




According to the desire theory of welfare, why are mental disorders harmful?

LSE-Bayreuth Student Conference
May 2-3, 2019

Paper Submission by:

Annabelle Riedesel
Matrikel No. 1401721
University of Bayreuth
BA Philosophy and Economics
Bismarckstr. 25
95444 Bayreuth
s5anried@uni-bayreuth.de



Contents

Introduction	3
Part 1: What is good for me?	3
1.1 <i>The desire theory of welfare</i>	3
1.2 <i>Two central problems</i>	4
Part 2: Harm and mental disorders	6
2.1 <i>What is harm?</i>	6
2.2 <i>What are mental disorders?</i>	6
Part 3: Why are mental disorders harmful?	7
3.1 <i>Two interpretations</i>	7
3.2 <i>The problem of harm vs. suffering</i>	9
Conclusion	10
Works cited	11

According to the desire theory of welfare, why are mental disorders harmful?

Introduction

In our general understanding, mental disorders are harmful. We may consider them harmful because they cause negative emotions, inhibit our life goals, or deprive us of valuable goods. If harm is the opposite of welfare, then mental disorders lower people's welfare. Welfare describes how well the life of a particular person is going for her. Knowing what constitutes welfare is crucial to understanding what exactly makes mental disorders harmful. The desire theory of welfare holds that welfare consists in getting what one wants. Along this interpretation, mental disorders are harmful because they prevent their victims from getting what they want.

A good theory of welfare should on the one hand be able to categorize those instances as harmful that we generally consider harmful. On the other hand, it should not find harm where we do not think it is present. I argue that at times, the desire theory of welfare is too narrow as it fails to account for certain harms caused by mental disorders. Other times, it is too wide as it allows to classify phenomena as mental disorders that we would consider to be within the boundaries of normal behaviour.

The first part of this essay is dedicated to outlining the desire theory of welfare. For this, I will contrast it with two other mainstream theories of welfare: the mental state account and the objective list account. I will then explain the challenges that each of the other accounts pose to the desire theory. In the second part, I discuss what is meant by harm and mental disorder. I will use George Graham's concept of mental disorders. In part three, I will analyse why mental disorders are harmful according to the desire theory of welfare. I will show that the same flaws of the theory highlighted in the first section cause problems in explaining why mental disorders are harmful.

Part 1: What is good for me?

1.1 The desire theory of welfare

Welfare can be judged subjectively or objectively. This is also the aspect in which the most well-known theories of welfare differ from one another. Mental state or hedonist accounts are usually purely subjective. By their simplest definition, welfare consists of pleasure and harm consists of pain. Whatever happens outside our perceptions is not necessarily relevant to welfare. Their main difference to desire theories of welfare is the Experience Requirement (cf. Griffin 1986, p.18). The Experience Requirement holds that welfare is only affected by some event when the subject is aware of it. For example, the mental state theory holds that if an important wish is fulfilled but I do not know it, my welfare remains unchanged. Both the content of my desire and its fulfilment have to be judged by myself, subjectively.

The desire theory of welfare denies this in stating that some things we value are outside of our experience. While the content of my desire is my own business, its fulfilment or frustration is only judged objectively (cf. Crisp 2017). Thus, the desire theory of welfare is partly subjectivist and partly objectivist. What I want should be up to me, but whether I get what I want is up to objective judgment.

Objective List accounts hold that there are goods that inherently contribute to our welfare. They are good for us whether or not we want them, and whether or not we realize that we have them. Usually, “getting what one wants” can be part of an objective list, too. But it would be weighed against other goods such as knowledge, health, and freedom. The desire theory of welfare cannot accept this, as it holds that nothing can be good for a person if it is not appreciated by them (cf. Heathwood 2014, p.202). Concerning welfare, desire satisfaction is the ultimate value and not one among others.

Table 1: simple contrast of the main theories of welfare

	mental state theory	desire theory	objective list theory
content of the desire	Subjective judgment	Subjective judgment	Objective judgment
fulfilment of the desire	Subjective judgment	Objective judgment	Objective judgment

Modifying the desire theory without transforming it into a sub-account of one of the other theories is a balancing task. The content of the desire should not be open to objective evaluation, or else the theory falls into the objective list account. And the fulfilment or frustration of the desire should be reality-oriented, or else the desire theory is nothing but a particular mental-state theory (it could be called preference hedonism). I will refer to these two constraints as the subjectivity constraint and the reality-orientation constraint.

