



Research Repository UCD

Title	Mills Captivating Proof and the Foundations of Ethics
Authors(s)	Baker, John
Publication date	1980
Publication information	Baker, John. "Mills Captivating Proof and the Foundations of Ethics" 6 (1980).
Publisher	Florida State University, Department of Philosophy
Item record/more information	http://hdl.handle.net/10197/4272

Downloaded 2024-03-28T04:02:09Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

Mill's Captivating 'Proof' and the Foundations of Ethics

by John M. Baker

Department of Ethics and Politics

University College

Belfield

Dublin 4.

Mill's Captivating 'Proof' and the Foundations of Ethics

Utilitarianism, like other philosophical theories, can captivate the intellect. To free ourselves from such a theory it is necessary to root out some of its basic commitments, and it happens that Mill's famous 'proof' of utilitarianism can help us. The central theme of this paper is that the mistakes in Mill's argument are not, on the whole, logical or formal, but consist in the tacit endorsement of widely attractive but disputable normative claims. If we accept these claims, we emerge with a moral outlook akin to Mill's: but having recognised them, we may not find such an outlook so appealing.

Since Mill's argument is an example of the attempt to provide a 'foundation' for ethics, the paper concludes with some remarks on foundationalism in general, and how to lessen its charm.

1. Utility, empiricism, and liberalism

Mill claims to belong to the 'inductive school' of ethics, and hopes in chapter four of Utilitarianism to place the first premises 'of our conduct' on the same kind of basis as 'the first premises of our knowledge.'¹ In the first step of his proof, he employs his much-abused analogy between 'visible' and 'seen', on the one hand, and 'desirable' and 'desired', on the other. Let's look at the analogy more closely.

An empiricist holds that the relation between a sensation and an object is contingent and inductive. For instance, he holds that someone might have a visual sensation with or without there being some object causing it, but that the sensation is evidence that some object is causing it. Moreover, the only way to show that an object is of the kind which causes such sensations (i. e., is visible) is to show that someone has such sensations--it remains possible, but not provable, that something is visible which no one will ever see. Finally, an empiricist might well allow that some things are not visible to everyone (e. g. objects of certain colours), and of course that many things are visible.

All this, we might say, is a grossly distorted picture of our language of perception. We might claim that the relation between sensation and visibility cannot be contingent and inductive, but belongs to the 'grammar' of our language.² The point, however, is that this picture does belong to the empiricist tradition. So is it, after all, surprising that Mill puts forward a similar picture of the relation between desire and desirability? Here, as before, the suggestion is that the psychological datum is related to the real object in a contingent, inductive way. It is possible that something is desirable although

no one desires it, and that someone has a desire for something not actually desirable. But the only way to show that something is desirable is to show that someone has a desire for it. Since, by analogy, it might be possible that different things are desirable to different people, or that many things are desirable, Mill takes care to deny these possibilities.

But if the empiricist claim about perception is merely 'tracing round the frame' of our language about vision, what is Mill tracing in his analogy? Let us start by noting that the claim is not nearly so implausible as it looks. Suppose I'm trying to prove that something (say, reading Mill) is desirable for you - or, in everyday English, good for you. How can I go about doing this? Well, I might claim that it would help you appreciate great minds, or that it would increase your vocabulary. If you ask why these are good for you, I might reply that they will help you to think more clearly yourself, and that this may, for example, help you resolve certain problems. What the conversation has done is to trace my claim that something is good for you to a claim about something you want. This connection, though misrepresented by Mill, does indeed seem to be part of the justificatory structure of our moral language.

There is a difference with the perceptual case, though, owing to the contestable character of moral justification. Someone who refused to allow that seeing something counts towards its being visible would be making either a linguistic or a philosophical error. But someone who refused to allow that wanting something counts towards its being desirable would be, simply, illiberal. He would be holding the view that people aren't the best judges of their own good, and hence that their desires might well conflict with what's good for them. Mill did not hold, as a crude liberal would, that people are never mistaken in this regard (though he thought it was wrong to interfere even if they were), but he maintains this central liberal tenet in the desired-desirable claim. The upshot of all this is that to take issue with Mill over the first step of the proof requires not merely an obvious perception about the ambiguity of suffixes, but at least a partial rejection of liberalism itself.

