

Revue d'études benthamiennes

25 | 2024

Queer Utilitarianism

Bentham and Effective Altruism

Bentham et l'altruisme efficace

JEREMIAS KOH

<https://doi.org/10.4000/129mo>

Résumés

English Français

In this paper, I explain how Bentham's utilitarianism is at odds with Effective Altruism's (EA) abstract use of numbers to calculate the most 'effective' ways to do good. This is interesting because Bentham is widely regarded as the father of modern utilitarianism and EA is a movement popularly associated with utilitarianism today. My paper is divided into 3 parts. In part 1, I explain how Bentham's utilitarianism is built on a view of pleasure as widely varied and inherently subjective. In part 2, I discuss a version of utilitarianism that bears a superficial resemblance to Bentham's but diverges from his theory in crucial ways. I show that it is with this second version, rather than Bentham's, that EA shares more in common. In part 3, I sketch out a Benthamite critique of EA.

Le présent article explique en quoi l'utilitarisme de Bentham s'oppose à une utilisation abstraite des chiffres pour calculer les manières les plus « efficaces » de faire le bien, caractéristique de l'Altruisme efficace (EA, pour Effective Altruism). Paradoxalement, Bentham est généralement considéré comme le père de l'utilitarisme moderne et l'EA est aujourd'hui un mouvement communément associé à l'utilitarisme. L'article se compose de trois parties. J'y explique d'abord comment l'utilitarisme benthamien est fondé sur une conception du plaisir entendu comme très varié et fondamentalement subjectif. J'aborde ensuite une version de l'utilitarisme qui s'apparente à celle de Bentham, mais diverge de sa théorie sur des points cruciaux. Je démontre alors que l'EA est plus proche de cette seconde version que de celle de Bentham. Enfin, j'esquisse une critique benthamienne de l'EA.

Entrées d'index

Mots-clés : Bentham, altruisme efficace, égalitarisme, utilitarisme**Keywords:** Bentham, Effective Altruism, egalitarianism, utilitarianism

Texte intégral

Introduction

- 1 Utilitarianism today is often associated with Effective Altruism (EA). In general, utilitarianism can be understood as comprising 2 claims: (1) that happiness is the only intrinsic good and (2) that we should always try to do the thing that produces the most happiness. William MacAskill gives the following definition of EA: 'Effective Altruism is: (i) the use of evidence and careful reasoning to work out how to maximize the good with a given unit of resources, tentatively understanding 'the good' in impartial welfarist terms, and (ii) the use of findings from (i) to try to improve the world.'¹ It should be clear from the above definitions that utilitarianism is distinct from EA; one can subscribe to EA without thinking that happiness is the only intrinsic good or that we should always try to maximize happiness (given that 'happiness' need not be all that characterizes 'the good' in impartial welfarist terms). A more interesting question is whether the reverse is true; can one be a committed utilitarian and yet



not subscribe to (or perhaps, even be critical of) EA? This question is one I will tackle in the course of this paper. This paper is divided into 3 parts. In part 1, I discuss Bentham's view of pleasure, and explain how his utilitarianism is built on this view. In part 2, I consider a version of utilitarianism that bears a superficial resemblance to Bentham's but diverges from his theory in crucial ways. I show that it is with this second version, rather than Bentham's, that EA shares more in common. In part 3, I sketch out a Benthamite critique of EA.

Part 1

Pleasure as Subjective Experience

- 2 Popular theories about pleasure can be broadly divided into 3 groups.² First, theories that say pleasure is a quality some of our experiences have. According to these theories, for some experience to be pleasurable is simply for it to be accompanied by a certain feeling or sensation. Second, theories that say pleasure is a complex of dispositions. Gilbert Ryle, who pioneered this view, argues that 'pleasure is not a sensation at all,'³ and that when we say a certain activity is pleasurable, we're not saying that there is some feeling or sensation that this activity produces. Rather, what we mean is that the activity fulfils a certain propensity; that by engaging in said activity we are doing something we want to do.⁴ Wilson D. Wallis offers a related definition of 'objective pleasure' as 'the doing of a thing for its own sake, or more accurately, that which, all things considered, should be done for its own sake.'⁵ Third, theories that say pleasure is a propositional attitude, like belief. In this vein, Fred Feldman characterizes pleasure as an attitude that 'takes propositional entities (or states of affairs) as its object,' 'a mode of consciousness [that] takes its place among such attitudes as hope and fear, belief and doubt.'⁶ I shall refer to the first group of theories as the Simple View of pleasure (SV), the second group as the Dispositional View of pleasure (DV), and the third group as the Propositional View of pleasure (PV).
- 3 Bentham does not seem to have provided us with a vigorous definition of 'pleasure.' Nevertheless, there is good reason to think that he would reject both DV and PV and subscribe to some version of SV. Consider this passage about the pleasure and pain of others from a manuscript written in 1800:

