

## Assignment

According to Nozick, there is something wrong with any view that takes the quality of experience to be what matters in a good life. What is Nozick's argument? Is it convincing? (Why or why not?) Does Brandt's version of the desire satisfaction theory avoid Nozick's argument? Do you think Brandt's view is an acceptable alternative to an experiential quality theory? (Why or why not?)

## Introduction

The question of what matters in living a good life has long been debated in philosophy. I explore what experiential quality theories, as a group, have to say about this question and its possible answers. I consider and argue for Nozick's refutation of such theories in "The Experience Machine," including my own justification for a crucial implicit assumption that Nozick seems to make. I consider desire satisfaction theories, and Brandt's in particular, explaining how Brandt's theory side-steps Nozick's experience machine. I extend Nozick's argument to criticize Brandt's theory, additionally bringing up other flaws. Finally, I suggest a slight modification of the concept of desire satisfaction, and briefly sketch how a theory based on this modification can resolve the flaws I point out in Brandt's theory.

## Experiential quality theories

Experiential quality theories assert that the quality of experience is what matters in a good life. By this I mean that we may only consider what a person experiences, or perceives, when judging whether or not that person leads a good life. We may not, for example, say that in order to live good lives, people must hold true views about the world; the truth of a view is not something that exists entirely in the mind.

### Support for experiential quality theories

While experiential quality theories, as I've described them, may seem counter-intuitive at first glance, there are valid arguments to adopt them. Imagine two people who are identical in every respect; they do the same thing in identical worlds. Any reasonable notion of what makes a good life should tell us that either both of these people lead good lives, or neither does; there is no means to distinguish between them. Imagine now that one of these people is actually just a 'brain in a vat'; there is a computer that provides the world, the interactions, the sensations, the experiences. Imagine that this illusion is so perfect, so complete, that, were you one of these people, it would be impossible for you to tell which of them you were. If we want to say that the person in the illusion can still lead a good life, then we must accept that what matters in leading a good life is the experience, the perception, and not the reality.

A similar argument appeals to our desire for certainty. If I ever want to *know*, for sure, whether or not I lead a good life, I must consider only things I can know for sure when I determine what makes a life good. Since a perfect illusion is, on the inside, indistinguishable from reality, I cannot require any things that exist only in reality and not in illusion; I must consider only experiences.

More formally, I assume that people may seek to live good lives, and that this seeking does not prevent them from living good lives. Since people who seek things generally desire to know when they have found what they seek, since they do not want to continue seeking something after

they have found it, people seeking to have good lives would want to know when they lead a good life. I make the reasonable assumption that some rational, knowledgeable people can lead good lives; I assume that the good life is not only for the irrational and the ignorant. I assume that it is irrational for people to attempt to achieve things that they know to be unachievable.<sup>1</sup> Then, if sufficiently knowledgeable, rational people attempt to live good lives, they would seek to be able to determine whether or not they are living a good life, and doing so could not be irrational. Since it is impossible to distinguish a perfect illusion from reality, and a sufficiently rational and knowledgeable person would know this, it follows that only our experiences, our perceptions, the things we can know for sure, can matter to us in living good lives.

### Nozick's "experience machine"

Robert Nozick argues<sup>2</sup> against experiential quality theories by proposing an experience machine, and asserting that while any experiential quality theory would say to use the machine, we should not and would not want to use the machine.

**Description** Consider a perfect experience machine, a perfect simulator.<sup>3</sup> From the inside, it is indistinguishable from reality. Any experience that is possible, this machine can simulate. It can cause you to feel any physically possible emotion, sensation, or feeling.

**Nozick's argument: Doing** "First," Nozick claims, "we want to do certain things, and not just have the experience of doing them." Nozick draws the distinction between things that we want the experience of doing, such as skydiving, and things that we want to have actually done, like community service and helping people. If I "go" skydiving, and later find out that it was a simulation, I think I would not be significantly saddened; I sought skydiving because I wanted to experience it. If, on the other hand, I "performed" community service, and later found out that it was just a simulation, I would feel like I had wasted my time. My purpose had been to help others, and this purpose was not achieved. The fact that there are actions, such as community service, that I seek to *do* and not just to experience, implies that I would have reason to not enter the experience machine.

