

Utilitarianism or Prioritarianism?

TORBJÖRN TÄNNSJÖ

Stockholm University

A simple hedonistic theory allowing for interpersonal comparisons of happiness is taken for granted in this article. The hedonistic theory is used to compare utilitarianism, urging us to maximize the sum total of happiness, with prioritarianism, urging us to maximize a sum total of weighed happiness. It is argued with reference to a few thought experiments that utilitarianism is, intuitively speaking, more plausible than prioritarianism. The problem with prioritarianism surfaces when prudence and morality come apart.

INTRODUCTION

Even thinkers who are, generally speaking, of a utilitarian bent, sometimes feel that utilitarianism faces problems when it comes to distributive matters. The most serious challenge to utilitarianism, I would say, comes from prioritarianism. However, there are other ideas confronting the utilitarian. Some hold, for example, with reference to Rawls and Nagel, that the separateness of persons means that loss in happiness in one individual cannot be compensated for by gains in other individuals. We need to opt for a lexical solution, and if there is room for improvement, we ought to improve the situation of the individual who is worst off, regardless of cost in the sum total of happiness. Others accept that trade-offs between lives can be made, but argue with Larry Temkin that it is bad when a person is worse off than another person. They defend some view to the effect that we need to seek an optimal distribution, catering both for the sum total of happiness and for its distribution. Roughly, the more happiness, the better, and the more equally it is distributed, the better. In the present context I will not discuss these views. This must be a matter for another occasion. This is so both because I find these views to pose less of a threat to the utilitarian position than does prioritarianism, but also simply for reasons to do with space. Here the focus is, then, on prioritarianism. I want to answer the question: which view caters best for our considered moral intuitions – utilitarianism or prioritarianism? With reference to three thought experiments, I reach the conclusion that, even if prioritarianism seems to get the upper hand in one of these

© Cambridge University Press 2015. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Utilitas Vol. 27, No. 2, June 2015

doi:[10.1017/S0953820815000011](https://doi.org/10.1017/S0953820815000011)

(that's why I am drawn towards it), the remaining two experiments show that it must go (and that I must debunk my intuition about the first one). Prioritarianism implies that it is better not to live a life that, from a prudential point of view, is worth living. This cannot be right.

In order to be able to conduct my investigation I need some common ground, a common currency, when I speak of a sum total of something to be maximized, weighed or not. I will simplify and speak of happiness. Here a few words about the common grounds.

COMMON GROUNDS

The hedonistic assumption is controversial, of course. I make it because I think it is reasonable.¹ But those who disagree, and feel that other things, besides happiness, matter in a life, are free to try to generalize my arguments to fit *their* favoured view. In many ways, I think it is possible to do so. The differences between utilitarianism and prioritarianism do not hinge on the assumptions made by each one of the theories about what it is we should ultimately be concerned with. But, of course, if more things than happiness matter, then some of the methodological assumptions made by the theories under scrutiny may turn out to be problematic. Even if we can make interpersonal comparisons of happiness, at least in theory, it is not so clear that we can make interpersonal comparisons of well-being, if well-being includes all sorts of items such as knowledge, friendship and virtue. It might even turn out to be difficult to construct ordinal, let alone cardinal, measures applicable in intrapersonal cases. I leave these problems to those who disagree with hedonism, however. Here I proceed with my discussion under the *assumption* that hedonism is on the right track.

Now, if we want to keep utilitarianism and prioritarianism apart, regardless of how else we conceive of well-being, we had better see to it that well-being is an empirical (non-normative) notion. Otherwise, it is likely that, if there are prioritarian weights in the first place, then we have already factored them in, when we assess how well off an individual is at a certain moment.² Moreover, we had better not calculate happiness (utility) in the standard manner we know from decision theory (in the style of von Neumann and Morgenstern) since, then, once again, if there are prioritarian weights, we will come to

¹ I have defended it in many places, for example in Torbjörn Tännsjö, *Hedonistic Utilitarianism* (Edinburgh, 1998).

