan's thought that facts about well-being must be limited to facts about a person's non-relational properties. Which non-relational properties ultimately count depends on what *beings like us essentially are*. If we are essentially minds, then well-being depends entirely on non-relational *mental* facts. If, however, our essence includes our bodies, then our well-being could include non-relational facts about our bodies as well.

Kagan reports being sympathetic to the idea that our essence includes our bodies. If that is right, and bodily facts count, then obviously something beyond the mental matters for well-being. However, even if one were to accept that idea, Kagan's conclusion, though not strictly ruling out things beyond the mental, would still be a radical departure from the kind of view endorsed by desire theorists who embrace MIE. Traditionally, such theorists have been interested in defending the idea that the kinds of extra-mental facts *picked out by our evaluative attitudes* affect our welfare. Most of these facts are neither mental nor bodily. However, it is important not to get too distracted by the question of the essence of persons. For whichever of these two claims one finds plausible, the real philosophical work (from the perspective of one interested in the topics of this paper) is done by Kagan's claim that changes in merely relational properties cannot affect well-being. It is this claim that rules out as relevant most of the facts traditionally defended by desire theorists as well as the kinds of facts a defender of PVK might want to include.

¹² Shelly Kagan, "The Limits of Well-Being," in *The Good Life and the Human Good*, eds. E. L. Paul, F. D. Miller, and J. Paul (Cambridge: Cambridge University Press, 1992), 169-89. Also see Shelly Kagan, "Me and My Life," *Proceedings of the Aristotelian Society*, 94: (1994): 309-24.

Kagan's argument for the radical conclusion turns on the notion of *benefit*. He claims, plausibly enough, that a theory of well-being must do more than simply capture the idea of value connected to, or associated with, an individual. It must capture value *for* the individual. If my well-being improves, it must be obvious how this benefits me. But Kagan also thinks that the connection with benefit can only be obvious if I am *directly affected* in some way. He thus assumes that a change in purely relational properties could not constitute a benefit. The argument goes like this:

- 1. Positive changes in well-being are benefits.
- 2. Benefits must involve changes in the person benefited, i.e. changes in her non-relational properties.
- 3. A person is no more than her mind and her body.

(C) Therefore, positive changes in well-being must be changes in the non-relational properties of the person's body or mind or both.¹³

This argument can be difficult to assess initially, because so much turns on the extent to which one finds each premise intuitive, and this in turn seems to depend largely on the kinds of cases one thinks about. In particular, it depends on the extent to which one is focused on MIE. It is both interesting and important that in his discussion leading up to this argument, Kagan rehearses a very familiar story about the development of philosophical theories of well-being. Hedonism, the story goes, was ultimately replaced by desire theory, in part because hedonism could not capture our sense that more matters than just mental states. Desire theory, whatever its other flaws, *seems* able to handle this well, because the satisfaction of desire requires that the world be a certain way. The fact that Kagan rehearses this particular history leading up to his argument, suggests that he is focused on MIE as

¹³ Of course, Kagan can run the same kind of argument in negative terms to explain why something cannot reduce well-being unless it directly affects a person in a negative way.

opposed to PVK when he makes his argument. There is certainly nothing in anything Kagan says to suggest PVK as an alternative thesis about what matters beyond the mental. And of course, whatever the truth of the matter about Kagan's focus, his rehearsal of this development makes it natural for *readers* to focus on MIE.

This is significant, especially if other people's intuitions follow mine, for I certainly find premise (2) more plausible if I am focused on particular implications of MIE. Suppose that A very much wants some event, E, to occur. In story one it occurs. In story two it doesn't. But in neither story does A ever learn what happens with respect to E. Desire theory claims that she is much better off in story one. But it is not too hard to understand the perspective of someone who finds this puzzling. To push the point a little further and establish the link with benefit, imagine that in story number one I am the person who satisfies A's desire. Moreover, I know I am satisfying her desire. However, it is still true that she will never learn what I have done, and I know she will never learn this. It certainly can seem very odd to describe this as a case where I benefit her. In short, the thought that benefit might require a change in non-relational properties [as premise (2) claims] may seem more attractive if one has been thinking about cases where no such change has occurred. and where benefit also seems absent or at least highly questionable. But as soon as one shifts one's attention to other types of cases and other potentially welfare enhancing relations (knowledge, for example) it is less clear that benefit is lacking. At any rate, given the variety of relations that might be thought to matter for wellbeing and the complexity and murkiness of the intuitions in this area, I see no need to treat Kagan's argument as decisive.

§2.0 Do Events Have Prudential Value Apart from Knowledge?

As we have seen, desire theory incorporates MIE—the thesis that certain mind independent facts or events have prudential significance in and of themselves and

independently of whether the subject knows them. Yet, as we have also seen, some of its implications can be highly counterintuitive.

