

Title	J. S. Mill's Proof of the Principle of Utility and Psychological Egoism
Author(s)	Ohishi, Toshihiro
Citation	Philosophia Osaka. 3 p65-p.74
Issue Date	2008-03
oaire:version	VoR
URL	https://doi.org/10.18910/11041
DOI	
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

Toshihiro OHISHI (Okinawa National College of Technology)

J. S. Mill's Proof of the Principle of Utility and Psychological Egoism

J. S. Mill presents the proof of the principle of utility and attempts to show the validity of the principle in his *Utilitarianism*. I think that psychological egoism plays an important role in this proof. The purpose of this paper is to examine the problem of whether this proof is valid or not by making this role clear.

There are two kinds of egoism, namely ethical egoism and psychological egoism. According to ethical egoism, whether we are really egoistic or not, we should be egoistic. Even though we are sometimes not egoistic, an ethical egoist recommends us to obey egoistic principles. On the other hand, psychological egoism is the view that as a fact our action is always motivated by our own self-interest and in essence we are all egoistic.¹ What is concerned in this paper is psychological egoism.

1. Mill's proof of the principle of utility

I should give an outline of Mill's proof of the principle of utility in Chapter 4 of his *Utilitarianism*. It is as follows:

- (1) The sole evidence it is possible to produce that anything is desirable is that people do actually desire it. (par. 3)
- (2) Each person desires his own happiness. (par. 3)
- (3) Each person's happiness is a good to that person. (par. 3)
(Therefore)
- (4) The general happiness is a good to the aggregate of all persons. (par. 3)
(Therefore)
- (5) Happiness has made out its title as one of the ends of conduct and, consequently, one of the criteria of morality. (par. 3)
(By the way)
- (6) Human nature is so constituted as to desire nothing which is not either a part of

¹ This is thought to be the description of facts about an individual. As I will argue in this paper, 'egoism' in the sense of the phrase 'psychological egoism' does not mean the character of being arrogant and selfish, but stands for a broader concept.

happiness or a means of happiness. (pars. 4-10)

(7) Happiness and means of happiness are the only things desirable. (par. 9)

(Therefore)

(8) Happiness is the sole end of human action, and the promotion of it the test by which to judge of all human conduct. Consequently happiness must be the criterion of morality. (par. 9)

The problem of the proposition (1) will hereinafter be placed at the central point of the argument in this paper. The relation between “desirable” and “desire” is referred to in the proposition (1).²

Mill argues about the proposition (1) as follows (I call the following quotation the quotation (a)).

The only proof capable of being given that an object is visible is that people actually see it. The only proof that a sound is audible is that people hear it; and so of the other sources of our experience. In like manner, I apprehend, the sole evidence it is possible to produce that anything is desirable is that people do actually desire it. (4.3)³

There are some objections against this explanation of the proposition (1). I would like to show the point of them next.

2. The criticisms of Mill's proof

When we read this explanation of the proposition (1), what we would first notice is that the analogy between ‘visible’ (‘audible’) and ‘desirable’ is not right. Certainly, since ‘visible’ means ‘can be seen’, to prove that some object is visible, we only have to see it actually. However, that does not apply to ‘desirable’ since ‘desirable’ does not mean ‘can be desired’ but ‘worth desiring’. Even though a person really desires an object, we cannot always say that the object is desirable. At first sight, Mill’s explanation makes the impression that it rests on a wrong analogy.

G. E. Moore points out that a wrong analogy is used in Mill’s proof of the principle of utility and suggests a more positive criticism. It is said that this criticism hit Mill’s proof

² As we can see from the propositions (1)-(3), Mill thinks that the words ‘desirable’ and ‘good’ are interchangeable.

³ The following way of quoting from and referring to Mill’s *Utilitarianism* is employed. I write a chapter number and a paragraph number in this order in the text of this paper instead of presenting a page number. For example, ‘4.3’ indicates ‘the paragraph 3 of Chapter 4’ in *Utilitarianism*.

severely. It still has a great influence on the argument about Mill's proof of the principle of utility.⁴

According to Moore, Mill commits 'the naturalistic fallacy'. Then, what is 'the naturalistic fallacy'? We can clearly comprehend what Moore means by 'the naturalistic fallacy' from his following description.

That fallacy [i.e. the naturalistic fallacy], I explained, consists in the contention that good *means* nothing but some simple or complex notion, that can be defined in terms of natural qualities. In Mill's case, good is thus supposed to *mean* simply what is desired; and what is desired is something which can thus be defined in natural terms.⁵

This description shows that 'the naturalistic fallacy' is one in a definition. That is to say, Moore regards Mill's proof as an attempt to define the word 'desirable' that applies to non-natural qualities as 'desired' that applies to natural qualities. He thinks this kind of definition is invalid.