2. The general happiness and the problem of social choice

The next step in Mill's proof is from 'each person's happiness is a good to that person' to 'the general happiness, therefore, a good to the aggregate of all persons.' In spite of the scorn of most commentators, there is a perfectly consistent interpretation of this step,

the shortcomings of which seem much more illuminating than traditional objections.

Let us start by considering the following, perfectly valid argument:

The sky above each person's house is visible to that person.

Therefore, the whole sky above Dublin is visible to the population of Dublin.

What makes the argument valid is that 'the whole of x is visible to A, B, & C' means something like 'every part of x is visible to A or B or C, and some part of x is visible to each', and that the whole sky above Dublin consists of the sky above the houses of Dublin. So there are two conditions for the success of Mill's argument. First, we must be able to take 'the whole of x is desirable to the aggregate' to mean that 'every part of x is desirable to someone, and some part of x is desirable to each.' This is quite plausible, though hardly self-evident. Secondly, we must be able to take the general happiness to consist of the happiness of each person. Mill himself confirms this view by referring to it in a now-famous letter as 'the sum' of A's happiness, B's happiness, etc.³ So interpreted, the conclusion follows.

But let us consider the second condition more closely. What is that thing which consists of the happiness of each person? It is the state of affairs in which everyone is happy. No doubt such a condition of universal bliss is the utilitarian ideal. But suppose that this ideal is unattainable - the utilitarian tells us to pursue the greatest happiness as the next best thing. What Mill fails to show is that of a set of distributions of happiness falling short of universal bliss, the one with the greatest happiness is better than any other. Nor can his proof be altered in order to do so, as in the following attempt:

Each person's greatest possible happiness is the best thing for that person.

Therefore, the greatest possible happiness of the whole is the best thing for the aggregate.

The conclusion doesn't follow precisely because neither of the two conditions necessary for its validity is acceptable. We cannot take 'x is the best thing for the aggregate' to mean 'every part of x is the best thing for someone, and some part of x is the best for each'; nor can we take it that the greatest possible happiness of the whole consists of the greatest possible happiness of each. Both conditions are too restrictive to be practicable, when the very point of the alternative argument was to deal with actual possibilities; and the second condition is certainly at odds with the utilitarian tradition. Mill has shown the desirability of the general happiness: but not of the greatest happiness.

Utilitarianism is often accused of failing to give an adequate account of distributive justice - Mill devoted part of chapter five of his essay to rebutting the charge. We have now seen that the failure is already present in Mill's proof, in his failure to construct what's best for society from what's best for the individual. But as Barry has pointed out, this very project - 'the problem of social choice' - is characteristically 'liberal'.⁴ Moreover, in the modern classics of this field, such as Arrow's Social Choice and Individual Values, Mill's two steps are conflated. The liberal tenet that individual preferences determine individual good is taken so much for granted that the whole problem of justification appears simply as the move from preferences on the one hand to social welfare function on the other. Were we to hold that some elements of the social good are not good for anyone in particular, or that a prior conception of social good is necessary for establishing the good of individuals, then we should be denying the basis of one or both of Mill's first two steps, with the consequence that even his argument for the general happiness would be undermined. But we should also be claiming that the liberal project is itself mistaken, and so striking at one of the roots of Mill's tree, not just at its branches.

3. Consequentialism and hedonism

Utilitarianism is often criticised for its consequentialism, i.e. the view that the rightness of actions is based on the goodness of their consequences. Mill takes this doctrine for granted in his proof - he doesn't bother to argue that the goodness of happiness makes it 'one of the ends of conduct, and consequently one of the criteria of morality.' Rawls comments:

The chapter title refers to the proof of the principle of utility; but what we are given is an argument to the effect that happiness is the sole good. Now nothing so far follows about the conception of right.⁵

