

The Experience Machine and Mental State Theories of Well-being

JASON KAWALL

Brown University, Providence, RI, USA

1.

In discussions of theories of human well-being it is now something of a tradition to mention mental state theories, usually crude forms of hedonism, and to quickly dismiss them with a brief summary of Robert Nozick's experience-machine thought experiment, and perhaps an additional example in which the particular author feels that a person's well-being is not exhausted by the mental states she experiences. I will defend mental state theories against such objections.

How are we best to understand the notion of personal well-being, or welfare with which we are concerned? A good initial grounding can be found in the following passage from L. W. Sumner:

Common sense tells us that a person's welfare, or well-being, is a matter of how well she is doing, or how well her life is going, or how well-off she is. To speak of a person's welfare is therefore plainly to evaluate her life. Lives, however, are complex things whose value can be assessed from a number of standpoints. Welfare assessments concern what we may call the prudential value of a life, namely how well it is going *for the individual whose life it is*.¹

The same basic notion seems to be what concerns utilitarians, according to James Griffin in his book, *Well-being*:

Utilitarians use our rough, everyday notion of "well-being", our notion of what it is for a single life to go well, in which morality may have a place, but not the dominant one. This does not mean that our job is merely to describe the everyday use. It is too shadowy and incomplete for that; we still have to be ready for stipulation.²

Griffin does allow that some stipulation may be necessary, but our common-sense notion is to play a central role in discussions of well-being.

What is a mental state account of well-being? The central feature shared by such accounts is that a person's well-being is improved or diminished only by events which enter into her mental or psychological life. If an event does not have an impact upon an agent's mental life then it can have no impact upon her well-being. Among mental states we may include all aspects of our mental lives, such as desires, beliefs, pleasures, pains, and character traits. Our well-being, according to any mental state account, will be constituted by our mental states, or some subset of them.

Let us consider the objections to mental state theories of well-being which are grounded in the intuitions drawn out by Nozick's experience machine. Nozick first presents an example that has captured much attention in *Anarchy, State, and Utopia*:

Suppose there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences, we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences for, say, the next two years. After two years have passed, you will have ten minutes or ten hours out of the tank, to select the experiences of your *next* two years. Of course, while in the tank, you won't know that you're there; you'll think it's all actually happening. . . . Would you plug in? *What else can matter to us, other than how our lives feel from the inside?* Nor should you refrain because of the few moments distress between the moment you've decided and the moment you're plugged in. What's a few moments of distress compared to a lifetime of bliss (if that's what you choose), and why feel any distress at all if your decision *is* the best one?³

The example is compelling. Even though we would have nothing but the experiences and mental states we desired, we still feel extremely little inclination to have ourselves hooked up to such a machine. This seems to suggest that our well-being involves more than our mental states. Nonetheless, there are crucial flaws in Nozick's example when taken as an instrument to throw mental state theories into doubt.

2.

Many authors seem to believe that the experience machine example poses a difficulty for mental state accounts of well-being in particular.⁴ However, the

example does not actually focus on a flaw in mental state theories. To see this, imagine that you are offered a chance to lead an extremely happy and productive life which would also produce more happiness for others. However, to lead the alternative life you must leave your home, family and friends, and never see them again. Most people would refuse such an offer, just as they would refuse the offer of being placed on the experience machine. We would reject alternative lives in which we would be guaranteed more pleasure, or in which we would maximize overall utility, or in which we would have great accomplishments, if the alternatives required us to abandon our current lives and commitments, whether or not an experience machine is involved. This shows that the experience machine itself plays little role in our attitudes when we consider the case. We are not rejecting a life on the experience machine in particular. We would reject almost any other life, in the actual world or otherwise, if it requires abandoning our commitments.

We might wonder whether our intuitions with respect to the alternative lives should be trusted, to the extent that we have not experienced the alternatives. Following Mill, we might hold that in order to compare two experiences competently, we must have undergone both experiences.⁵ However, such a requirement is too strong. Suppose that a man has never experienced being shot with a pistol. Surely he can be justified in preferring to have the experience of sipping a fine whiskey to being shot. His knowledge of guns, injury, and pain allows him to have competent, rationally justifiable preferences to avoid being shot. More generally, we can reasonably predict the nature of experiences we have not undergone given relevantly similar experiences and general knowledge. Given this, our preferences and intuitions concerning the alternative lives and the experience machine can be taken as rationally justified and worthy of consideration.