² See Henry Sidgwick, *Methods of Ethics* (New York, 1966), p. 127, and J. J. C. Smart, 'Utilitarianism', *Utilitarianism: For and Against*, ed. J. J. C. Smart and Bernard Williams (Cambridge, 1973), pp. 3–76, at 22, for views where happy states are conceived of as *desirable* states. In Sidgwick I think this is just a lapse, and indeed one we had better avoid in the present context.

factor them in when we assess the utilities, and the distinction between utilitarianism and prioritarianism collapses.

Both utilitarianism and prioritarianism come with solid rationales. This is why both views, on the face of it, may seem plausible. Since they contradict one another, they cannot both be true, however. Here the problem is to see, not which one of them is true, but which one of them gains the best intuitive support in the thought experiments I have constructed (leaving the possibility open that they are both false while some third theory is the correct one).

THE RATIONALE BEHIND UTILITARIANISM

The rationale behind utilitarianism is the thought that happiness and happiness alone is what matters. And the importance of one additional 'hedon' (or however we measure happiness) is the same, wherever it obtains in the universe. Hence, we ought to maximize the sum total of them; we ought to maximize the sum total of happiness in the universe. Not to do so would be to waste what is of sole importance, happiness. The rationale behind (hedonistic) utilitarianism is as simple as that – and as closely connected to the doctrine as that. The rationale and the doctrine are really just two ways of making the same claim.

It has sometimes been argued that utilitarianism presupposes the existence of someone who experiences the sum total of happiness, and since no one does, utilitarianism is false. This is a mistake. The idea of someone experiencing the sum total of happiness is not part of the rationale behind utilitarianism. The utilitarian concern is a concern for the sum total *of* experienced happiness; the idea is not to waste happiness. No assumption is made, nor need be made, about any real or merely hypothetical individual experiencing the sum. It is of note that a similar argument, if it had been valid when directed against utilitarianism, could have been directed at egoism just as well. No one experiences (at any time) the sum total of happiness contained in a life.

THE RATIONALE BEHIND PRIORITARIANISM

It is more complicated to spell out the rationale behind prioritarianism. First of all, prioritarianism comes in two versions. The idea is that to give more moral weight to an increment in happiness given to a person at a time, the more unhappy the person is. It is not clear what it means for one person to be less happy than another, however. Do we here refer to the individual who, at the moment when the increment in happiness is given to her, is less happy? Or, are we referring to the individual who, overall, leads the least happy life? It is my firm belief that the best version of prioritarianism is presented by the idea that the crucial thing is how an individual fares at the *moment* when the increment in

happiness takes place. I will proceed under that assumption and I will give my reasons for it in the sequel.

A way of stating the priority view as a view where happiness at moments matters to the assignments of weights would be as follows. The better a person fares at a moment, the less weight should be given to further benefits to this person at this moment. And in particular, the more a person *suffers* at a moment, the more important is it to benefit him (to lessen his suffering) at this moment.

Of course, even on this understanding of the view, we sometimes have to cater for the needs of the happier person, rather than the miserable one, since this intervention brings with it a larger weighed sum of happiness *on the whole*. When we decide to whom an intervention should be directed, we must take all the future consequences into account.

My (moment-focused) interpretation of the priority view has been inspired by Dennis McKerlie, who is one of the few authors to have explicitly addressed the temporal matter. He too thinks that the application to moments is 'easier to defend' in relation to the priority view than to egalitarianism; this is, in my view, an understatement.³ I will say nothing here about egalitarianism, however.