Rawls is concerned with the 'teleological' character of utilitarianism, and this, on his account, includes both its consequentialist and maximising aspects. But when Mill equates the ends of conduct with the criteria of morality, there is as yet no commitment to a purely aggregative theory. I do not intend to enter the debate about consequentialism here, but only to emphasise the appeal of consequentialism in this broad sense, freed from aggregative notions - for it is nothing more than an overriding concern for promoting human welfare.⁶ In contrast to the intuitionist, the consequentialist refuses

to accept any obligation which makes no difference to this welfare, or does nothing but decrease it. That such obligations might be given some other justification, in terms of the authority of tradition, or moral sense, seems to him simply reactionary mystification. Such consequentialism strikes many people as fundamental, or even self-evident. Yet there are well-known alternatives - views which hold that some kinds of behaviour are required, or intolerable, whatever their consequences. And so here again Mill has relied on an appealing but disputable moral belief in order to prove his point.

The final element of Mill's proof is the claim that nothing is desired but happiness. It is unnecessary to review here the way in which this seductive view has been widely investigated and convincingly undermined. But it is worth noting that in response to these criticisms, utilitarians are apt to choose one of two options. The first, to retain the traditional doctrine but deny its dependence on psychological hedonism, could not have appealed to Mill: if people want things besides happiness, it would be illiberal to deny these things any moral weight. No doubt Mill was sufficiently concerned about ignoble desires that he wrote into the very meaning of utilitarianism the notorious doctrine of the quality of pleasure; but the test of quality remained, in typically liberal fashion, what people actually preferred (or in conflict majority vote!). So it is not just that, in the absence of psychological hedonism, utilitarianism remains ungrounded, but that the denial of the psychological doctrine provides a liberal with a good reason for rejecting utilitarianism altogether. Or for taking the second option - for revamping the Greatest Happiness Principle as the Greatest Want-Satisfaction Principle. This certainly involves problems akin to those which beset the classical doctrine (inter-personal comparison, cardinality, distribution, etc.), as well as the difficulty of clarifying the very concept of want-satisfaction. But it is more important in the present context to observe how much the revamped principle owes to the central elements of Mill's proof. It remains consequentialist. It endorses not just the liberal project of constructing social from individual good, but Mill's ill-considered way of doing it. And it is even more firmly in the liberal tradition of the argument from desire to desirability than Mill himself.

4. Foundationalism and its alternatives

We may characterise generally any attempt to give an ultimate justification of morality--to base morality on incontrovertible grounds--as 'foundationalist'. Mill's utilitarianism is foundationalist, but so is classical intuitionism; whereas what Rawls calls intuitionism is anti-foundationalist. This paper is intended to contribute to the criticism of foundationalism in general, by treating a particular case of it. Mill's proof seems a good example for three reasons. First, because it is usually criticised using basic and incontrovertible criteria like the ambiguity of suffixes and the citation of recognised fallacies, while leaving the foundationalist project intact. Secondly, because the real source of its plausibility can be seen to lie in its implicit value commitments. And thirdly, because it seems genuinely to capture much of the appeal of utilitarianism.

Of course, the criticism of one form of foundationalism cannot show that foundationalism as such is mistaken, although it can suggest that this is the case. At the very least, similar criticisms of other forms are also needed. But besides showing how in each case foundationalism runs aground, a different strategy is necessary for showing, not that foundations are impossible to obtain, but that it is possible to live without them. It is not enough for this purpose to rest content with a view like Barry's: that we just have the values we have, and that justification comes to an end when we reach these basic values. First of all, anyone's value system is much more complicated than Barry's indifference curves suggest. But more importantly, in the absence of an overview of the origin and use of our evaluations, the Barry model seems like second best, to which we are forced by the lack of an adequate foundation. As Rawls remarks, it remains a challenge to provide the sort of criteria it claims impossible.⁷

So what the second strategy requires is an overview of the way moral and political language is used--of what we might call, following Wittgenstein, the language game of moral and political argument, or moral and political 'language play'. Such an overview is necessary to 'give philosophy peace'.⁸ Its boundaries must be vague, as there is no sharp boundary between what counts as moral argument and what is, say, merely a matter of prudence or etiquette--not to mention the fact that where to draw the line is itself morally controversial. We have to rely initially on our own poorly understood ability to discriminate between moral and other kinds of judgment; an overview would throw some light on this ability. Only a few salient features of such an overview are relevant to our present concerns.