The kinds of cases that test MIE are like (3) and (4) in our earlier ranking exercise. In (3) I desire X, my desire is frustrated, but I never learn this fact. In (4) I desire X, my desire is satisfied, but I never learn this fact. A desire theory tells us that (3) is better for me than (4). Suppose I have dedicated myself to bringing up my children, and define myself in terms of my parenting skills. Of course, I want my children to do well for whatever reason. But I also want it to be true that my parenting makes a large contribution to their flourishing. I would be disappointed if I discovered that my parenting caused them to fare poorly or even if it made no difference at all. Suppose my children do fare well largely because of my parenting, but they live far away and for one reason or another, it is not possible for me to know how their lives are going. My desire is satisfied though I never know it.

Various things seem unobjectionably true in this case. Certain counterfactuals are true: namely, that *if* I knew I *would* be pleased, and *if* the opposite had been true I *would have been* disappointed. Certain evaluative facts are also true: It is certainly good *for my children* that they flourish. It may also be good for others who rely on them or who care about them and who (unlike me) can regularly benefit from seeing how well they do. But it seems strange to say that it benefits *me*.

This is not just a hedonistic intuition. I feel no inclination to say that facts can benefit me only by changing *how I feel*. Moreover, in other cases where I know how things stand in the world, it seems that the facts are the locus of value, and my mental state important only because it is my mode of access to that value. Suppose I accomplish a goal (I run a marathon) and know this. Given that I wanted to run a marathon it now seems that what matters (what contributes value to my life) is *the fact that I succeeded in running one*.

This suggests the following alternative to MIE. Certain events (whether picked out as relevant by desires or picked out as relevant in some other way by

some other theory of well-being) have the potential to affect welfare positively or negatively. But their ability to actually affect welfare is *conditional*. Only if I have knowledge about these facts can their value enter my life. If I have false beliefs or no beliefs about X then the facts about X cannot contribute to my welfare. It is not, as a hedonist would insist, that only experiences possess intrinsic value. On this view, facts external to the mind have intrinsic welfare value. But they only have it if they are mentally grasped. Call this the conditional value of events thesis or CVE.

CVE is not the same as PVK—the thesis I originally contrasted with MIE and which I discuss next. To see the difference compare the following three claims:

- (1) An event, E, can contribute to A's welfare (positively or negatively) independently of whether A knows E. (MIE)
- (2) An event, E, can contribute to A's welfare (positively or negatively) only if A knows E. (CVE)
- (3) Knowing E contributes positively to A's welfare. (PVK)

CVE says that potential value can only be realized when there is a positive epistemic link. But the nature of the value depends on the nature of the facts. In some cases, if the facts are negative (e.g. I learn about a desire frustration) the welfare impact will be negative. In other cases it will be positive. Moreover, CVE does not imply that more true beliefs are better than fewer. CVE places intrinsic value on events or facts in the world and makes an epistemic relation the condition for the realization of this value. PVK, on the other hand, assigns intrinsic value to a positive epistemic relation in itself. PVK does imply that more true beliefs are better, as each additional true belief contributes something positive to welfare regardless of the what the belief is about.

What merits does CVE have? There are certainly some things to be said in its favor. Many people who reject experience machines feel that welfare value lies in real world happenings. If I climb Mount Everest, it is the fact that I really did this and not simply the fact that I think I did it that matters. Nozick is persuasive when he says, "we want to *do* certain things, and not just have the experience of doing them." CVE allows facts to play a role in many cases, but eliminates the bizarre cases where facts I never encounter nonetheless make my life much better or much worse that I think it is.

This view also rids the desire theory of the desires of the dead (which only some will see as an advantage). If the satisfaction of my desire only benefits me when I know about it, then obviously it can't benefit me once I am beyond the point of having epistemic relations. I have never felt any temptation to think we can benefit the dead in the welfare sense, so this seems like an improvement in desire theory from my perspective, but it will not seem so to everyone.

More problematically, the view gives us a motive not to learn certain unpleasant truths. If the negative value of negative events can only enter my life when I know about them, then it seems I should try not to discover these facts. I am not sure what to make of this aspect of the view. At least two points are worth considering here. First, the counterintuitive implications mainly concern third parties—people who are in the position to either hide or share information with us. But it does not give the subject any obvious motivations one way or the other. This is because in real life we don't have much direct control over what we learn. Moreover, we don't know whether something we care about will turn out well or badly until we do know, at whichpoint the damage-if it is damage--is done. The only sure way to avoid learning negative facts would be to try to avoid learning anything. Since on CVE this would typically deprive a person of as many welfare goods as welfare bads, this hardly seems like a promising strategy.

¹⁴ Nozick, *Anarchy*, 43.

Second, and more importantly, if one combines CVE with another thesis such as PVK, whatever residual problem remains would be largely resolved. For then learning something bad would be, at one and the same time, a welfare good and a welfare bad (which actually seems right). Whether it is *overall* beneficial to learn some fact will vary from case to case, depending on how bad the news is, how this compares to the positive value of true belief, and also what the psychological effects of learning the bad news will be.