1.2 Two central problems

The demarcation against the two other theories of welfare is also where we can find major problems of the desire theory. First, supporters of the mental state account ask, how can something be good for us that we do not experience? And second, objectivists argue, what if we want things that are bad for us?

The first charge reveals fundamental differences in the conceptions of welfare that I cannot resolve here. It is about whether or not we value reality-orientation. A famous example in favour of the reality-orientation constraint is Nozick’s experience machine (cf. Nozick 2013, p. 264f.). Nozick holds that most people would not prefer a more pleasurable life that is merely simulated to a normal life. However, some mental state theorists may disagree with him on that. All we can safely say at this point is that if the reality-constraint is to be respected, then we have to accept that imagined desire fulfilment or frustration is not affecting our welfare. For now, it is important to remember this as either a central flaw or fundamental characteristic of the desire theory of welfare. Concerning desire

fulfilment, it “ignores the individual’s positive experiences and emotional states (Badhwar 2014 p.53)” and only focuses on states of affairs¹.

More work has been dedicated to solving the second issue. Desire fulfilment can lead to unwanted outcomes. Thus, it seems problematic to claim that it always contributes to welfare, let alone that it equals welfare. According to Daniel Hausman (2011, p.85), the fact that something is preferred does not make it good for that person. He admits that most of the time our preferences are good evidence of what constitutes our welfare, but they do not equal welfare without constraints. A response to this objection are informed desire theories. In distinction to actualist desire theories that take into account any present desire felt by an individual, these theories set quality standards for desires. They seek to find out under which conditions desire satisfaction constitutes welfare, and when it does not. Desires grounded in false beliefs, logical error and psychosis are malformed and therefore cannot lead to good outcomes (cf. Superson 2005, p.112).

However, it is difficult not to violate the subjectivity constraint here and fall into an objective list account. The informed desire must be expressed in a *value-neutral manner*. If we admit that some things contribute directly to our welfare, whether we desire them or not, we move into objective list accounts (cf. Heathwood p.213). For desire theories, all goods other than desire satisfaction are merely instrumental. By finding a way to judge someone’s desires without referring to an objectively valuable outcome, informed desire accounts may be saved. Richard Brandt calls this process “cognitive psychotherapy” (cf. Brandt 1998, p.113). Suppose George wants B, but B is bad for him. Thorough reflection should make him arrive at the conclusion that B is bad for him. Though plausible, this reply still carries an objectivist taste, for B is still bad no matter what George thinks of it.

A truly value-free situation would sound more like this: George wants A, and because he thinks that B brings about A, he wants B. However, B does not bring about A. Therefore, George is worse off by getting B. Here we only judge George’s reasoning process and not the desired end itself. Instead of calling a desire bad because it is directed towards a bad thing, we may call it bad because it stands in the way of getting what we really want. This way, we need to prove that something went wrong in the construction of the desire, without judging the content of the desire itself. Thus, value neutrality is preserved by replacing the judgment of *what* is desired with a judgment of how the desire came about.

Another example illustrates that ordering desires can help judging them. Suppose that George wants both A and C. A is more important to George than C. If fulfilling his desire for C directly frustrates his desire for A, then desiring C clearly was not informed. Hence, whether or not more important, deeper or global desires are prioritized is a good indicator for informed desires. Not all uninformed desires are also incoherent with deeper desires. For example, I may have the spontaneous, uninformed desire to read a philosophical book. It turns out that this helps to fulfil my deeper desire to know something about the world. To avoid confusion, I will

¹ For this reason, I refer to desire fulfilment rather than desire satisfaction in this essay.

call uninformed desires that are incoherent with more important or global desires *distorted desires*.²

Part 2: Harm and mental disorders

2.1 What is harm?

Most straightforwardly, I can assert that the frustration of informed desires is harmful. By definition, distorted desires are causing the frustration of informed desires. Thus, the fulfilment of distorted desires is harmful as well.