**Nozick's argument: Being** Second, Nozick claims, "we want to be a certain way, to be a certain sort of person." Nozick asserts that this is improbable, if not impossible, in an experience machine. For example, if I wanted to experience successfully negotiating peace between warring nations, the machine would present me with such an experience. Because such an experience would require me skilled at diplomacy, if the machine gave me only consistent experiences, the machine would cause me to think the things required to be diplomatic, despite the impossibility of me otherwise thinking those things. Similarly, if the machine were to give me the experience of leading an ancient army to victory, it would cause me to not feel fear, or to not feel the desire to flee from fear, when faced with danger. Thus, it is impossible to determine the character of a person who has long been in an experience machine, because we cannot tell the difference between thoughts and reactions given to the person by the machine, and thoughts and reactions that the person would have come up with

---

<sup>1</sup>Some may argue that we may accomplish meaningful things by attempting the impossible. I would counter that, if this is a sufficient argument, then what you actually care about is achieving those (possible) goals, and not achieving the impossible goal; anything that can be achieved by being irrational in this manner can also be achieved by being rational, and mimicking the actions of the irrational person in order to achieve possible goals.

<sup>2</sup>Nozick, "The Experience Machine".

<sup>3</sup>Wikipedia, *Brain in a vat* — *Wikipedia, The Free Encyclopedia*.

naturally. Nozick goes further, saying that such a person *is* not any particular way; so long as I am in such a machine, my actions are no more determined by my “natural” reactions than by what the machine deems necessary to the experience.

**Nozick’s argument: Meaning** Third, Nozick argues, many people want to live a meaningful life. People like to believe that they have a purpose in living, that they are a part of something bigger and more meaningful than themselves. An experience machine ensures that any meaning in your life is shallow, no deeper than what the creators of the machine can provide to you, and ensures that you are *not* a part of something bigger or more meaningful than yourself. Nozick cites the conflict over psychoactive drug as evidence; some shun them as experience machines, while others embrace them as a means to a deeper reality, a more meaningful life.

Though this concludes Nozick’s argument against experiential quality theories, he goes on to delve deeper into the issue of why we wouldn’t want to use such an experience machine. Nozick proposes a series of machines, each of which fixes deficiencies in the previous one. Nozick concludes that the most disturbing thing about such machines is that they live our lives for us; for any but the most basic of uses, these machines control so much of the experience that who I am becomes muddled with what the machine is, and it can no longer be said that *I* am living or experiencing things. As Nozick says, “perhaps what we desire is to live (an active verb) ourselves, in contact with reality.” The more we utilize an experience machine, the less contact we have with reality, with ourselves, with the universe outside the experience machine.

**Implicit assumption: Our wants and desires matter in determining what is a good life.**

Without this assumption, the discussion is useless as a refutation of experiential quality theories, because it would tell us only about what we want, and not about what is meaningful to leading a good life. I justify it based on what I see to be more well-accepted assumptions. I assume that morality and ethics, that what makes for a good life, have some internal basis; I assume that there are justifications for doing the right thing and for leading a good life, and that these do not consist simply of following a plethora of unjustifiable rules (e.g., given to us by a higher being). I claim that any reasonable justification for action is given in terms of desires. “It’ll make me happy,” you tell someone. “You want me to be happy,” you implicitly assume. “If you don’t do this, you’ll regret it later,” you warn. “You don’t want to regret having done something,” you implicitly assume. You wouldn’t expect the argument “it’ll bring about world peace” to convince a warmonger to act, just as you wouldn’t expect the argument “it’ll show that our country is militarily superior to all others”<sup>4</sup> to convince a pacifist to act. Similarly, any reasonable justification to live a good life must appeal to our desires; at least some of our desires must matter in what makes a life good. I call any theory that abides by this principle a *desire-based theory*.

Even if we take this assumption, that our desires matter in what makes a life good, for granted, Nozick does not justify that these particular desires matter in considering what is a good life. It is conceivable, for example, that the desire of a drug addict to consume more of a particular drug is not relevant to living a good life, and should in fact be removed in order to live a good life. While this may be true of these desires for some people, I assert that it is possible for a rational, knowledgeable person to hold these desires, or desires are effectively the same. If these desires are sufficiently important to these people, such that there is no argument that would convince these people to give up such desires, and these people can live good lives, then it must be possible for some considerations other than the quality of experience to matter in what makes a life good.