However, according to another interpretation it is the fallacy to derive 'an evaluative (value) judgment' from 'a non-evaluative (factual) judgment' that Mill commits. Mill deduces the evaluative conclusion that an object is desirable from the non-evaluative (factual) premise that the object is desired.⁶

Moreover, there is another interpretation.⁷ It is the interpretation that Mill commits the fallacy to equate 'a natural property' with 'a non-natural property'. That is, Mill cannot tell 'a natural property' that is represented by the word 'desired' from 'a non-natural property' that is represented by the word 'desirable'.

It is not clear if Moore thinks of the latter two fallacies as 'the naturalistic fallacy'. However, what we are concerned here is whether Mill commits such fallacies or not. If these criticisms (I call them Moore's criticism) are true of Mill's argument, the argument will not be valid. Before I enter into this problem, I would like to take up the argument that attempts to show that Moore's criticism is implausible and Mill's argument is acceptable.

⁴ Cf. Susan Leigh Anderson, *On Mill* (Wadsworth, 2000), p. 54; Tooru Hirao, *Kourisei Genri [The Principle of Utility]* (Houritsu Bunka Sha, 1992), pp. 165-166, 172.

⁵ G. E. Moore, *Principia Ethica* (Cambridge, 1903), revised ed., p. 125. Cf. *ibid.*, pp. 111, 118-119. Hereinafter the words in square brackets mean a supplementary explanation I added.

⁶ Cf. Roger Crisp, *Mill on Utilitarianism* (Routledge, 1997), p. 74; Hirao, *op. cit.*, p. 171.

⁷ Cf. Crisp, *loc. cit.*

3. The vindication of Mill's proof

Mill writes, "If the end which the utilitarian doctrine proposes to itself were not, in theory and in practice, acknowledged to be an end, nothing could ever convince any person that it was so" (4.3, cf. 1.4). We can say that Mill is convinced that the general happiness is the end of moral. Moreover, whereas Mill writes, "It has already been remarked that questions of ultimate ends do not admit of proof, in the ordinary acceptation of the term. To be incapable of proof by reasoning is common to all first principles, to the first premises of our knowledge, as well as to those of our conduct" (4.1), he claims, "Considerations may be presented capable of determining the intellect either to give or withhold its assent to the doctrine [i.e. the utilitarian doctrine]; and this is equivalent to proof" (1.5, cf. 4.12). That is, the principle of utility is premised as the first principle in Mill's proof and in the ordinary sense of the term 'prove' we cannot prove this principle right. Mill attempts to present an argument to persuade the intellect to accept the principle of utility and calls the argument proof.⁸

In the orthodox vindication of Mill's proof, the above implication of the proof of the principle of utility is emphasized and the quotation (a) is explained on the basis of this implication. Naturally, there is not a unified interpretation as an orthodox view. Next, I would like to present the point of some orthodox interpretations of the quotation (a).⁹

Since Mill does not intend to prove the principle of utility right in the ordinary sense of the term 'prove', the proof of the quotation (a) is not a strict proof. In fact, in the quotation (a) Mill uses the word 'proof' regarding 'visible' and 'audible', whereas he uses the word 'evidence' regarding 'desirable'. Moreover, what Mill argues is not that what we can say of 'visible' and 'audible' applies to 'desirable' *in the same way*, but that what we can say of 'visible' and 'audible' applies to 'desirable' *in like manner*. That is, Mill does not argue that 'desired' entails 'desirable' as 'seen' entails 'visible'. The analogy on which Mill's argument depends in the quotation (a) is between 'our factual knowledge' and 'ultimate ends of our conduct' (cf. 4.1). In the case of 'our factual knowledge', we rely on the empirical faculty of senses such as sight and hearing to make a judgment on facts. Similarly, in the case of 'ultimate ends of our conduct', we rely on the empirical faculty of desire. As we can talk about 'visible' just because we have the experience of 'see', we can talk about 'desirable' just

⁸ Cf. J. Seth, "The Alleged Fallacies in Mill's 'Utilitarianism,'" *Philosophical Review* 17 (1908), 474-476; E. W. Hall, "The 'Proof' of Utility in Bentham and Mill," *Ethics* 60 (1949), 8-10; D. D. Raphael, "Fallacies in and about Mill's Utilitarianism," *Philosophy* 30 (1955), 346-349; N. Cooper, "Mill's 'Proof' of the Principle of Utility," *Mind* 78 (1969), 278-279; Takashi Koizumi, *Miru no Sekai [Mill's World]* (Koudan Sha, 1988), pp. 166-172; Crisp, op. cit., pp. 75-76; H. R. West, *An Introduction to Mill's Utilitarian Ethics* (Cambridge, 2004), pp. 125-126.