While the experience machine example does not pose a difficulty unique to mental state accounts of well-being, it still might seem that any such account must fail, given that Nozick's example does show us that we value various commitments more than mental states. This is the case even if the thought experiment does not focus exclusively on mental state theories. Thus, we must consider the relationship between our deep commitments, those which we would not voluntarily abandon, and our well-being. In particular, we must consider whether a mental state account of well-being can accommodate such commitments.

3.

In attempting to explain why they believe mental state theories of well-being are inadequate, both Nozick and Griffin stress that we value many things besides our experiences and mental states. Nozick writes: "We learn that

something matters to us in addition to experience by imagining an experience machine and then realizing that we would not use it.”⁶ Griffin makes a similar claim:

According to the enjoyment account [a particular mental state theory of well-being], what affects well-being can only be what enters experience, and the trouble is that some of the things that persons value greatly do not. My truly having close and authentic personal relations is not the kind of thing that can enter my experience; all that can enter is what is common to both my having such relations and my merely believing that I do. And this seems to distort the nature of these values. If I want to accomplish something with my life, it is not that I want to have a sense of accomplishment.⁷

Griffin is unfair to mental state theorists, when he insinuates that they must claim that we hold our various values simply as means to experiences or mental states. Will Kymlicka makes a similar misleading statement:

What we want in life is something more than, or other than, the acquisition of any kind of mental state, any kind of “inner glow,” enjoyable or otherwise. We do not just want the experience of writing poetry, we want to write poetry; we do not just want the experience of falling in love, we want to fall in love. . . . Human well-being is something more than, or other than, getting the right sequence of mental states.⁸

All that a mental state theorist is required to claim is that the impact on well-being of such commitments and values is a matter of the mental states which they produce. She is not committed to the further, dubious claim that the values are held in order to produce the mental states. A mental state theorist need not hold that we merely want the experience of writing poetry. She can consistently maintain that we desire to actually write poetry. She is also not committed to the claim that we desire to write poetry in order to produce certain mental states. Simply put, for a mental state theorist, there need be no limits on what we may value or on why we value what we value. The only claim she is committed to is that our values and commitments will affect our well-being only through our mental states.

We can outline a simple argument intended to show that mental state theories of well-being are inadequate:

- (1) We would not abandon our commitments and values, even if such behaviour were to lead to better mental states for us. This is reflected clearly in our reactions to the experience machine.
- (2) Hence, it is clear that we value more than our own mental states.
- (3) Therefore, there is more to our well-being than mere mental states.

This argument is not presented so explicitly by Nozick, Griffin, or Kymlicka, but it does seem to accurately capture the basis of their rejection of mental state theories. If this is an accurate rendering of their reasoning, we find that their rejection of mental state accounts is ill-founded.

Even if we grant that we value our various commitments, this does not show that we must abandon mental state theories of well-being. While it is true that we value many things besides our mental lives, it could well be that the other values contribute to our personal well-being only through the effects they have on our mental lives. For example, consider a collector of stamps. He treasures his stamps, and is extremely pleased when he finds a rare Penny Black. While the collector may value the stamps themselves, and completing his collection, it is compatible with this to suggest that his hobby contributes to his well-being only through the mental states it leads to: from the pleasure he derives from finding rare pieces, the patience he develops, or the goals it gives him to aim for.

Similarly, while we may value our jobs and personal relationships and be unwilling to give them up for a life on the experience machine, this does not show that it is not through the mental states that they produce that they contribute to our own well-being. Again, it seems plausible to maintain that the impact of our commitments upon our well-being is simply a matter of the pleasures we derive from them, the good character they build, and the goals they give us.