At this point it is important to note that there are two ways that desire fulfilment contributes to welfare. Desires are on the one hand instrumentally good because they lead to the satisfaction of other desires. On the other hand, getting what one wants is also good in itself (cf. Heathwood 2014, p.214). As I think it reflects how we perceive desires, I take it as given that there is some intrinsic value to any desire satisfaction. If there is some harm to every desire frustration, then coercion is always harmful, even if it is justified for other reasons.

Because desire fulfilment is both intrinsically and instrumentally contributing to welfare, desire frustration can be intrinsically and instrumentally harmful, too. In this sense, distorted desires are harmful when fulfilled, but there is also harm when they are frustrated. Thus, their mere existence is harmful to varying degrees. Whether or not a distorted desire should be frustrated to ensure the fulfilment of other desires depends on how much worth is attributed to the intrinsic value of desire fulfilment. If the concept of deep or global desires is used, the intuition is that the deeper the desire, the more intrinsic harm is caused by its frustration.

There is another relevant aspect of harm within the desire theory of welfare: harm is not the same as suffering. Suffering is the subjective experience of harm. As there is no Experience Requirement in the desire theory of welfare, harm and suffering do not necessarily occur at the same time. The gap may be bridged by formulating suffering as the frustration of the desire to feel good. However, this formulation moves the desire theory very close to a mental state account of welfare.

2.2 What are mental disorders?

Mental disorders not only cause desire frustration, but change and distortion of the desires themselves. In contrast to physical illnesses, it is therefore difficult to view mental disorders separately from the person suffering from them. This is why their discussion is relevant to the desire theory of welfare. George Graham (2013, p.164f) proposes a concept of mental disorders in which four key features have to be considered. First, they are caused by an impairment in a person's *basic psychological capabilities*. Second, they cause "harmful or potentially harmful symptoms [...] for the person (and perhaps others)" (Graham p.164). Third, they

² Similar terms have been used in the literature on desires. For example, Anita Superson (2005) writes about the "deformed desires" of oppressed women.

are caused both by psychological activity and “brute arational causes or mechanisms” (Graham p.164). And fourth, this leads to a *partial* truncation in rationality, so some rationality is always preserved.

Graham provides a list of basic psychological capabilities that are essential to mental well-being. They include bodily/spatial self-location, historical/temporal self-location, general self/world comprehension, and recognizing and acting on opportunities, among others (Graham 2013, p.157). It is the impairment of these capabilities that is harmful.³

Part 3: Why are mental disorders harmful?

3.1 Two interpretations

The informed desire theory’s interpretation of mental disorders could look like this:

Interpretation 1:

Mental disorders are harmful because they reflect an impairment in someone’s basic psychological capabilities which causes an *inability to form informed desires*.

This interpretation implies that any desire of a person suffering from a mental disorder will be distorted. However, Graham’s last condition holds that some rationality is always preserved. The incapacity of reason is only partial. Therefore, being diagnosed with a mental disorder cannot disqualify all of the subject’s desires. All we can safely say is that a part of her desires will be distorted. Şerife Tekin (2011, p.20f.) addresses the double-edged sword of the mental disorder diagnosis. On the one hand, it provides clarity and potential help for the problems some patient faces. On the other hand, such a diagnosis will diminish the confidence that she and others will have in her desires. To avoid stigmatization and unnecessary frustration of the remaining informed desires, a more nuanced interpretation of mental disorders may be helpful.

Interpretation 1 is insufficient in explaining what exactly constitutes the harm caused by mental disorders in a given moment. It also does not account for her remaining informed desires. Another way to explain why mental disorders are harmful is to refer to the conflict between the different types of desires. A person who suffers from a mental disorder is still partially rational and partially autonomous. Therefore, both distorted and informed desires may be present.

Interpretation 2:

Mental disorders are harmful because the impairment in basic psychological capabilities leads to the *formation of distorted desires that conflict with the remaining informed desires*.

³ Note that following Graham’s concept gives ground to doubt whether certain conditions should be classified as disorders. For example, the author refuses to count Down’s syndrome as a disorder, as it does not cause suffering. The presence of harm is therefore a prerequisite of the phenomena I am writing about.