⁹ Cf. Seth, loc. cit.; Raphael, loc. cit.; Cooper, loc. cit.; Crisp, loc. cit.

because we have the experience of 'desire'. In other words, it is not a *sufficient* condition but a *necessary* condition of something's being desirable that it is desired.

If this kind of vindication is convincing, Moore's criticism referred in the section 2 of this paper is a misunderstanding of Mill and is not true of Mill's argument. In my understanding, however, there are two flaws in this vindication.

4. Against the vindication of Mill's proof

The first problem is that in the quotation (a) Mill clearly states that the sole evidence it is possible to produce that anything is desirable is that people do actually desire it. According to the orthodox vindication of Mill's proof, as above mentioned, this statement should not be literally interpreted and we should look at the statement in the right context. However, this statement is obviously equal to the claim that if you desire an object, the object is desirable. Hence, if Mill's real intention is such as the orthodox advocates of Mill's proof claim, Mill should not have made such a misleading statement. It is all the more so if Mill intends to persuade the intellect to accept the utilitarian principle. We cannot help saying that Mill was very careless to make such a statement.

We can see similar statements in other places. Mill writes as follows (I call the former the quotation (b) and the latter the quotation (c)):

If the opinion which I have now stated is psychologically true—if human nature is so constituted as to desire nothing which is not either a part of happiness or a means of happiness—we can have no other proof, and we require no other, that these are the only things desirable. (4.9)

I believe that these sources of evidence, impartially consulted, will declare that desiring a thing and finding it pleasant, aversion to it and thinking of it as painful, are phenomena entirely inseparable or, rather, two parts of the same phenomenon—in strictness of language, two different modes of naming the same psychological fact; that to think of an object as desirable (unless for the sake of its consequences) and to think of it as pleasant are one and the same thing; and that to desire anything except in proportion as the idea of it is pleasant is a physical and metaphysical impossibility. (4.10)

Mill presented the description of the quotation (b) after he gave concrete consideration to conclude that people desire only happiness. From this description, the thesis (8) of the proof referred in the section 1 of this paper, namely that happiness is the sole end of human action, and the promotion of it the test by which to judge of all human conduct, and consequently happiness must be the criterion of morality, is drawn. In this context, in the quotation (b) Mill

claims that if happiness alone is desired, then happiness alone is desirable.

There is a little difference between the quotations (a) and (b), that is to say, the conditional in the quotation (b) seems to be different from the invalid conditional that if happiness is desired, then happiness is desirable and would be acceptable to the orthodox advocates of Mill's proof.¹⁰ In this case, we should think it is presupposed in this valid conditional that although there are a lot of cases where happiness is not desirable, happiness alone is desirable.¹¹ However, if we understand this conditional to mean that if happiness alone is desired, then all happiness alone is desirable, this conditional is unacceptable to the orthodox advocates of Mill's proof. The reason for this is that in this case we derive 'being desirable' from 'being desired'.

In the quotation (c) Mill refers to two correspondences, namely the correspondence between 'desiring a thing' and 'finding it pleasant' and the correspondence between 'thinking of an object as desirable' and 'thinking of it as pleasant'. If these two correspondences are accepted, the correspondence between 'desiring a thing' and 'thinking of it as desirable' should also be accepted. That is, we should accept that if an object is desired, it is desirable. This conditional, however, is not acceptable to the orthodox advocates of Mill's proof.

If we stick fast to the orthodox vindication of Mill's proof, we should say that Mill attempts to persuade the intellect to accept the principle of utility although Mill is careless in his argumentation regarding the principle. However, as has been noted, Mill is *much too* careless in the argumentation. Hence, we cannot accept his argument about the principle. It is difficult for us just to understand that Mill intends to convince the intellect to accept his principle of utility by such a careless argument.

The second problem is that the orthodox vindication of Mill's proof ruins the proof of the principle of utility.

According to this vindication, in the quotation (a) Mill does not argue that if an object is desired, it is desirable, but that if an object is desirable, it is desired. To accept this argument is to accept the argument that although not all desired things are desirable, what is desirable is desired. It implies that some desired things are undesirable.