A related point is that almost everyone believes that we can value more than our own well-being; indeed, this seems to be commonsense wisdom. A parent might greatly value the welfare of his child and a scientist might be committed to finding the truth or developing a new theory. That people have such values and would be willing to sacrifice their mental states for such values does not directly show that they contribute to our well-being. After all, it is possible for a person to sacrifice her own well-being for the sake of one of her commitments. Parents can sacrifice their well-being for the well-being of their children, just as scientists can give up many valuable experiences to devote themselves to their labs and the search for the truth.

If we value things besides our mental states, they can produce rival values to our own well-being. The values can come into conflict. Thus, it is clear that the mere existence of other values does not show that they must necessarily contribute to our well-being. There is a large gap between the premiss that we value more than our mental states and the conclusion that there is more to our personal well-being than our mental states. Seeing our commitments and values as rival values to our well-being allows us to properly understand cases in which it might seem that more contributes to our well-being than our mental states. Consider a soldier who throws himself upon a grenade in order to save his comrades. In this case it seems most natural to hold that the soldier has sacrificed his own well-being for the well-being of others, for the sake of his

commitments to their well-being, or for the well-being of his country. Griffin's position should be seen in a similar light. He would be willing to sacrifice his own well-being for the sake of his commitment to genuine interactions should he be forced to choose. The other commitments and values are rivals to well-being, and do not play an obvious role in our well-being beyond the positive impact they can have upon our mental states.

Still, we should consider one more objection. Someone may hold that there is no conflict between our various commitments and our well-being. When we consider the soldier who throws himself on the grenade, his action seems rational and even admirable. He has accomplished something and obtained what he values, even if this does not enter into his mental states. It may be said that we have been presupposing a narrow view of well-being that involves only mental states. But a great deal more can enter into our well-being. The noble act of the soldier shows this. The essential point of the objection is that in cases in which we sacrifice pleasant or valuable mental states for the sake of a commitment it seems there must still be a contribution to our well-being in some sense. After all, it helps us to achieve or obtain something which we value.

But a mental state theory can account for the intuitions which push us to hold that in cases of apparent sacrifice, the agent's well-being is still improved. In the case of the soldier we admire his courage and integrity, which are mental traits. The core intuition is based on the good mental states that such an act reflects and produces. As well, our intuitions are shaped by our imagining what the agent would feel. The soldier would likely feel extremely guilty throughout his life if he were not to jump on the grenade. Griffin would be devastated if he were to eventually find out that he had been living through a pleasant deception, without genuine interactions. We are imagining the mental states the agents would have, and if they do not actually have them it becomes much harder to see how their actual well-being is affected. Finally, the reason that we admire such actions is just that they involve a sacrifice of the agent's own well-being. The objection would seem to make such actions merely prudent on the part of the agent, as all such actions would contribute to the agent's well-being. This seems to run contrary to our ordinary, intuitive notion of well-being. Surely we can make genuine sacrifices for the sake of others. Griffin does wish to allow for some stipulation in his characterization of well-being, but treating acts of sacrifice as mere prudence seems to stretch our commonsense notion of well-being too far. Mental state theories are much more in accord with our intuitive notion of personal well-being.⁹

Notes

1. L. W. Sumner, "Two Theories of the Good," in Ellen Frankel Paul, Fred D. Miller, Jr., and Jeffrey Paul, eds. *The Good Life and the Human Good* (New York: Cambridge University Press, 1992), p. 4.

2. James Griffin, *Well-being: Its Meaning, Measurement and Moral Importance* (New York: Oxford University Press, 1986), p. 7.
3. Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 42–43.
4. See, for example, Griffin, *op. cit.*, p. 9; Nozick, *op. cit.*, p. 44; and Will Kymlicka, *Contemporary Political Philosophy* (New York: Oxford University Press, 1990), pp. 13–14.
5. See John Stuart Mill, “Utilitarianism” in J. M. Robson, ed. *Collected Works of John Stuart Mill* (Toronto: University of Toronto Press, 1969), vol. X, pp. 211–214.
6. Nozick, *op. cit.*, p. 44.
7. Griffin, *op. cit.*, p. 19.
8. Kymlicka, *op. cit.*, p. 14.
9. I would like to thank Dan Brock, Jeremy Fantl, Catin Lechapeau, an anonymous referee, and the editor for their helpful comments and suggestions.