No one man in short has any accurate and minute measure of the intensity of the sensations of any other: much less is there any such thing as an instrument that is to all persons of man a common measure for the intensity of the sensations of all. I cannot be content just to deliver my judgment and take my chance for its concurring with his.⁷

- 4 Here, Bentham claims that pleasure is essentially subjective, and argues that we should not extrapolate from personal judgments about what we find pleasurable to judgments about what other people find (or should find) pleasurable. As Carrie Shanafelt writes, 'In the case of physical pleasure... there is no common sense of what is private to one's own body, and no common sense of private physical pleasure.'⁸ Combine this with his further claim that pleasure is the only thing that is intrinsically good, and we can see why Bentham might be wary of DV and PV. Conceiving of pleasure as either a complex of dispositions or a propositional attitude raises the obvious question 'which dispositions/attitudes are good or best to have?' The same question doesn't arise for Bentham's view, according to which pleasure itself is the only good. If pleasure is understood dispositionally, then there would need to be some way to say that 'pleasure A is better than pleasure B' if A and B represent mutually exclusive sets of dispositions that come into conflict. Much the same can be said for PV; if pleasure is a propositional attitude like belief, then some pleasures would be more appropriate than others, in the same way that some beliefs are more rational to have than other beliefs. Under DV and PV then, pleasure becomes something which is no longer essentially subjective or good-in-itself, but measurable against some extrinsic objective standard.
- 5 It is also important to note here that while Bentham most likely thought that pleasure always involved some sort of pleasant feeling, his talk of pleasure is not limited to sensory pleasure, as is often alleged, but also includes intellectual, emotional, social, and other pleasures. Indeed, his *Table of the Springs of Action*⁹ lists a total of fourteen classes of pleasures and pain.

The Greatest Happiness of the Greatest Number

- 6 In *A Fragment on Government*, Bentham gives what is commonly regarded as the fundamental axiom of his utilitarian theory, stating that 'it is the greatest happiness of the greatest number that is the measure of right and wrong.'¹⁰ But why not just 'the greatest happiness' ? Taking 'the greatest number' as an additional factor in moral evaluation seems to conflict with statements made in his other works; for example, in *An Introduction to the Principles of Morals and Legislation (IPML)*

Bentham writes that 'it is for [pain and pleasure] alone to point out what we ought to do' and that the 'standard of right and wrong' are 'fastened to their throne.'¹¹ Ben Eggleston's discussion (and rejection) of 'the greatest number' as a utilitarian criterion for right action is representative of the way many moral philosophers today think of the phrase, and is worth quoting in full:

On its face, this phrase suggests that an act should not only produce as much happiness as possible but should also produce happiness for as many people as possible. That makes this phrase problematic as a criterion of right action, since it is often the case that the most beneficial act is different from the act that will spread the benefit most widely (since, in many choice situations, a small set of people has much more at stake than the rest of humanity does). In contrast, when overall well-being is conceived simply as the sum of individuals' well-being (as explained above), the 'for the greatest number' part of the phrase proves otiose. Maximizing overall well-being might often result from the act that benefits the most people, but even in that case the act is right (according to act utilitarianism) simply because it maximizes overall well-being, not because it benefits the most people. We may conclude, with Russell Hardin, that 'No philosopher should ever take the dictum of the greatest good for the greatest number seriously except as a subject in the history of thought.'¹²

7 In the passage above, Eggleston seems to assume that happiness or well-being is something that can be accurately quantified and measured, such that it is sensible to say quite abstractly, for example, that the units of happiness or well-being a certain group of people have, is greater than that possessed by a larger group of people. Though the passage quoted in the previous section suggests otherwise, Bentham himself is often taken to affirm something like this with his theory of the 'felicific calculus' (though it seems that Bentham himself never uses this term). However, a closer examination of this theory reveals the opposite and explains why 'the greatest number' is not an otiose consideration (as Eggleston puts it), but central to Bentham's utilitarianism.

8 In Chapter 4 of *IPML*, Bentham tells us how to go about measuring pleasure and pain: for an individual, 'the value of a pleasure or pain (considered by itself) will be greater or lesser according to (1) its intensity. (2) its duration. (3) its certainty or uncertainty. (4) its nearness or remoteness.'¹³ To determine the value of an act for a community, Bentham tells us to take one person from those 'whose interests seem to be most immediately affected by the act' and 'sum up the values of all the pleasures [that the act is likely to produce] on one side and of all the pains on the other. If the balance is on the side of pleasure, that is the over-all good tendency of the act to the interests of that person; if on the side of pain, its over-all bad tendency.'¹⁴ Then :

[R]epeat the above process with respect to each person whose interests appear to be concerned; and then sum the results. If this balance is on the side of pleasure, that is the over-all good tendency of the act with respect to the interests of the community; if on the side of pain, its over-all bad tendency.¹⁵

9 Bentham is also careful to acknowledge that applying this calculus is impractical and necessarily inexact, presenting it as an ideal model of calculation that 'can always be kept in view'¹⁶ rather than strictly followed. If happiness is, as Bentham has it, an aggregate of pleasures and pains, and pleasures and pains are inherently subjective, then the only way for us to sensibly estimate total happiness or well-being for any group of people would be to start with an individual subject (rather than an impersonal abstract unit of happiness) and work our way up from there.