However, this interpretation will make the argument in the paragraph 3 of Chapter 4 inconsistent.

In the paragraph 3 of Chapter 4 Mill argues about the propositions (1)-(5) presented in the section 1 of this paper. If we adopt the vindication in question, we cannot help accepting the following argument. The proposition (2) (each person desires his own happiness) means

¹⁰ Cf. R. F. Atkinson, "J. S. Mill's 'Proof' of the Principle of Utility," *Philosophy* 32 (1957), 161-162.

¹¹ However, in this case, the same criticism as I will present in the following also applies to this interpretation of the quotation (b).

that my own happiness is desired. Since part of the desired things is desirable, the proposition (3) (each person's happiness is a good to that person) means that although not all my happiness which I desire is desirable (a good) to me, part of my happiness is desirable (a good) to me. Mill derives the proposition (4) from the proposition (3). Hence, we can interpret the proposition (4) in the same way as the proposition (3). That is, although not all the general happiness is desirable (a good) to the aggregate of all persons, part of the general happiness is desirable (a good) to the aggregate of all persons. Consequently, some general happiness is undesirable to the aggregate of all persons.¹²

However, I think that Mill cannot accept this consequence since in his proof of the principle of utility he considers all the general happiness desirable (a good) to the aggregate of all persons.

5. Two meanings of 'desirable' and psychological egoism

So far I have been arguing about Moore's criticism of Mill, the vindication of Mill and the defects of the vindication. This vindication appears to fail in defending Mill's proof. Then, does it show that Moore's criticism of Mill is valid? I think that Moore's criticism is valid in a sense, and is not in another sense. To give the reason for this I would like to focus attention on the meaning of 'desirable' ('good').

How do we use words such as 'desirable' and 'good'? When we take notice of the use of the words in our daily life, we can see an important distinction in the meaning. To take an example, when I say that if anyone is in trouble, it is desirable (good) to help her/him, I mean that we should help such a person as a member of society. In this case, I use the words such as 'desirable' and 'good' from a social point of view. In other words, it is the use of the words in a moral sense. On the other hand, for example, when I say that since I am inclined to fat these days, it is desirable (good) to do some exercise, I mean that I should do some exercise for myself. In this case, I use the words from my own personal point of view. It is the use of the words in a non-moral sense, that is to say, in an individual sense.

The word 'good' has some uses in a non-moral sense. To take another example of the use of 'good', we sometimes say that he is a good baseball player. The word 'good' in this sentence is used in a non-moral sense, but not in an individual sense. The word 'good' has some other meanings than an individual meaning. However, what we must consider here is the contrast between a moral meaning and an individual meaning.

¹² Is it possible to avoid this consequence by manipulating the transition from the proposition (3) to the proposition (4)? I do not think that it is possible, but I do not discuss this matter here.

Furthermore, some desirable (good) things in an individual sense are not desirable (good) in a moral sense. Let us suppose that you find a child drowning in a river. If you race into the water, you will be able to rescue the child from drowning. However, you are not very good at swimming. You believe that there is a strong probability that you will be drowned. You are afraid of drowning and do not want to die. Hence, you cannot possibly dive into the water to save the child's life. The child will die in the river while you are hesitating to save her/him. You will be blamed for not having dived into the river.

In contrast, some desirable (good) things in a moral sense are not desirable (good) in an individual sense. As an illustration, there is the case where somebody gets down from a station platform to rescue a person who fell down the platform and is run over by a train. She/he is killed in the railroad accident but does not really want to die. She/he will be praised for trying to save the person. However, this conduct is the cause of her/his death and is not desirable for her/him. We can say that she/he should not have got down from the platform.

Thus, there is a notable distinction between a moral meaning and an individual meaning regarding such words as 'desirable' and 'good'. It is a fact in the eyes of the use of these words in our daily life that we use these words, distinguishing these two meanings of these words. From this viewpoint I would like to examine Mill's proof of the principle of utility.

Mill argues about the relation between 'desired' and 'desirable' (the quotation (a)) and states that each person desires his own happiness (the proposition (2)) several lines later. Right after the description of the proposition (2) Mill states that happiness is a good, that is to say, each person's happiness is a good to that person, and the general happiness, therefore, is a good to the aggregate of all persons (the propositions (3) and (4)). In particular, the remark that each person's happiness is a good to that person is worthy of notice. The point of this remark is that each person's 'desire', 'desirability', 'good' and 'happiness' are on Mill's mind. That is, up to here Mill has argued from an individual point of view. On the basis of a sequence of arguments from an individual point of view Mill concludes that the general happiness is a good to the aggregate of all persons. This conclusion is about whole society. Mill makes the transition from an individual viewpoint to a social (moral) viewpoint at a blast. The statement of the propositions (3) and (4) is followed by the proposition (5). Next Mill starts a new paragraph and states that in fact human beings psychologically desire nothing except happiness (the proposition (6)), observing his own and others' psychology. This shows that an individual at issue is a concrete individual living in actual society.