Let us now briefly consider what objective theories say about MIE. Is it part of an objective theory to claim that extra-mental facts, of which we remain forever ignorant, nonetheless affect our welfare? The answer is not obvious. Desire theorists use an individual's desires to pick out a set of facts in the world that are welfare affecting. Given that desires can range over many things, it is possible for many of the facts selected to be at some remove from the subject whose well-being is in question, and so the issue of whether these facts still count when not known arises quite naturally. It is more common, however, for objective list theorists to focus on identifying the most significant life goods, and to then talk about how subjects need to *have* or *possess* these goods. The very terminology of 'having' or 'possessing' suggests a much tighter relationship between the subject and her putative goods, and so one might wonder whether the issue of having a good but not knowing about it even arises.

Nonetheless, some objective theorists are probably committed to MIE. Take for example the claim that health is an objective good. Many objective theorists have been drawn to this claim. To take it seriously (as opposed to saying that health is instrumental in the production of other intrinsic goods) is to leave room for facts that might not be known. A person can have a disease and fail to know it. Indeed, she might have it for some time without having any symptoms. In most cases, at some point, the disease will affect her and she will become aware of it. But if it is the disease iself that counts as intrinsically bad, then the subject can fail to know about one of the things directly affecting her well-being. It is possible that she might never

know, if she dies in some accident before her illness becomes manifest. This seems to me analogous to the desire theorists claim that a desire frustration never known about nonetheless makes one's life worse.

Just as I find MIE implausible in the desire theory, I find it implausible in objective theories. Better to say that health is an extremely valuable all-purpose instrumental good. It contributes to things that matter intrinsically such as our ability to pursue our goals, enjoy our relationships etc. etc. But it is these things that matter, and they matter when we know about them.

Recently, some theorists with objectivist leanings have defended hybrid views according to which well-being is enjoying the good. On these accounts, there are certain objectively valuable things. Insofar as positive welfare value enters a life it is only because of these. But for that value to enter a person's life the individual herself must appreciate, love, or enjoy the thing in question. Interestingly, loving something or appreciating it presuppose knowing about it. If that's correct, then these hybrid views seem not to embrace MIE, but rather something more like CVE.

I have argued that MIE has some strongly counter-intuitive results and have introduced CVE as an alternative. How appealing theorists find CVE remains to be seen. I myself am undecided about its merits. However, there is nothing in principle to stop either a desire theorist or an objective theorist from adopting CVE. In order to avoid some conterintuitive implications of CVE, however, theorists may also want to adopt PVK or one of its cousins. And as we shall see, that *may* pose a problem for desire theorists in particular, and subjectivists more generally.

§3.0 Is it Prudentially Good to Have an Accurate Picture of One's Life?

¹⁵ Robert Adams, *Finite and Infinite Goods* (Oxford: Oxford University Press, 2002), 83-101; Shelly Kagan, "Well-Being as Enjoying the Good," in "Philosophical Perspectives," suppl., *Nous* 23: (2009): 253-72; Stephen Darwall, *Welfare and Rational Care* (Princeton, NJ: Princeton University Press, 2002), 73–104.

I now consider PVK—the thesis that personal life knowledge has intrinsic welfare value. A familiar tradition in moral philosophy assigns value to knowledge in general, with a special emphasis on *theoretical* knowledge. Knowledge in this sense appears on lists of objective goods, alongside pleasure, virtue, the appreciation of beauty, and loving relationships. However, my concern is not with theoretical knowledge, but with *personal* knowledge. I wonder whether it is good for a person (at least usually) to be epistemically in touch with the reality of her life. Personal knowledge is what deceived spouses and experience machine inhabitants lack. However, the scope of personal knowledge is much broader than simply knowledge of the status of one's desires, though it presumably includes that. Suppose I desire a loving relationship with my sister. My desire is satisfied if we have a loving relationship. But having robust knowledge about my own life requires knowing more than just the fact that we have such a relationship. It requires that I know quite a bit about my sister, that I remember a good many of the things we have done together, and so on and so forth.

PVK seems initially quite plausible. For example, it does seem important to know whether those who act lovingly towards you, really do love you. If you have children that you love, it seems important to know this fact and to know *them* in the sense of knowing and appreciating many significant things about them. If I develop dementia late in life and forget my children, my life seems worse because of this, even if I do not realize I have forgotten and so feel no regret. The loss of memory in such cases, and the loss of personal knowledge it entails, are prudential bads because they cut one off from the reality of one's life. It does seem that *part* of what

1

¹⁶ Probably the most famous historical example is Book X of Aristotle's *Nicomachean Ethics*. Theorists such as G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903) and W. D. Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930) thought of knowledge as having objective intrinsic value, though they did not recognize welfare value (goodness *for* a person) as a distinct value category. A contemporary defender of this view is Thomas Hurka, *Perfectionism* (Oxford: Oxford University Press, 1993). A contemporary objective theorist of *welfare* who assigns intrinsic value to knowledge is John Finnis, *Natural Rights and Natural Law*, 2nd edition (Oxford: Oxford University Press, 2011). Richard Kraut, *What is Good and Why: An Ethics of Well-Being* (Cambridge, MA: Harvard University Press, 2007), argues that there is objective intrinsic welfare value in enjoying the acquisition of knowledge.

bothers people about experience machine cases is the fact that the subject in the machine is so deceived about the real conditions of her life.