The rationale behind this moment-focused understanding of prioritarianism is as follows: while happiness always matters, unhappiness not only matters, but it is of marginally increasing moral importance. And the core of this rationale is the asymmetry, i.e. the idea that suffering has a special weight. In a way, it accounts for Popper's claim to the effect that:

[T]here is, from the ethical point of view, no symmetry between suffering and happiness, or between pain and pleasure . . . In my opinion human suffering makes a direct moral appeal, namely, the appeal for help, while there is no similar call to increase the happiness of a man who is doing well anyway.⁴

This idea has sometimes been dismissed because of its alleged anti-nativist and even murderous implications. If Popper is right, would it not be better to put an end to all sentient life, since there is bound to exist some suffering among sentient beings in the future? This is of course the famous argument of R. N. Smart against what he calls 'negative' utilitarianism:

Suppose that a ruler controls a weapon capable of instantly and painlessly destroying the human race. Now it is empirically certain that there would be some suffering before all those alive on any proposed destruction day were

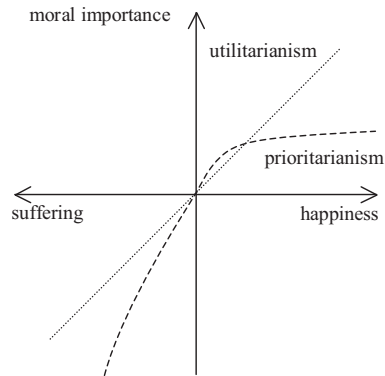
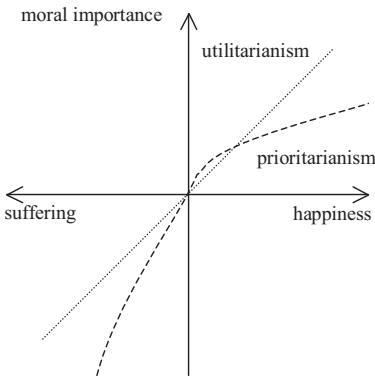
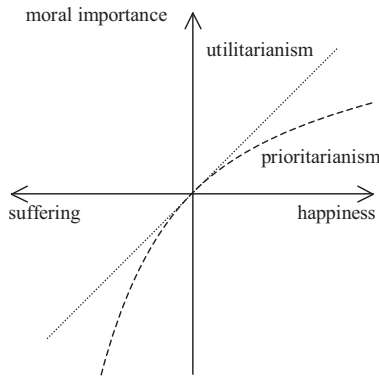
³ Dennis McKerlie, 'Dimensions of Equality', *Utilitas* 13 (2001), pp. 263–88, at 263 (abstract).

⁴ Karl Popper, *The Open Society and its Enemies, Volume 1: The Spell of Plato* (London, 2002), pp. 284–5.

to die in the natural course of events. Consequently the use of the weapon is bound to diminish suffering, and would be the ruler's duty on NU [Negative Utilitarian] grounds.⁵

However, in the passage from Popper quoted above it is only said that suffering has a 'direct' moral appeal, while the call to increase the happiness of a man who is doing well anyway makes no 'similar' call. This is consistent with both things being of importance, only that the one is more important than the other. I think the priority view nicely captures this intuition. We should avoid a lexical understanding of the view, then.

The difference between the priority view and utilitarianism can be illustrated by the following graphs:



The straight lines illustrate the utilitarian view that each equally large amount of happiness is equally morally important. The (concave) curves, on the other hand, illustrate the idea that the lessening of the

⁵ R. N. Smart, 'Negative Utilitarianism', *Mind*, New Series, 67 (1958), pp. 542–3, at 542.

unhappiness (the suffering) of those who are worse off counts for more than the lessening of the unhappiness (the suffering) of those who are better off. The exact shape of the concave function reflects the strength of the priority view.

Obviously, what we meet with here is really a family of very different views with one thing in common: the idea that unhappiness has an increasing marginal moral importance. To this is often added the view that it is the other way round for (positive) happiness (as shown in the curves above). I will be exclusively concerned with the unhappiness aspect of the theory in this context, however. One could think of prioritarian theories following the linear (utilitarian) function for (positive) happiness. One could also think of sufficientarian versions, where after having crossed the utilitarian straight line the curve has an upper limit, meaning that additional (positive) happiness, once you are happy enough, has no moral importance any more. One can also think of concave curves that start out above the utilitarian straight line for (positive) happiness, and cross the straight line, but without any upper limit. Some of these differences have been illustrated in the graphs above. I am eager to defeat *all* these version of the prioritarian view.