Under this interpretation, harm can arise in various ways. First, it may stem from the direct frustration of an informed desire. The formation of distorted desires is not the only consequence of impairments in the basic psychological capacities. For example, the impairment can cause anxiety attacks, confusion about the world, and a sense of lost autonomy. There may also be harm from the frustration of a distorted desire. This refers to the intrinsic value of desire fulfilment. If desire fulfilment is valuable in itself, then fulfilling as well as not fulfilling my distorted desires is harmful to me. Still, most importantly, harm is caused by the fulfilment of a distorted desire that in turn frustrates more important informed desires. For example, a depressed person may want to stay in bed all day, thereby frustrating her informed desire to care for her children.

There are some problems with this account as well. First, I concluded before that the most value-free conception of an informed desire is to define it as one that is coherent with more important desires. As established in part 1, the only way a desire is disqualified is if it interferes with higher-order desires. It cannot be disqualified by referring to an objective value. However, we might have difficulty to interpret the deepest or global desires caused by a mental disorder this generously. For example, someone could develop a deep longing for death because of their disorder. If someone's deepest desires are distorted that much, their ultimate welfare may consist in something we cannot reasonably accept. This is even more tricky considering that I stated that the intrinsic harm of desire frustration is greater the deeper the desire.

However, without violating the subjectivity constraint we cannot judge such distorted deep desires. It is difficult to find a reply to the objectivist charge here. No informed desire was frustrated but we would still say that the person is worse off by fulfilling their desire. Some informed desire theorists may allow for a value judgement concerning our most important desires. This could be a judgement of excellence or perfection. However, I think requiring that our life choices point towards something valuable is still too objectivist as it ignores the possibility that we just do not know what is best for a specific person.

A pragmatic reply is that in extreme cases we have to accept that a desire we cannot relate to is now central to that person and thus should be respected. Examples like the man who suddenly is happy to spend his days counting grass (cf. Parfit 1984, p.499f.) would not be considered mental disorders. And, the reply holds, if welfare is truly assessed from the subjective point of the grass-counter then we should accept that there is no harm present. Still, this could only work if we do not believe in treatment for mental disorders. Treatment may revise the desires and cause higher welfare in future. Thus, as long as there is the possibility to at least partially restore the impairment in psychological capabilities, we cannot simply ignore and dismiss the most problematic cases. Surely, if there is no hope for recovery, then we are unable to deny that counting grass is now the best option for that person. It is also relevant to know how great the harm from treatment will be. If it is significantly greater than the desire fulfilment created by re-established informed desires, then treatment is questionable on utilitarian grounds.

Additionally, the desires felt before the impairment took place have to be considered as well. Arneson (1990, p.174) responds to the grass-counter problem by asking whether the person desired to become a grass-counter in the first place.

He imagines a procedure that changes desires so that one only wants to count grass. As this activity is easy to perform, now there is more desire fulfilment than before. However, he holds that this does not count if we did not want to undergo the procedure in the first place. Similarly, a mental disorder may turn someone into something like a grass-counter. Along Arneson's line of argument, the disorder would be an unwanted change. Then, the harm from the disorder would be determined by the extent to which desires prior to the disorder are frustrated.

A second issue that applies to both interpretations is that there still may be social conventions influencing what counts as a mental disorder and what does not. The use of basic psychological capabilities attempts to determine this value-freely. However, mental disorders are usually diagnosed by their symptoms. Often the impairment, which in itself can be hard to measure, is explained backwards by referring to the symptoms it causes. For example, an impairment in general self/world comprehension may be detected by the formation of the distorted desires it caused.

However, we are then unable to distinguish which desires stem from a disorder and which are just socially deviant. Consider a 20th century man who feels attracted to other men. His desire conflicts with his informed global desire to get married. This conflict causes profound harm to him, so that he is diagnosed with the disorder "homosexuality". While the harm to his global desire is significant, its causes do not lie in impairments of psychological capabilities. To rule out such diagnoses, the connection between the impairment in psychological capability and desire conflicts needs to be very sound. This issue highlights the importance of establishing what exactly is causing harm. Conflicts between distorted and informed desires can arise in any normal person. The fact that there is significantly great harm should not be reason to diagnose a disorder.

3.2 The problem of harm vs. suffering

The charge from mental-state accounts introduced in the first part poses an important problem to both interpretations of harm as well. The desire theory holds that some things are good for us even if we do not experience them. In the same way, some conditions cause harm even if there is no suffering or vice versa. The first direction is covered by the desire theory. For example, a severely deluded person who does not realize that her desires are frustrated may not suffer. However, the desire theory would correctly detect that her desires are harmful, and that she needs treatment.