Problems emerge, however, if one is inclined, as I am, in the direction of subjectivist theories of well-being. Desire theories claim that knowledge (of whatever sort, personal or otherwise) is only valuable if the subject desires it. To alter desire theory to accommodate PVK might reasonably be thought to conflict with the basic motives that drive subjective theories. Above all, subjective theories aim to avoid directly stipulating what is prudentially valuable. They prefer, whenever possible, to let substantive values emerge from individual attitudes. Objective theories have no problem with PVK, and for that reason some will prefer to incorporate what I say about PVK into an objective theory. But objective theories have other problems at the metaethical level. So there is a real question about how best to square the initial plausibility of PVK with a plausible overall theory of welfare.

Before turning to a consideration of these matters, I will first refine the thesis a bit in the hope of identifying its most plausible form. PVK is not really a single thesis, but a placeholder for a large set of claims one might defend about how standing in a certain epistemic relation (but which one?) to objective facts in the world (but which facts?) can have prudential value.¹⁷ Thus two issues need to be addressed. First, which facts intuitively count as significant life facts? *If* it turns out to be true that we benefit prudentially from standing in some particular epistemic relation to significant life facts, which facts will these be? Second, is it really *knowledge* of these facts that matters or might some other, less demanding epistemic relation—such as true belief or justified true belief—have just as much *prudential* value?

¹⁷ There may also be other kinds of relations that matter—e. g. agential relations for example. Something like this is suggested by Nozick, *Examined Life*, 106. For now I have simply assumed that such matters could be captured in a discussion of the kinds of extra-mental facts that matter, i.e. if it matters that you climb Mt. Everest, then welfare would be enhanced only if you actually climb Mt. Everest.

§3.1 What Are the Significant Life Facts?

Can we say more precisely what the significant facts of a life are? How do we know when a person qualifies as having significant life knowledge?

It is easy to think of intuitively important *categories* of life knowledge, such as knowledge of what one has done, and knowledge of the people in one's life. However, even though we can identify *categories*, we should not assume individuals need to know all or even most of what can be known in each category. Assuming we select the categories appropriately, then if an individual failed to know *anything* in a category, we would think she had a serious gap in her life knowledge. But it remains open how much detail individuals must grasp *within* particular categories.

For example, it seems important to know the various things that you have *done* in your life and, if you acted for a conscious reason, *why* you did what you did. Yet we all forget many of the details of our lives, and while some of this loss might also be a prudential loss, surely not all of it is. When I try to remember most years of my life I quickly generate a short list of significant events, and around each item on the list a thick web of particular memories. My memories of these particulars are dense. Yet it is still true that I cannot say what I did during most of the days in any given year of my life. Would I be significantly better off if I suddenly remembered all of those details? I doubt it.

Although it can seem maddeningly imprecise, it seems intuitively right that, in order to count as knowing the significant facts of our own lives we need to know (a) the big things we have done or that have happened to us—the actions and events which have pushed our lives in particular directions, and (b) a large number of lesser actions and events. We need (b) in order to have a sense of the texture of our lives, in order to fill in the gaps between the 'big' events. But insofar as these "fill-in" events are the ones that have had less of an impact on the shape of the whole, it is less important which specific events a person knows. For example, I would be missing an important chunk of my life if I were to forget that I went to graduate

school in philosophy. But if I only remember *that I went to graduate school for philosophy*, but remember almost nothing from that time period, this seems a lack as well. I need, in addition, a fairly hefty collection of memories from that time that help me to relate it to other times in my life. However, the particular events I remember from then will only be a subset of all the events that happened. Arguably, as long as the set remembered is large enough, it shouldn't matter which particular things I remember. For each person and for each significant category of life knowledge there are presumably many sub-sets of true propositions from that category all of which cross the quantitative threshold for adequate "fill in" knowledge. My suggestion is that knowing the propositions in any one of these sets is sufficient to count as having significant life knowledge.

It also seems plausible that we must know something about the *people* in our lives, at least those who have played a significant role. Moreover, the larger a person's role, the more we must know about her and our relations with her. We must not only know facts about the person, but have a roughly correct sense of the course our relationship with her has taken, including awareness of what we have sought from the relationship, the impact our actions have had on her, some sense of what she has sought from us, and how she feels about us. Again, I am describing *categories* of knowledge. We don't need to know everything in a category, so there is no implication that we need to know everything (or even most) of what others feel about us. But to be completely off-base about such facts in relation to an important person in our life, does seem to count as an instance of failing to know "the significant facts about one's life."