In brief, in the quotation (a) Mill argues about the relation between 'desired' and 'desirable' on an individual level in actual society. What someone desires is desirable for her/him, and nothing is desirable for her/him except what she/he desires. On an individual

level we can say that if an object is desired, it is always desirable.¹³ In other words, I can see that Mill makes use of psychological egoism for showing the validity of the utilitarian principle. This applies to the descriptions of the quotations (b) and (c). At this stage, 'desirable' in a moral sense is not at issue.

6. Conclusion

In my understanding, Moore's criticism of Mill mentioned in the section 2 of this paper is not applicable to Mill's proof as it is. Mill's analogy is acceptable since he is talking of 'desired' and 'desirable' on an individual level in actual society. What is desired is desirable for each individual. Conversely, what is desirable for someone is what she/he desires. If an object is desired by someone, the object is desirable for her/him at the same time, and vice versa. Hence, what each individual desires is good, pleasant and happiness for her/him. Each person's happiness is a good for her/him and each of human beings desires her/his own happiness, that is to say self-interest (psychological egoism).¹⁴

Hence, I find it incorrect to consider that Mill intends to define the word 'desirable' in a moral sense. I think that Moore assumes that Mill attempts to define the word 'desirable' in a moral sense in terms of natural qualities. However, it is not 'desirable' in a moral sense, but 'desirable' in a non-moral (individual) sense that Mill deals with in the argument at issue. In addition, when Mill is criticized for inferring 'an evaluative (value) judgment' from 'a non-evaluative (factual) judgment', it would be supposed that Mill infers 'a judgment on a moral value' from 'a judgment on an individual fact'. In the case of the claim that Mill equates 'a natural property' with 'a non-natural property', the point would be that Mill fails to distinguish 'a property of an individual' from 'a moral property'. These criticisms do not hit the mark. The critical mistake of them is that they overlook the distinction between a moral meaning and a non-moral (individual) meaning in the use of the words such as 'desirable' and 'good'. The orthodox advocates of Mill's proof make the same mistake and thus it is natural

¹³ Hirao supports the view that what is desired is desirable by the concept '*zettaiteki tandokusha* (the absolute single person)'. He sets '*zettaiteki tandokusha*' as an epistemological hypothesis. '*Zettaiteki tandokusha*' is a nonsocial existence and does not have any connections with other people. Hirao suggests that we should understand our social conditions by means of the concept '*zettaiteki tandokusha*' (cf. Hirao, op. cit., pp. 175-179). However, I think that his interpretation is not very convincing. In my understanding, as I have said, 'an individual' in Mill's proof is a flesh and blood individual living in actual society. I interpret Mill's proof, focusing attention on each of human beings in actual society.

¹⁴ We should go on to a more detailed examination of the concepts 'happiness' and 'self-interest' to avoid misunderstandings about the concepts, but I leave the matter open for the moment.

that their arguments get into incoherence.¹⁵

Mill attempts to persuade us to accept the proposition (4)—the general happiness is a good to the aggregate of all persons—on the basis of the observation of the everyday use of ‘desirable’ (‘good’) on an individual level. The important question we have to ask here is whether we can make the transition from an individual (non-moral) level to a social (moral) level. In this case, it is not the transition from ‘desired’ on a individual (non-moral) level to ‘desirable’ on a social (moral) level, but the transition from ‘desired’ and ‘desirable’ on an individual (non-moral) level to ‘desirable’ on a social (moral) level that we should deal with. This is not the very same problem that Moore is concerned with. However, if we can regard the argument about ‘desired’ and ‘desirable’ on an individual (non-moral) level as the argument about a fact of an individual, we can claim that Moore’s criticism is not dead. I do not think that this transition from ‘desired’ and ‘desirable’ on an individual (non-moral) level to ‘desirable’ on a social (moral) level is possible (cf. the section 5 of this paper), but we need closer examination to solve this matter entirely. I leave the matter a future subject.

©2008 by Toshihiro OHISHI. All rights reserved.

¹⁵ Whether Mill clearly grasped this distinction or not is a controversial issue. Here I would only like to point out that if he was definitely conscious of this distinction, Mill should have argued in another way.