UTILITARIANISM OR PRIORITARIANISM?

Our choice then, is between utilitarianism and prioritarianism, applied to moments in our lives. In one crucial respect many may find prioritarianism superior to utilitarianism. Prioritarianism does indeed take suffering seriously. So does utilitarianism, but many may find that it does not take suffering seriously *enough*. I must confess that, when I contemplate, say, a case with two patients, one suffering terribly, and the other in mild pain, I tend to hold the intuition that, all other things being equal, we should care for the one who suffers terribly; this is so even if this would mean *some* waste of resources (happiness). A problem with this intuition, however, is that it may trade on the fact that it is put forward without precision. As soon as we have to specify how much waste in terms of happiness we should tolerate, in the interest of priority, it tends to go away (I think) and, anyway, even if it stays, it should not be accepted as evidence in its own right. Even if it stays, there must be something wrong with it, since there is a very strong argument against prioritarianism, which should stop us from having any hope that it might be on the right track. Here it is.

A CRUCIAL TEST

Think of a person whose life is threatened by a disease. If he is not treated he will die immediately. If he is treated he will live one

additional year. However, there will be ups and downs during this last year. In order to stay alive he will now and then have to go through short sessions with painful therapies. Assume that, when we sum the happiness in his remaining year, the net will be +100. However, when we add the weights given by prioritarianism, the moral value of his additional year, because of the extra weights given to his downs, will be -1.

Note that examples of this kind can be designed regardless of which version of prioritarianism we have opted for, as long as we apply it to moments. The example trades merely on the fact that momentary suffering is given an extra moral (negative) weight.

Suppose we do save the life of the patient. On utilitarianism we are now allowed to congratulate ourselves. We may have made the right decision when we decided to spend resources on this patient. Unless there were more pressing needs we should have addressed, we did the right thing. However, on prioritarianism, we must conclude that, irrespective of whether there were any competing needs to address, we have wasted resources. We have used our resources in a way that is proscribed by the theory. In the example, on prioritarianism, the moral value of a life and the prudential value of the same life have, as it were, come apart. This implication of prioritarianism seems very problematic indeed. It is my firm intuition that the life of this patient should be saved. I suppose most people would agree about this. So, is there a way for the prioritarian to avoid it? Why not accept trade-offs within lives after all and apply prioritarianism to whole lives rather than to moments?

This does not seem to be a viable option, at least not if we are concerned with the impression that utilitarianism does not take suffering seriously enough. By focusing on moments, and by not accepting that intrapersonal compensation for severe suffering is possible, we arrive at the desired result: prioritarianism takes suffering more seriously than does utilitarianism.

Suppose instead that we give up on the idea that intrapersonal compensation is possible and apply prioritarianism to whole lives. Then it may well happen that a person who is suffering from a slight headache should be tended to rather than a person suffering from intense pain – only because his life has gone, and/or will go, much worse on the whole than the life of the person with the intense pain. But that cannot be right.

Suppose both these people will as a matter of fact lead very long lives. The person with the slight headache will each moment of his long life be just below the level where life becomes worth living, while the one with intense pain in our example will at each day of (the rest of) his life be just above this level (with the sole exception from the day when he is in terrible pain). On average, then, these two people live on almost the same level. However, the one who is all the time just

below the level where life becomes worth living will, during his entire life, garner an enormous sum of unhappiness; the one just above the same line will garner an equally enormous sum of happiness in his life, however. This means that the one in our example with the intense pain leads, on the whole, a *much* better life than the one with the slight headache. So, from a moral point of view, if we take prioritarianism to apply to entire lives and fix the moral weights with reference to this sum of happiness/unhappiness, his intensive pain matters much less than the corresponding slight headache in the life of the unhappy individual. The 'unhappy' individual with the slight headache should be tended to, rather than the one in intense pain, on such a whole-life understanding of prioritarianism. But this is not acceptable. On such an understanding of prioritarianism it does not take (severe) pain seriously. It takes pain in some circumstances *much* less seriously than utilitarianism does. On such an understanding of the priority view, therefore, it poses no serious threat to utilitarianism.