Yet another aspect of life knowledge is knowing how your own projects, goals, concerns etc. have (or have not) been realized in the world, as well as why this is so. This kind of life knowledge—knowing and understanding one's successes and failures—typically also requires a person to situate her knowledge in some larger picture of the society she lives in and the time period she inhabits. It seems impossible to say how much such contextual knowledge a person would need. Yet it

is hard to imagine that someone lacking awareness of all such contextual details could count as having significant life knowledge.

And, of course, much more could be said, both in the way of adding categories and refining the ones so far listed.

Someone might worry about PVK, given that certain kinds of false beliefs about the self are extremely common. Psychologists have maintained for some time that most people hold exaggeratedly positive beliefs about themselves, their ability to exert control over their lives, and about the future. Not only are these positive illusions pervasive, and so completely normal, but there is much evidence that they are *highly adaptive* as well. In other words, certain distorted beliefs appear to have greater prudential value than true beliefs.

However, what the psychological findings suggest is that certain false beliefs have prudential value of the *instrumental* sort, whereas PVK is a thesis about intrinsic value. Moreover, it seems that these common self distortions have instrumental value (as opposed to the delusions characteristic of mental illness, which do not) only because they (1) co-exist well with other *correct* beliefs, and (2) are responsive to environmental feedback.²⁰

Common self delusions can co-exist with correct beliefs, because they do not depart from reality entirely, but rather exaggerate the real in a positive direction. For example, a person who is moderately funny may view himself as very funny, but will not typically see himself as a patient person if, in fact, he lacks patience altogether. Most important of all, common self-delusions are responsive to feedback, even though this is still no guarantee of accuracy. For example, someone with an exaggeratedly positive sense of his own wit will revise his view of himself downward when confronted with clear evidence that others do not see him as

¹⁸ S. E. Taylor, *Positive Illusions: Creative Self-Deception and the Healthy Mind* (New York: Basic Books, 1991), 6.

¹⁹ Ibid, ch. 2.

²⁰ Ibid, 123-43.

funny. But since his original starting point was exaggerated, his revised view, though less flattering, may still not be fully accurate.

The simplest answer to the worry is that we are concerned with intrinsic welfare value, and these beliefs have, at best, instrumental value. There is no contradiction in maintaining that positive illusions have intrinsic disvalue, even though they also have positive instrumental value. However, this strikes me as too simple. In its simple form PVK implies that more knowledge is always better than less. But a modification here may be plausible in its own right *and* help to explain my sense that positive illusions are often entirely harmless. What has intrinsic welfare value is having significant life knowledge—a picture of one's life that is accurate, covers all the main categories, and is fairly rich in details. Significant life knowledge is thus a threshold concept with a fairly high threshold. For someone who already counts as possessing significant life knowledge—someone who has met the threshold conditions—we should say that acquiring additional knowledge is value neutral. For many people then, if their positive illusions form a relatively small part of their total set of personal beliefs, and if they are otherwise well-informed, their illusions need not detract from their welfare.

§3.2 Do We Really Need Knowledge?

From a certain perspective it is natural to assume that knowledge is what matters, particularly since in ordinary life we talk rather loosely about knowledge, as if all of our true beliefs have this status. Yet there are other epistemic relations, such as true belief (TB) and justified true belief (JTB), which *might* also matter prudentially. Moreover, we should be open to the possibility that from the standpoint of *prudential* value, TB and JTB might be *just as good as knowledge*.

It is helpful to distinguish here between epistemic value and prudential value. The relationships between these two are complex, and it is only by carefully

untangling them that we can hope to answer our question about prudential value. I shall discuss each in turn.

In the realm of *epistemic value*, knowledge is the supreme value, superior to both TB and JTB. It has this status because, relative to TB and JTB, it has more of what matters from the epistemic point of view: freedom from epistemic luck.²¹ The traditional story about why true belief is not knowledge is that true belief could simply be the result of luck. For example, suppose Mary believes that her project has been a huge success because she is the kind of person who wants to believe good things about herself and will believe such things *no matter what*. The results of her efforts only become obvious at a point when Mary herself is abroad with no way to get any updates on how her project is doing. Still, she believes things are going well, and in this case she turns out to be right. We would hardly say that she *knows* this.

The traditional response to such cases was to analyze knowledge as justified true belief. In contemporary epistemology two different visions of what justification might be have emerged, but there are problems with equating knowledge with justified true belief on either version. On the first view (JTB^E), S is justified in believing p if, "it is reasonable or rational, from S's own point of view, to take p to be true."²² In other words, S is justified in believing p, if S's *evidence* uniquely and strongly supports belief in p. Gettier type cases have led most people to conclude that JTB^E is not knowledge, again because of the potential role for luck.²³ Suppose Claire looks at a clock to see what time it is and forms the correct belief that it is 4pm. Unbeknownst to her, the clock she is looking at is broken and always says 4pm. She has good reason for her belief (clocks are usually good sources of information about the time, and she has no reason not to trust this clock), but in this case the truth of her belief is a mere accident. This disqualifies it as knowledge.