10 We can see that some kind of democracy is clearly implied by Bentham's utilitarianism; if the happiness of the state is the happiness of all its members, then the interests of each member should be considered, and each member's interests should be represented in the relevant processes of democratic deliberation. Whether it implies egalitarianism is perhaps less clear. Some might think that given Bentham's calculus, the misery of some minority population, even if democratically represented, could possibly be justified by promoting the happiness of the majority. Bentham anticipated this worry,¹⁷ and responds with an argument that can be summarized as follows:

1. Assume that there are 4001 people in some community. Suppose the majority number 2001 and the minority number 2000. Take each happy person as counting for 1 unit of happiness.
2. Take all the happiness from the minority and give it to the majority. In the place of the happiness taken, substitute in as large a quantity of unhappiness (i.e. unhappy people as units of negative happiness).
3. The result is a net loss in happiness; before there were 4001 (2000+2001) units of happiness, after there is only 2001 (2001+2000-2000).
4. This result holds for any community divided into 2 unequal parts (though the net loss in happiness is most obvious when the difference between the sizes of both parts is small).
5. Therefore, in any community, the greatest aggregate happiness is achieved by promoting the happiness of each member.

11 Here again, we see Bentham taking the individual as the starting point and having everyone's happiness count equally. The conclusion of the above argument is one that is radically egalitarian;

maximizing utility is antithetical to oppressing a minority, however small, for the benefit of a majority, however large. For Bentham, promoting the happiness of a group of people requires promoting the happiness of each member of the group. The democratic and egalitarian implications of Bentham's moral theory are evident in his politics, which revolved around enfranchising the disenfranchised. Indeed, for Bentham, the two are inseparable, as is clearly seen in his advocacy for pleasure as the basis of political enfranchisement.

Part II

Abstract Utilitarianism

¹² What I will refer to as abstract utilitarianism is by far the dominant version of utilitarianism today. It is characterized by its commitment to cost-benefit analysis and tendency to frame utility (i.e. happiness) in terms of countable units (usually economic ones) that are abstract, fungible, and detached from subjective experience. A prime example of abstract utilitarianism in action is the use of Quality-Adjusted Life Years (QALYs) to evaluate the (moral) value of medical interventions. One QALY is equal to one year of life in perfect health. QALYs are calculated by weighting each year of life with a quality-of-life score, on a scale of 0 to 1. For example, if a certain medical procedure is estimated to give a patient 5 more years of life, with each year having a 0.5 quality of life score, then this medical procedure is expected to produce 2.5 QALYs. In *The Allocation of Health Care Resources*, John McKie, Jeff Richardson, Peter Singer, and Helga Kuhse embrace the view that according to utilitarianism, we should discriminate against the disabled in the allocation of health care resources.¹⁸ McKie et al. argue that since disabled people have, on average, lower quality-of-life scores and since health care resources are scarce, health care resources that are allocated to a nondisabled person would yield greater utility (i.e. more QALYs) than they would if they were allocated to a disabled person. Under this reasoning, it is not only right and rational to discriminate against the disabled in the distribution of health care resources, but it would also be wrong and wasteful not to.

¹³ Bentham's own views on disability were problematic and fraught with tension. As Michael Quinn explains, on the one hand, Bentham anticipated 'the central insight of the social model of disability, in seeing that the life prospects of people with impairments could be enhanced or diminished by the way in which society reacted to their impairment.'¹⁹ On the other, 'Bentham's commitment to a deficit understanding of impairment was allied to his conviction of the central role of individual responsibility for individual subsistence in the creation of wealth. This prevented him from recognizing the full implications of his 'social model' insight for the individualised and medical model to which he, for the most part, subscribed.'²⁰ That said, there is some reason to think that Bentham would have rejected the discrimination endorsed by McKie et al. As discussed, Bentham's arguments can be interpreted as going against the idea that happiness can be abstracted and measured on an objective scale. Furthermore, given the subjectivity of pleasure, Bentham would warn us to be wary of narrow and rigid definitions of happiness (or quality-of-life in this case) that centralise the experiences of certain groups of people, especially when these definitions recommend discriminatory practices.