---

<sup>4</sup>Thanks to Amy Zhou for the advice that “patriotic arguments motivate warmongers.”

I now resolve the apparent contradiction between this argument and the argument I presented earlier in favor of experiential quality theories. The simplest resolution is a refutation of the assumption that rational people may seek a good life. I argue by analogy: I do not, for example, seek a single thing called a “happy life,” and I do not seek to know when I have achieved a “happy life.” I seek to be happy. Similarly, I seek to live *as good a life as I can*, and I seek to know which choices are *better* than other choices. More generally, I do not seek to know whether or not I have achieved a goal, but when I am approaching and when I am receding from a particular aim. My argument in support of experiential quality theories falls apart when considered in this new context.

## Desire satisfaction theories

A desire satisfaction theory is one that argues that what is important in leading a good life is the satisfaction of desires. In the strictest sense, these hold that if all of my desires are (fully) satisfied, then my life is good. Various theories make refinements on this claim, specifying the relevance of different (classes of) desires.

### Brandt’s theory

Brandt proposes<sup>5</sup> a theory of “rational desires,” and asserts that a good life is any life that a person would rationally desire to live.<sup>6</sup>

### Rationality

Brandt defines a rational desire to be a desire that persists in the face of careful cognitive psychotherapy. Brandt defines cognitive psychotherapy to be repeated presentation of all available relevant information in an ideally vivid way at appropriate times.<sup>7</sup> He cites empirical evidence to support his claim that such “cognitive psychotherapy” can effect changes in desires.

### Nozick’s machines: yes or no?

Brandt’s theory, as stated, can easily side-step the issue of Nozick’s experience machine. If, after careful cognitive psychotherapy, I still desire to not use an experience machine, Brandt’s theory would say that I should not use it. My assertion above, that knowledgeable, rational people can desire the things that make experience machines repulsive, is precisely the claim that the desire to not use an experience machine is rational for at least some people. Set in this light, Brandt’s theory, and in fact any desire satisfaction theory, provides a meaningful context in which to ask questions about Nozick’s machines; they give us a means for evaluating the relevance of our (potential) desires to not use the machines. In this context, it is not meaningful to propose a sequence of machines, each of which shores up any deficiencies in the last; there need not be any particular reason to be repulsed by the experience machine, nor any particular way to fix it. Perhaps what we (rationally) want is, as Nozick claims, to live and be in contact with reality. If this is the case, then Brandt’s theory supports Nozick’s argument.

---

<sup>5</sup>Brandt, “Goodness as the Satisfaction of Informed Desire”.

<sup>6</sup>Page 635: “The first is that there is no sentence in which the word ‘good’ appears, at least in that core complex of uses which have been important for philosophy, which makes an identifiable point which cannot be made by a sentence containing ‘rationally desired’, doubtless in some complex clause but in which no ‘value-word’ is present.”

<sup>7</sup>Brandt’s definition of cognitive psychotherapy and the relevant terms is sufficiently nuanced that I could not do it justice with a brief explanation. I instead assume the reader’s familiarity with the terms, and include a large excerpt from “Goodness as the Satisfaction of Informed Desire” at the end of this paper (Appendix A).

There are, however, two more sinister cousins of Nozick's machines which we must consider.

**Genies** Consider what I call the *genie machine*, a machine which grants to us anything we wish. Brandt's theory says that we should first wish to hold only rational desires, and then wish for all of our desires to be fulfilled. Whether or not we should view a life achieved in this manner as "good" or desirable is not clear to me. This raises the question of whether or not I would have only rational desires if I were living the life I rationally desire to live. I return to this point shortly, but leave the larger question of genie machines for the reader to ponder, as beyond the scope of this paper and my current ability to decide or argue.