²¹ Steup, Matthias, "Epistemology," in *Stanford Encyclopedia of Philosophy* ed. Edward N.Zalta (Stanford, CA: Stanford University, 2012), http://plato.stanford.edu/archives/win2012/entries/epistemology/. Section

^{1.1 22} Ibid, Section 1.1.

²³ Gettier, Edmund, "Is Justified True Belief Knowledge?" *Analysis* 23: (1963): 121-23.

On the other more contemporary approach to justification (JTB^R), S is justified in believing p if p originates from, "reliable cognitive processes or faculties."²⁴ On this view, justification ensures that S's belief has a high objective probability of being true. But once again examples have called into question whether JTB^R is really knowledge. Goldman's famous barn examples make this point.²⁵ Suppose I am, unbeknownst to me, driving through barn-façade country where most of what look like barns are really just elaborate barn facades. I happen, however, to see the only real barn in the area and develop the belief that I am seeing a barn. My belief is true, and it was developed through a reliable cognitive process (vision). But again my belief seems too much based on luck to count as knowledge.

I have rehearsed this familiar story because it explains the natural hierarchy of epistemic values within the realm of epistemic value. In this realm, what matters intrinsically is reduced reliance on luck. Luck, in turn, provides a natural entry point into a discussion of the prudential value of various epistemic goods. To begin with, if anything is true, it is that TB has instrumental prudential value. Without at least a roughly accurate picture of our world we could not achieve other intrinsic welfare goods. Moreover, given the kinds of creatures we are and the kind of world we inhabit, our natural interest in true belief gives us a strong prudential reason to want to reduce the role of luck in our belief formation processes. This prudential interest in reducing luck is broad based. At the general level we are less concerend about any particular belief, and more concerned that large numbers of our beliefs be true over lengthy periods of time. Luck just is another word for the absence of reliability. The presence of luck indicates that we cannot reasonably hope for the same outcome twice.

These considerations explain why the natural hierarchy of value we find in the realm of epistemic value is not mirrored in the realm of prudential value, at least insofar as we focus on *prudential instrumental* value. In the case of any particular

²⁴ Steup, Matthias, "Epistemology," Section 1.1.

²⁵ Alvin Goldman, "Discrimination and Perceptual Knowledge" *The Journal of Philosophy* 73: (1976): 771-91.

true belief p, there is no greater *instrumental* prudential value if p is also justified (in either sense) or if p qualifies as knowledge. Suppose I very much want to play professional baseball and suppose that this is not an unreasonable hope given my talents and that playing professionally would in fact be good for me. I live in a remote town where talent scouts never visit, but I get a not-to-be-repeated chance to travel to a city to play in a statewide "all-stars" game where many talent scouts will be. My coach first learns of this chance, but he is drunk at the time and writes down the wrong date. He then passes this on to me. Later in the day the wind blows the sheet of paper out of my hand, and I must simply rely on my vague memory of what it said. I misremember, but (by fluke) my faulty memory of what was written is correct about the date. I show up, play, and am recruited. In this case, having a true belief about when the game will be is incredibly instrumentally important. But there is no greater instrumental value present in a slightly different story in which my belief about the date is justified (ITBE or ITBR) or in one where my belief is knowledge. It is certainly true that since I can't possibly hope for the world to treat me so well in general, I have strong instrumental reasons for wanting more reliable epistemic links to the world, and for trying to cultivate in myself a degree of epistemic virtue. But in any given case, all that seems to matter from the standpoint of prudential instrumental value is having TB.

What then about *intrinsic prudential* value? If one is at all inclined to assign intrinsic welfare value to epistemic statuses, then it seems one should, at the very least, say that it is intrinsically better to have true beliefs than false beliefs. But is it intrinsically better to have JTB as opposed to TB, and better still to have knowledge? I suspect not. The value hierarchy we see in epistemology reflects our concern with epistemic luck. And I have argued that we have strong *instrumental prudential* reasons to be concerned with luck. But I doubt that epistemic luck makes an intrinsic welfare difference. Indeed, we find further support for this view if we consider other types of welfare goods. In these cases too, luck seems to have little or no impact on intrinsic welfare value. It doesn't generally matter whether we obtain

welfare goods through some reliable method or simply by luck. What matters is just that we attain them.