The first thing to note is that this language game, like any other, is activity, not crystalline theory--it is dynamic, not static. Utilitarianism and other foundational approaches, as well as Barry's and Rawls's theories, construct one's moral and political

outlook as an edifice which, completed, furnishes an answer to every moral question. By contrast, when we look at the use of moral language we see that it is quite open-ended. It involves past applications so far as these are remembered and still cited as examples, and rules governing common cases and types of behaviour, and exceptions to these rules. But there is no reason to suppose that what we now feel about some wildly removed science fiction case would withstand the gradual development of the sort of world in which it is situated. Another aspect of the dynamic nature of the moral language game is that it is applied to problems which impose their own deadlines. Although moral language can be employed in seemingly endless inquests over whether certain past actions were really justified, even these take their character from the fact that moral problems can't be agonised over indefinitely, that they arise in the context of being forced by circumstances to act, to choose. This feature is reflected in moral reasoning typically being a matter of adjustment rather than deduction. (Rawls's 'reflective equilibrium' is an idealisation of this.) A case arises - how do I feel about it? What sorts of reason support this feeling? But what counts against it? Are these counter-reasons weak enough to be overridden, or do they call for an adjustment in my initial reaction? And so on. Moral argument reaches to various depths, according to who the arguers are and what they have in common. Two people (or, as Rawls says, a person in two minds) may be in complete agreement in their reaction to some issue, in which case the grounds for their agreeing beliefs may not even come into question; in the case of disagreement, the dispute works back to common ground, if there is any. Why should we suppose that once common ground is reached, there is a whole continent of common ground behind that? The aim of a moral discussion is to come to a judgment mutually agreed, not that this always happens; so all that is required is that starting from some common views, agreement can be secured. 'Justification comes to an end' - but not at any particular place; it ends where, in each case, it is forced back to ending.

This picture of the structure of moral argument is reinforced if we ask: how did any of us come to learn his values? A crucial part of the answer is that we were taught them piecemeal, and on the authority of our teachers. These teachers were themselves taught that way, though they changed their values under the impact of experience, just as we have done. We did not sit up in the pram and demand a detailed account of the bases of the values we were taught; luckily enough, because our teachers could not have provided one, any more than we can. Once having learnt them, or some of them, we came in life

to discover various conflicts and contradictions in them--or, more accurately, they came in our application of them to conflict and to contradict each other in ways unapparent to us previously, or to our teachers. We learn to resolve these conflicts as best we can, relying on the inner logic of the conflicts (the relations we perceive between the conflicting values themselves) and on feeling. Sometimes these conflicts require major surgery--as when one finds that one's whole outlook is permeated by sexist attitudes. In such times of upheaval, one is particularly prone to the attraction of foundationalism. It might be objected that to observe that people don't know the grounds for their moral outlooks, and were never taught them, is not enough to show that such grounds are impossible to supply. But the objection misses the point, which is not to show that foundations are impossible, but only to attack one reason for considering them necessary.

This account of the use of moral language--of the way we play that language game--does not itself have direct normative consequences. Of itself, it is neither for nor against liberalism and consequentialism. But it does make room for an outlook that would differ radically from Mill's. First of all, it would allow that we learn that some things are good for us even though we don't (or don't initially) want them--that our knowledge that they are good for us precedes our developing a taste for them. Think, for example, of the ancient problem which dominates a third of Rawls's book: is a sense of justice an element of individual good? Rawls tries to argue that, in general, it is, since there are generally good reasons for individuals to choose to have this sense. I would suggest, by contrast, that we are taught that it is in our own interests to be just, and that by the time the question of justifying this comes up, we are already hooked. Secondly, it would allow that some of the things which are good for a society can't be justified entirely in terms of the resulting benefits for its members, or that, sometimes, one of the reasons for saying that something is good for an individual is that it is good for society. For instance, how often can national sovereignty be justified in terms of its benefits for the people? Isn't it often just the opposite, that what the individual gains is simply to live in a free country? Finally, the present picture of the moral language game would allow that some actions are right for reasons other than the promotion of human welfare. Purported examples abound in anti-utilitarian literature; perhaps the clearest is genuinely retributive punishment.