**Inverse-genies** I now come to the more immediately troublesome machines, which I call *inverse-genie machines*. Consider a machine that, instead of molding reality to our whims, molds our whims to reality. Imagine a machine that constantly replaces the unsatisfied desires of a person with the desire for the world to be the way it is. A simpler machine might simply remove any presently unsatisfied desires. Do not imagine that a person using such a machine would be unsatisfied with life; the machine would not allow this. If we were to integrate this type of machine with a person, then all of this person's desires, and lack thereof, would pass Brandt's test for rational desires. Indeed, such a person would likely be the most satisfied of any of us, never having any unsatisfied desires. However, I would claim that such a life repulses most, if not all of us.<sup>8</sup> Assuming this reaction, and that this reaction is rational (which I believe it is), it is clear that such a life is not good, is generally not rationally desired.

## Problems

Consideration of the genie machines brings up ambiguities in the idea of a "desire satisfaction theory." Do we demand the satisfaction of current desires, whatever they may be, or do we permit modification of these desires? If we permit modifications, how do we decide what modifications are permissible?

There are additional problems in Brandt's theory in particular, in general due to his lack of consideration of "meta"-questions. The most obvious such question is that of the desirability of rational desires. Brandt does not eliminate the possibility of rationally desiring to not undergo cognitive psychotherapy. Given how time consuming such therapy seems to be, it should not be surprising that it may be better to allow most, if not all, of our desires to exist, whether or not they are rational. More troubling, his theory suggests that some desires, which we generally consider rational, are actually irrational, and that, despite, this, it is rational to want to have these desires. Consider, for example, the desire for education. Ideally vivid presentation of the information that would be learned via education clearly eliminates the desire to learn such material; having already learned it, I have no desire to be taught it again. Nonetheless, if there is something that it benefits me to know, and I do not know it, I am better off if I want to learn it than if I don't care; it is rational to want to desire to learn some things that I do not know. Brandt's theory thus admits inconsistencies, where it is irrational to desire something, but rational to want to desire that thing.

Finally, in the chapter given us, Brandt does not actually present his theory as a desire satisfaction theory; a good life is not, according to Brandt, one in which our rational desires are satisfied,

---

<sup>8</sup>An example of what a person using the simpler inverse-genie machine can be found in the case of António Damásio's patient, "Elliot."<sup>9</sup> A brain tumor caused him to lose the ability to experience emotion. The result was a complete change in character and pathological indecision. Since most emotions, I think, are fundamentally related to desires, a person using an inverse-genie machine would likely be similar to this Elliot.

<sup>9</sup>Ergenzinger, *Conversations With Phineas Gage: A Neuroscientific Approach to Negotiation Strategies*

nor one in which we desire only rational things, but one that we would rationally desire to live. This allows the question of who is doing the desiring. If I, currently, rationally desire to live a particular life, does that make it a good life? Must I continue desiring to live it while I actually live such a life?<sup>10</sup> If, having a particular life, I rationally desire it, is my life good? What if, as I am now, I (rationally) do not want that life?<sup>11</sup>

## Solutions

I propose some possible modifications to Brandt's rational desire satisfaction theory, which solve the problems that I brought up, and, hopefully, provide an acceptable alternative to both experiential quality theories and desire satisfaction theories.

I begin by analogy. Consider a person acutely aware of the troubles caused by poverty. Imagine growing up, never being sure of how palatable your next meal would be, when you'd get it, or even if you'd ever have a meal again. Imagine never being sure of whether you'd come home to a meal and a bed, or the landlord telling you that your family had been evicted. Imagine further that, having grown up only a few blocks away from a lavishly rich neighborhood, you are also very aware of the luxury of the rich. It is conceivable that a person who grew up in a situation like this would come to identify money with all the things that were lacking from childhood and, perhaps, teenage years. If this person lives an unexamined life, money might slowly come to be valued as intrinsically good; as memory of the specifics of childhood were slowly forgotten, so too might the reasons for desiring money be forgotten. We would say that this person has mixed up causes and effects, has turned a means to ends into an end in itself.

I claim that we may be doing the same thing with desire, satisfaction, happiness, *et cetera*. Consider pain. Some hedonistic theories would say that pain is intrinsically bad. I argue that these theories have misidentified cause and effect. In most cases, at least when we restrict ourselves to the historical origins of pain, the realm of biology, it is not the case that something is bad for us *because* it causes us pain, but that we have evolved to respond to things that are bad for us *by* experiencing pain. Foods are not good for us because they taste good; we have evolved to find (natural) foods that are good for us pleasant tasting. It is a small jump to infer that, if our desires are relevant to morality, to ethics, to a good life, then it is not the satisfaction of our desires that makes something good, but the goodness of something that causes us to desire it.