Consider a non-epistemic case. Suppose it would be intrinsically good for Rebecca to marry George. Let us suppose that this is grounded in the fact that they are temperamentally suited to one another, and if they were to marry they would have a wonderful life together, supporting each other in many positive ways. Now consider two possible stories. In story one, Rebecca and George have known each other for a long time, and everyone around them recognizes what a good fit they are for each other. When they begin to date, their friends are very happy for them and expect good things. They grow to love each other, and no one is surprised when they marry. Their marriage is a great one. In story number two, Rebecca and George do not know each other. Rebecca's parents meet George and his family and come to believe he would be a good match for Rebecca. But Rebecca refuses even to be introduced to George, resenting what she sees as her parents' meddling. She makes plans to go abroad to study for the next eight years. But one evening, a few weeks before she is supposed to leave, she runs into George in a local park and they start to talk. She is surprised to discover how well they get along. She has a change of heart, puts off her plans to go abroad, and starts seeing George. Shortly afterward they get married and their marriage is a great one. In the first story a good thing comes to Rebecca in a straightforward and reliable way, while in the second story it is a fluke of luck that it comes to her at all. Yet this hardly seems to make a difference to our assessment of how well her life goes. All that matters prudentially is that she marries George.

In a similar way, I think that from the standpoint of prudential value, what matters intrinsically is true belief. JTB and knowledge also have intrinsic prudential value because they include true belief. But my claim is that they do not add intrinsic value to true belief. Consider the miraculous case of Julia. Julia had a rich and interesting life, and while she was alive she had true beliefs about all the significant facts of this life. But despite having such a large set of true beliefs, all of these beliefs

were true by fluke. This is why her life was miraculous. She was, let us say, the most epistemically lucky person ever to live. Of course, it is hard to imagine how such a case could come about. Given that many of her beliefs are about quite personal things it is hard to imagine how she could have had a large number of such beliefs that were both true and yet never related in the proper way to her evidence (JTBE). Similarly, because she had so many true beliefs it is hard to imagine how the cognitive processes that gave rise to them could all fail to be reliable (JTBR). After all, consistent truth as an outcome is normally good evidence for reliable cognitive processes at work. So it is indeed hard to get a good grip on a case of someone who has a large set of merely true beliefs. Only with some hefty and no doubt bizarre science fictions could we make it work. But let us import some such science fiction (I leave it to the reader to develop one she likes). Though a life like Julia's is an unlikely reality, we can still ask whether it was lacking a degree of intrinsic prudential value in virtue of its epistemic failings. I don't see how it is.

Remember that, unlike someone in the experience machine, everything she believed was *true*. The world was a certain way, and Julia had an acurate picture of how it was. For example, when alive, she believed that she loved Charles, that she had been married to him for years, and had traveled the world with him. She believed they both enjoyed their African travels the most, that they both loved unusual scenery and gourmet food, and that they had enjoyed these things together on many occasions. She also believed that Charles loved her, and that he appreciated her best features. Since her beliefs were true, she *really did* travel the world with Charles and so forth. He really did love her, and really did appreciate her best features. In short, her life contained valuable actions and happenings, and she was aware of these. I find it hard to believe that the further removal of reliance on epistemic luck would have made her life intrinsically better *for her*. As she was living this life, it would have made sense for anyone concerned about her to want her to have reliable epistemic access to truth (which she didn't have). But that desire

makes sense, given that no one outside bizarre philosophy examples can rely on epistemic luck.²⁶

I have not yet committed myself to the idea that any epistemic status has intrinsic welfare value, though I find the idea intuitively plausible. What I have done is suggest that insofar as one is inclined to explore this territory, one should agree that true belief has immense *instrumental prudential* value. Moreover, because of this, *epistemic virtue* and *a reliable epistemic connection to reality* have instrumental prudential value. JTB and knowledge, however, have no additional instrumental prudential value over and above true belief. And when we turn to questions of *intrinsic prudential* value, I suggest we should say the same. If epistemic statuses have intrinsic prudential welfare value, then TB has such value. But JTB and knowledge are not superior to TB in intrinsic prudential value. They have equal intrinsic prudential value to TB, because they are themselves instances of true belief.

Despite the fact that my preferred version of PVK has turned out to be a thesis about true belief, for the rest of the paper I will continue to talk about knowledge. In ordinary speech it is actually quite hard to substitute "true belief" for all the ordinary references to knowledge. I shall also for the sake of continuity continue to use PVK as the label for the thesis, though a more accurate acronym would be PVTB.

§3.3 What Should We Ultimately Say About PVK?

²⁶ No doubt many readers will not be fully persuaded that TB is as good as JTB. I confess I have struggled with this myself. However, I have ultimately come to think that what matters prudentially is simply having an accurate picture of reality. Interestingly, Nozick in *The Examined Life*, (106) suggests that knowledge is important because the kind of contact with reality we want is one in which we are tracking reality. Many readers might feel the same though be willing to settle for JTB. However, Nozick's point is ambiguous. Why suppose that only knowledge or JTB involve tracking reality? After all, if you go through life with all true beliefs, then even if this happened by luck there is still a sense in which you were tracking reality (your *tracking* was merely a matter of luck, but you were tracking reality since your beliefs continuously adjusted to the world).

Should a theorist of well-being try to incorporate something like this revised version of PVK? And if so, can a subjectivist theory accomplish this, or would adoption of PVK push us towards an objective theory?