IS THE NOTION OF PRUDENCE PROBLEMATIC?

The problem I have identified is that prioritarianism implies that a life that, from a prudential point of view, is worth living may, when the priority weights are in, turn out to be not worth living. Could the prioritarian not bite the bullet and claim that the life I have described in my thought experiment is not worth living? If it sums up to -1 , when the moral weights are in, then it is not worth living. It does not matter that, from a prudential point of view, it is worth living. The prioritarian may even want to question the very notion of prudence used in the argument against the view. Does it make sense?

One may wonder, indeed, if the notion of prudence is really applicable in a contest between utilitarianism and prioritarianism. Admittedly, there is a (normative) notion of prudence, which has no place in either theory. I think of the (egoistic) idea that each person has an obligation (perhaps a rational obligation, as it is sometimes put) to pursue his or her own best interests, at least to some extent, even when it means that other persons stand to lose from this. This notion of prudence is rejected for moral reasons by both utilitarians and prioritarians. They either deny that it is real or, if they think of it as a separate 'sphere' of normativity, they still think that prudence is trumped by morality.⁶ However, we may think of prudence, not in normative but in empirical terms. We can define an action as 'prudent' if and only if it maximizes the sum total of happiness in the life of the agent. Then we can separate normative matters from linguistic matters and query

⁶ See Torbjörn Tännsjö, *From Reasons to Norms* (Dordrecht, 2010).

whether we (ever) ought to perform prudent actions. Now, on ethical egoism, we should always be prudent. On utilitarianism, we ought to be prudent if and only if being prudent maximizes the sum total of happiness in the universe. As a corollary to this utilitarian claim we can also see that in 'Robinson Crusoe' situations, where our actions do not affect the happiness of anyone else, we ought to be prudent. It is in this sense that prudence and morality come apart when we judge our example from the point of view of prioritarianism.

Could not the prioritarian invent a notion of prudence of her own, where the priority adjustment of happiness at different happiness levels is made part of the notion? This is possible, of course. On such a revisionary notion of prudence, it is imprudent to go on with the life in our example. However, the problem with this tack is exactly to do with the fact that the prioritarian has to *invent* her own notion of prudence. This notion seems to have no place in common-sense thinking about cases such as this one. To be prudent is to garner as much happiness in your own life as possible – where happiness is seen as an empirical (non-normative) phenomenon. This may come at the cost of all sorts of religious and moral ideals you may hold. If instead you follow the dictates they give you, you may come to sacrifice your happiness for the sake of these ideals. This may be seen a noble thing to do. Yet, when you do, you cannot be said to act 'prudently'. Here my intuition is that common sense has got things right.

Admittedly, this linguistic observation is not a very strong argument in defence of my position.⁷ To appreciate my (perhaps rather weak) point, however, we should keep in mind that the priority-adjusted value of well-being is a *moral* notion. When I am at a low level of happiness, say at -1 , then the priority-weighted value of my happiness (or, my suffering) is more than -1 (perhaps -2). But this does not mean that it *feels* any worse than its nominal value. It only means that added happiness (or, rather, relieved suffering) at this level is supposed to be more important than added happiness (or, rather, relieved suffering) at a higher level. So the conflict between prudence and morality, according to prioritarianism, is still there. It is real. But I would not like to sacrifice a life (my own) on the altar of (prioritarian) morality if it contains a net surplus of happiness. This is not like sacrificing your life for a noble end.