Some might argue that different people desire different things, and that this must mean that things are not intrinsically good. I would say instead that different people perceive different aspects of things and understand things to different degrees, and that what and how they desire is a reflection of their understanding. I claim that those who fully understand and fully experience something

---

<sup>10</sup>Some might counter that it is not rational to desire to live a life that, having it, we would not want. While I generally agree with this statement, it might be possible to concoct a scenario, like the one involving education above, were different levels of desire would disagree.

<sup>11</sup>We may go further and consider opposite extremes of this situation. Plato's "The Allegory of the Cave"<sup>12</sup> describes a primitive person chained to a cave wall who sees only shadows. Plato argues that it is right to forcibly drag this person into the sunlight, into the three-dimensional world. It is, however, irrational for the cave-dweller to want to ascend, because the information about the non-shadow world is not publicly known or knowable (from the cave-dweller's point of view). We might like to say that the life of the ascended cave-dweller is better than the life of the unascended cave-dweller.

On the other hand, we may consider someone who is to go under the influence of an inverse-genie machine. Despite such a machine causing this person to desire to live the life he or she is currently living, we might want to say life under the influence of this machine is not as good as life free from its influence.

Brandt's theory, as stated, does not allow us to distinguish between these cases.

<sup>12</sup>Plato, "The Allegory of the Cave"

will not disagree on its inherent desirability relative to other fully understood and fully experienced things. I say that a desire is *appropriate* in the extent to which it results from understanding of that which is desired. I say that a desire is *complete* to the extent to which it approaches the desire that would result from complete understanding of that which is desired.

Recall that, in order to avoid the argument for experiential quality theories, I asserted that I do not seek a “good life,” but instead seek to live as good life as possible. Following this, I say that people may better their lives by bettering their understanding of things and better appreciating goodness, and by satisfying desires with preference to more true and more complete desires.

## Resolutions

I now go over the objections I raised to Brandt’s theory and desire satisfaction theories in general, and explain briefly how a theory like what I propose here can avoid or resolve the objections.

The resolution to the question of inverse-genies and modification of desires is immediately obvious: changes that result from improvements in understanding improve the quality of life, and changes that result otherwise degrade the quality of life, and similarly with the admissibility of desires in deciding what makes a life good.

It is also clear to me that consistency is not a hard condition to satisfy in any theory based on these ideas; it is reasonable to assume that the universe is consistent, and that as we improve our perception and understanding of it, the consistency of our view of the universe improves. A concept of good that is, in some sense, intrinsic, should therefore be consistent as well. More concretely, if we assume that good is a self-advocating concept (that is, it is good to be good, it is better to be better, it is bad to be bad), then it follows that full understanding of something yields consistent desires on all levels, and as we approach full understanding, we improve consistency.<sup>13</sup>

## Conclusion

NOT YET WRITTEN

---

<sup>13</sup>Incidentally, this provides a rough metric of possibility for goodness; if we desire something, but desire to not desire it, or vice versa, then this might indicate that we do not fully understand this thing.

## References

- Brandt, Richard B. “Goodness as the Satisfaction of Informed Desire”. In: *A Theory of the Good and the Right*. Revised. Prometheus Books, June 1998, pp. 622–636. ISBN: 9781573922203. URL: <http://amazon.com/o/ASIN/157392220X/>.
- Ergenzinger Jr., Edward R. *Conversations With Phineas Gage: A Neuroscientific Approach to Negotiation Strategies*. Dec. 2002. URL: <http://www.mediate.com/articles/Ergenzinger.cfm>.
- Nozick, Robert. “The Experience Machine”. In: *Anarchy, State, And Utopia*. Basic Books, 1977, pp. 42–45. ISBN: 9780465097203. URL: <http://www.amazon.com/o/ASIN/0465097200/>.
- Plato. “The Allegory of the Cave”. In: *The Republic*.
- Wikipedia. *Brain in a vat — Wikipedia, The Free Encyclopedia*. Sept. 2010. URL: [http://en.wikipedia.org/wiki/Brain\\_in\\_a\\_vat](http://en.wikipedia.org/wiki/Brain_in_a_vat).