A beguiled utilitarian will oppose these possibilities as reactionary; but to do so is to reject values many people would be reluctant to abandon. This brings us to

what is really at stake: namely, that when it comes to the detailed arguing out of these issues, no point of view occupies the privileged position foundationalists seek.

Utilitarianism appeals to some features of one's moral outlook and of the language game to which it belongs, involving particular commitments to which there are coherent alternatives; and recognising this may free one from its grasp. As for foundationalism in general, there is undoubtedly much more to be done. First, there is the detailed criticism of foundationalist attempts other than Mill's. Secondly, it is necessary to pursue the sources from which foundationalism as such derives its appeal. The most important of these is the fear of philosophical relativism. What needs to be shown is that just as the mere existence of moral argument is possible without foundations, so are such 'objectivist' activities as the criticism of the views of other cultures, and the search for moral truth. In the present context, I can only suggest that what needs uprooting is not the possibility of saying that something is really right, but a particular picture of what that means--what Williams calls the 'mid-air position'—, and that the role of moral truth in our lives is better appreciated by looking at the way we actually reflect, argue, develop, reassess, and so on, than by employing a picture of the right values being stored away somewhere, waiting to be discovered.⁹

5. Conclusion

If Mill's proof is unacceptable, it is not because of simple logical errors, but because of its deep commitment to certain ways of thinking. These ways of thinking are not easy to give up, and the most important of them involve not philosophical mistakes but the central tenets of liberalism. The proof is thus no mere appendage to the rest of Mill's ethics. But it is also no foundation for it; it rests on attractive but disputable ~~normative~~ claims. Have these themselves got some foundation? Mill himself hardly provides them, since he claims in On Liberty that his utilitarianism supports his liberalism. I should like to suggest in closing that in this respect Mill is no worse off than foundationalists of other persuasions, or in other areas of philosophy. But it is perhaps even clearer in ethics than elsewhere that Wittgenstein's remark applies: 'my reasons will soon give out. And then I shall act, without reasons.'¹⁰

Notes

1. J.S. Mill, Utilitarianism, chs. 1 & 4. All subsequent references are to chapter 4 unless otherwise indicated.
2. Cf. P.M.S. Hacker, Insight and Illusion (Oxford: Oxford University Press, 1972), chs. 9-10.
3. H.S.K. Elliott, ed., The Letters of John Stuart Mill (London: Longmans, Green & Co., 1910), II, 116.
4. B. Barry, The Liberal Theory of Justice (Oxford: Oxford University Press, 1973), ch. 11.
5. J. Rawls, A Theory of Justice (Oxford: Oxford University Press, 1972), 562n.
6. Or perhaps something more than that, such as the welfare of 'all sentient beings'. Cf. J.S. Mill, A System of Logic, VI, xii, 7.
7. B. Barry, Political Argument (London: Routledge & Kegan Paul, 1965), 4-8; Rawls, op. cit., sec. 7.
8. L. Wittgenstein, Philosophical Investigations (Oxford: Blackwell, 1963), sec. 133. The translation of 'Sprachspiel' as 'language game', though now well-established, is for many reasons misleading. In English we have no word which covers both 'game' and the activity of a game, i.e. 'play'; but for understanding 'Sprachspiel' we need to keep this multiplicity in mind, and consider whether it might not sometimes make more sense translated as 'language play' (e.g. secs. 7, 138, 669, and p. 179). However, I have reluctantly conformed to established practice in the text.
9. B. Williams, Morality (Harmondsworth: Penguin, 1973), 42ff.
10. op. cit., sec. 211. For their comments on earlier drafts of this paper, I'd like to thank Peter Hacker, Attracta Dunlop, and members of the Department of Philosophy at Trinity College Dublin and of the Irish Philosophical Society.