Wlodek Rabinowicz has acknowledged that the moral value and the prudential value of a life can come apart. In situations where we do not affect other people by our actions we are allowed by prioritarianism to act prudently, however. Or so he claims. But the reason he gives for this

⁷ I thank an anonymous reviewer for having pressed this point.

claim sits ill with the priority view. On Rabinowicz's understanding of the view, it is based on Rawls's concern for the separateness (or distinctness) of persons:

[I]f prioritarianism is an expression of the Rawlsian concern for the distinctness of persons, then the priority weights should only be used to balance against each other the interests of different persons, but not when it comes to *intrapersonal* balancing. In particular, we should not apply these weights when we ask whether an individual's loss in one possible outcome is *ex ante* compensated, *for that same individual*, by his gain in another possible outcome.⁸

However, if we are impressed with the idea that lives are separate (distinct) we are likely to follow Rawls and Nagel and claim that compensation between lives is impossible. We then opt for a maximin or leximin view according to which those who are worst off should be given absolute priority. Prioritarianism is different, however. It is an aggregative view allowing for compensation between lives.

The observations that we lead separate (distinct) lives often come together with another idea, however, to wit, the idea that our lives exhibit a kind of *integrity*; within lives compensation *is* possible. However, we have already seen how this idea, when wedded to prioritarianism, leads to a neglect of the special importance of suffering, taken to be a problem for utilitarianism.

Of course, if you are not prepared to bite the bullet and argue that the life we are discussing in our thought experiment is not worth living, you could simply claim that, since no one else is here affected by your decision whether to stay alive or not, you need not bother with prioritarian thinking. In Robinson Crusoe situations you may do as you see fit with your life. After all, you could claim, when a decision concerns only your own life, it is your decision and you need not then be concerned with moral thinking at all.

This strikes me as *ad hoc*, though. Why should morality not guide also merely self-regarding decisions? There is an intellectual price to be paid for this restriction of the scope of our moral theory. More importantly, however, this move seems to be of little avail if we complicate the case a bit and add a small interpersonal component to it.

THE INTERPERSONAL CASE

Let us abstract from this complication by adding a competing interest. Suppose that we could add a few minutes to an extremely happy life, or

⁸ Wlodek Rabinowicz, 'Prioritarianism and Uncertainty: On the Interpersonal Addition Theorem and the Priority View', *Exploring Practical Philosophy: From Action to Values*, ed. Dan Egonsson, Jonas Josefsson, Bjorn Petersson, Toni Ronnow-Rasmussen and Ingmar Persson (Aldershot, 2002), pp. 139–65.

save the life of our patient. Both want to have the (now) scarce medical resources. What should we do?

According to utilitarianism we should save the life of our original patient, provided the addition of a few minutes to the life of the extremely happy person did not balance out the net happiness in the remaining life of our original patient. Let us make the very plausible assumption that it does not. According to prioritarianism, we should still add a few minutes to the extremely happy individual, rather than save the life of our original patient.

But that cannot be right. And in relation to this example it would not even be of any avail to claim that prioritarianism is not applicable to merely self-regarding actions. Certainly, when there is a conflict of interests, as there is in the present example, it does apply. But its implication is horrific.

In order to avoid this implication, in relation to a specific example, the prioritarian can always move her concave function closer to the linear utilitarian one. However, as long as there remains a difference between them, a new example with the same implication can be constructed. The extra weight to suffering is always there, and, even if there exist moments where positive happiness carries a heavier moral prioritarian weight than its nominal (utilitarian) weight, according to some versions of prioritarianism, we can construct the example so that the happy periods take place above the limit where the prioritarian concave function crosses the utilitarian linear one.

CONCLUSION

Prioritarianism has to go, even in its most plausible version (where it applies to moments). Utilitarianism has gained the upper hand in our moral contest. It has proved to be superior to prioritarianism, regardless of the form in which prioritarianism has been cast – which does not show, of course, that utilitarianism is true. Some even better theory might be in the offing.⁹

torbjorn.tannsjo@philosophy.su.se

⁹ I thank the higher seminar in practical philosophy at Stockholm University, and Frans Svensson in particular, for valuable comments on an earlier draft of this article. It was also presented to a conference on hedonism in Oxford in March 2014, where I received further useful comments, in particular from Gustaf Arrhenius, Hilary Greaves and Roger Crisp. Finally, I am *extremely* grateful for comments from an anonymous reviewer for the journal; they were extraordinarily detailed, clear and constructive. This work was funded by the Bank of Sweden Tercentenary Foundation.