## A Cognitive psychotherapy in Brandt’s “Goodness as the Satisfaction of Informed Desire”

The critique to come will show that some desires, aversions, or pleasures would be present (or absent) in some persons if their total motivational machinery were fully suffused by available information; and will show how to identify such desires, aversions, or pleasures for a given person. Less metaphorically, the aim is to show that some intrinsic desires and aversions would be present in some persons if relevant available information registered fully, that is, if the persons repeatedly represented to themselves, in an ideally vivid way, and at an appropriate time, the available information which is relevant in the sense that it would make a difference to desires and aversions if they thought of it. By ‘ideally vivid way’ I mean that the person gets the information at the focus of attention, with maximal vividness and detail, and with no hesitation or doubt about its truth. I mean by ‘available information’ ... relevant beliefs which are a part of the ‘scientific knowledge’ of the day, or which are justified on the basis of publicly available evidence in accordance with the canons of inductive or deductive logic, or justified on the basis of evidence which could now be obtained by procedures known to science.

We need to restrict further the kind of information that qualifies as ‘relevant’ in order to guarantee that the effectiveness of the information is a function of its content. If every time I thought of having a martini, I made myself go through multiplication tables for five minutes, the valence of a martini might well decline. But obviously the desire for a martini is not misdirected simply if it fails to survive confrontation with the multiplication table in this way. Any desire would be discouraged by this procedure. We want to say that a thought is functioning properly in the criticism of desires only if its effect is not one its occurrence would have on any desire, and only if its effect is a function of its content. It must be a thought in some fairly restricted way about the thing desired; for instance, a thought about the expectable effects of the thing, or about the kind of thing it is, or about how well one would like it if it happened, and so on.

The claim is that relevant available information, if confronted on repeated occasions, affects our desires. But how often is ‘repeated’? We cannot be, and need not be, precise on this. If several representations have no effect, we can reasonably infer that more of them would do no better. But if several representations have some effect, the question



arises what would happen if there were more? Presumably what we are after is the asymptote: what desires, aversions, and pleasures a person would have if the number of representations increased without limit, if the reflection had maximal impact by representation as often as you like.

Finally, I have said that the self-stimulation by representation of relevant information should come at ‘an appropriate time’. When is that? Obviously, when the effect will maximize counterconditioning, inhibition, or relevant discrimination. For instance, the appropriate times for a smoker to reflect on the bad consequences of his habit are (1) just after inhaling, when the reflection may destroy any pleasure he ordinarily takes in the cigarette; (2) when he wants or is thinking about lighting a cigarette, when the reflection will tend (a) to set up an association between the idea of the state of affairs and contrary motivation, or (b) to make it clear that the total outcome will in fact not be pleasant with the inhibiting effect of that reflection; and (3) when he is having thoughts or images which tend to make the idea of smoking exciting, so that the reflection will associate these mental occurrences with a less glamorous image. But it is not clear that such reflections will be fruitless on any occasion when one is thinking about an outcome.

This whole process of confronting desires with relevant information, by repeatedly representing it, in an ideally vivid way, and at an appropriate time, I call *cognitive psychotherapy*. I call it so because the process relies simply upon reflection on available information, without influence by prestige of someone, use of evaluative language, extrinsic reward or punishment, or use of artificially induced feeling-states like relaxation. It is value-free reflection.

I shall call a person’s desire, aversion, or pleasure ‘rational’ if it would survive or be produced by careful ‘cognitive psychotherapy’ for that person. I shall call a desire ‘irrational’ if it cannot survive compatibly with clear and repeated judgements about established facts. What this means is that rational desire (etc.) can confront, or will even be produced by, awareness of the truth; irrational desire cannot. It is obvious, of course, that desires do not logically follow from the awareness which supports them; the relation is causal and sometimes involves other desires, aversions, or pleasures.<sup>2</sup>

Brandt, “Goodness as the Satisfaction of Informed Desire” in *A Theory of the Good and the Right*, pages 623–623.

---

<sup>2</sup>If a desire required to support another desire, the rationality of which is being assessed, is itself irrational, we must say that the desire being appraised is irrational.