If one is sympathetic to PVK then it is clear that current versions of desire theory won't do. For they claim that personal knowledge is only important when desired. Even if many people in fact want to know about their lives, presumably some do not. However, it may seem to run counter not just to desire theory but to subjectivism more generally to insist that personal knowledge has intrinsic value. An objective theorist doesn't have this problem.

It seems to me, however, that there are certain moves open to desire theory that might allow one to accommodate PVK. And whether or not the revised theory is one that desire theorists would willingly recognize as a variety of their own view, it would clearly be a form of subjectivism. The first step is to recognize some of the ways in which personal knowledge differs from other kinds of substantive prudential goods. Personal knowledge is *knowledge of (and knowledge related to)* the things we care about. Arguably, however, part of what it is to care about something is wanting to know about it. Caring is a complicated attitude, but at least part of concern for X is the desire to know X: to see it, or appreciate it, or know that it is. And another part of caring about X is caring what happens to X, and this in turn makes one want to know how things stand with X, what is happening in the vicinity of X and so on. Indeed, to lack all concern for the facts about X is good evidence for a lack of concern about X.

Now this suggests a way to argue that, in fact, people do quite universally want to know about the things that matter most to them and which affect their welfare. Of course, it may not be a desire they recognize having. A person may only have the desire in the sense that she possesses certain unconscious dispositions. But that is enough. A desire does not have to be conscious for a desire theorist to appeal to it. It is also true that many people will be conflicted, also possessing conscious desires *not* to know certain kinds of facts. Suppose I devote much time to a

particular project. On the view under consideration my genuine concern for the project entails that I desire to know about it. So the modified desire theory could say knowledge about the project is intrinsically good for me. Yet I may now suspect that it has not gone so well and, fearing confrontation with my failure, I may actively desire *not to know* how the project turns out. Thus the theory would also have to allow that not knowing is also good. The reality of conflicted desires means that in many cases knowledge would be both prudentially good and prudentially bad, and the net value of knowing would depend on many details of the case. If one wants to adopt a pure form of PVK where personal knowledge is always an unalloyed good, this would not appeal. But I actually find the pure form too simplistic. It seems right to me that in many cases the value of personal knowledge is a complex question with reasons pointing in different diretions. So if this kind of subjectivist view could be made to work, I would not see that particular feature of it as problematic. The view would still have an advantage over current desire theory in that it would allow us to see personal knowledge as having presumptive intrinsic welfare value.

However there is one further kind of case that theorists of either persuasion should take seriously and that seems to pose a problem for desire theory even in the revised version I have sketched. The original problem was that there might be cases where intuitively personal knowledge counts as a good but where a subject doesn't desire it. I think there are also cases where, intuitively, knowledge is not good, but where the subject does desire it. These are cases where learning something is psychologically devastating. Many types of news might make us less than fully happy, but this doesn't seem to count against the idea that knowledge might be intrinsically good. What I have in mind are extreme cases.

L.W. Sumner describes a case that might—depending on how we fill in the psychic details—count. Consider the case of Asta, whose son has been killed in war. She wants for it to be true that he did not suffer, but she also wants to know the

truth whatever it may be.²⁷ Eventually she learns both that he suffered greatly and many of the grisly details of his suffering and death. She is devastated by what she learns, forever altered, forever haunted by the images created for her by the soldier who tells her the truth. I do not know the details of how Asta's story continues. But there is a plausible continuation of Asta's story in which she is psychically undermined by this news. Yet it also seems plausible to me that some individuals might *want* such news, and might want it even knowing what it will do to them. Our desire to know can be deep and in such cases can be our undoing. Of course, I don't mean to imply that all cases of bad news are alike. There may be people who could learn what Asta learns and yet have their happiness only temporarily blighted. The real issue is about the long-term psychic effects on the individual—the extent to which knowledge of certain facts can cause suffering. I think knowledge that causes suffering cannot count as good for a person even if she wants it. Subjectivists and objectivists alike will have to find a way to limit PVK in such cases.

§4.0 Conclusion

I have argued that there is more than way to go beyond the mental. I have distinguished between two broad types of claims, which I labeled MIE and PVK respectively—both suggested by traditional examples like the experience machine. Each of these is really, in turn, a placeholder for a number of views one might develop about (a) how facts in the world contribute to welfare and (b) how positive epistemic relations do. I remain ambivalent about MIE and its cousin CVE, but I find PVK intuitively plausible. Yet, I am also drawn to subjectivist views, and it is not entirely clear that a revised desire theory (or some other subjective theory) can easily incorporate PVK. Objective theories have no problems of this sort—objectivists are free to pick and choose among the theses I have described. However, there are other disadvantages to being an objectivist about well-being. Sadly, there

²⁷ This story is cited by L.W. Sumner, *Welfare*, 96-7.

is no space to work out all of these matters here. But I hope to have opened up a new, hopefully exciting area of inquiry for theorists of well-being and to have at least laid the groundwork for the development in the future of my own considered view.