The distinction between harm and suffering within the desire theory causes more controversy in the opposite case. There may be suffering without apparent desire frustration. What makes many mental disorders tragic is that people think their desires have been frustrated, when in reality, nothing seems to have happened. For example, a depressed person may think that she is not truly loved, no matter how much her family wants to convince her of the opposite. While her desire was fulfilled in reality, she suffers because she does not believe it. Desire theories cannot count this as an impairment of welfare without violating the reality-orientation constraint and moving too close to mental-state accounts of welfare. In this example, I can accept the impairment of a basic psychological capability

but not this specific type of harm brought about by it. The desire theory may be able to account for some of her experienced harm, but not sufficiently.

Concerning suffering while no desire frustration is present, the desire theory is too narrow. It allows to downplay and disregard harm that I think should be considered real. The distinction between harm and suffering may be resolved by the desire not to suffer. But as established before, this lets the desire theory move very close to a mental-state theory. I think this flaw of the desire theory of welfare is greater than the objectivist charge. While the latter can be avoided for most instances using informed desires, the former is based on the main distinction between the mental-state and the desire theory. The only escape I can see is to embrace a mental-state conception of welfare in favour of the desire theory. This, however, comes with new problems as it moves further away from objective justification. The account would move into a very wide conception of harm that at the same time excludes harmful disorders that cause no direct suffering.

Conclusion

I found that the desire theory of welfare needs some supporting constraints to account for the harm caused by mental disorders. There are two main ways in which the desire theory of welfare is too narrow and fails to detect harm. First, it excludes harm that is caused when we cannot rationally judge whether a desire was fulfilled or not (charge from mental state). Second, it excludes harm caused when no desire is frustrated but we would still say that the person is worse off (charge from objective list).

If the impairments in basic psychological capabilities are not measured directly, then the desire theory can also be too wide. This is because using the extent of desire frustration to detect an impairment allows for diagnoses where no real disorders are present. The desire theory of welfare may also account for more harm than necessary if the intrinsic harm of desire frustration is too highly estimated. After all, frustrating distorted desires is necessary to protect patients from themselves. However, I think giving weight to desire frustrations whether or not the desire was informed is a great advantage of the theory. Assigning intrinsic harm to desire frustration expresses a respect for patients' autonomy that I think does justice to a modern understanding of mental disorders. I therefore hold that the desire theory of welfare provides a useful explanation of the harm caused by mental disorders that in some respects needs to be complemented by other theories.

Works cited

Arneson, Richard J. (1990): *Liberalism, distributive subjectivism, and equal opportunity for welfare*. *Philosophy & public affairs*, 158-194.

Badhwar, Neera K. (2014): *Well-being: Happiness in a worthwhile life*. Oxford University Press.

Brandt, Richard B. (1998): *A theory of the good and the right*. Amherst, N.Y.: Prometheus Books.

Crisp, Roger (2017): *Well-Being*. The Stanford Encyclopedia of Philosophy, edited by Edward N. Zalta, Fall 2017, Metaphysics Research Lab, Stanford University. *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/fall2017/entries/well-being/>.

Graham, George. (2013): *The disordered mind: An introduction to philosophy of mind and mental illness*. Routledge.

Griffin, James (1986): *Well-Being: Its Meaning, Measurement and Moral Importance*. Clarendon Press, Oxford

Hausman, Daniel M. (2011): *Preference, value, choice, and welfare*. Cambridge University Press.

Heathwood, Chris (2014): *Subjective theories of well-being*. in B. Eggleston and D. Miller (eds.), *The Cambridge Companion to Utilitarianism*. Cambridge University Press, pp. 199-219.

Nozick, Robert. (2012): *The experience machine*. *Ethical Theory: An Anthology*, 14, 264.

Parfit, Derek (1984): *Reasons and Persons*. Oxford: Clarendon.

Superson, Anita (2005): *Defomed Desires and Informed Desire Tests*. *Hypatia*, 20: 109-126. doi:10.1111/j.1527-2001.2005.tb00539.x

Tekin, Şerife (2011): *Self-concept through the diagnostic looking glass: Narratives and mental disorder*. *Philosophical Psychology*, 24(3), 357-380.