¹⁴ Bentham's writings on sexual nonconformity reveal an alternative to abstract utilitarianism; instead of framing utility in terms of abstract units, in these writings, Bentham formulates a discourse grounded in the bodily pleasure of individuals, that recognizes (and is supportive of) the many ways in which pleasure might be experienced, and the variety of things that make different people happy.²¹ This discourse is grounded in Bentham's definition of the 'taste for any object' as an 'aptitude or disposition to derive pleasure [from] that object,'²² and his argument against the classical distinction between good and bad tastes. As he writes in *The Rationale of Reward*:

There is no taste which deserves the epithet good, unless it be the taste for such employments which, to the pleasure actually produced by them, conjoin some contingent or future utility: there is no taste which deserves to be characterized as bad, unless it be a taste for some occupation which has a mischievous tendency.²³

¹⁵ What this means is that tastes in themselves should not be classified as good or bad. The goodness or badness of particular tastes, like everything else, should be judged according to the pleasure or pain they can be expected to produce, both for individuals and for the wider community. As such, Bentham advocates for governments to openly assert the liberty of tastes, so as to eliminate any socially deleterious effects that might be produced by punishing or encouraging antipathy towards those with minoritized tastes that have no mischievous tendency. This advocacy for the liberty of tastes and focus on the subjectivity of pleasure is in stark contrast with the approach favoured by abstract utilitarianism, and the differences here have direct relevance to issues of distributive justice and political equality. In the case of deciding how best to allocate medical resources, we see abstract utilitarians advocating for discrimination against disabled people. The basis of this discrimination is

the idea that we can produce a general objective scale to accurately score any individual's quality-of-life. By rejecting this assumption, Bentham's utilitarianism recommends the opposite; we should not discriminate against the disabled by virtue of their disability. Instead, we should do our best to include them in our deliberative processes and treat their happiness as just as important as that of the non-disabled.

16 It is important to note here that accepting Bentham's utilitarianism doesn't amount to a wholesale rejection of abstract measurements relating to happiness. Rather, if Bentham is right, when we talk about and weigh 'quantities' of happiness against each other, we should always be keenly aware regarding the severe limitations of these measurements, and the fact that it is invariably those in positions of power who devise the tools by which we do the measuring.

EA and Utilitarianism

17 In *The Most Good You Can Do*, a key introductory text for EA, Peter Singer argues that in choosing a charity or cause to donate to, it is not enough that your donation will do some good.²⁴ Instead, you should only donate to the most cost-effective charities that 'help[] the most people the greatest amount per dollar.'²⁵ For EA, cost-effectiveness is calculated by looking at the expected marginal rates of return per additional dollar donated to a charity. According to GiveWell, a non-profit charity assessment and EA focused organisation, such estimates for cost-effectiveness 'include administrative as well as program costs and generally look at the cost per life or life-year changed (death averted, year of additional income, etc).'²⁶

18 This notion of cost-effectiveness is based in a similar assumption as that which motivates the use of QALYs in allocating healthcare resources. In the case of cost-effectiveness, the assumption is applied twice; first in quantifying the goodness that each donated dollar is expected to produce (hence assuming that we can produce some general objective scale by which goodness can be measured), then again in evaluating the effectiveness of charities (hence assuming that we can produce some general objective scale by which effectiveness can be measured).

19 Accordingly, the assumptions of abstract utilitarianism are the assumptions of EA. EA is heavily dependent on these assumptions; they are what give EA its distinctiveness and action-guiding power. As Nicolas Côté and Bastian Steuwer argue, this great action-guiding power is an 'incontestably attractive feature of EA':

EA provides a very simple, evidence based, and at first blush easy to apply method for assessing the donation-worthiness of any charity. It is a feature rather than a bug of this method that it generates a restricted menu of permissible options for individuals to choose from as they please. With thousands of charities in the world one could support it is impossible to make well-informed and rational philanthropic decisions unless this gigantic option space is restricted to a manageable menu of options.²⁷

20 I mentioned in the introduction of this paper that we shouldn't conflate EA with utilitarianism. One can believe that there are things other than happiness (e.g. justice, equality, political freedom) that are intrinsically good and worth promoting, and also believe that one should seek to do the most good (however one understands goodness), with the 'most good' understood in terms of cost-effectiveness. As such, one can subscribe to EA without subscribing to utilitarianism (abstract or otherwise). For abstract utilitarians however, the reverse doesn't seem to hold true; by understanding happiness in terms of abstract countable units, it naturally follows that doing the most good involves being as cost-effective as possible. It might be argued then that one cannot subscribe to abstract utilitarianism without also subscribing to EA, because abstract utilitarianism is at the heart of EA (insofar as they both rely on the same assumptions), at least as things currently stand.

Part III

The Institutional Critique of EA

21 The institutional critique is best thought of as a family of arguments directed at EA, with different arguments from this family focusing on different aspects of EA. What these arguments have in common is the claim that EA doesn't pay sufficient attention to the importance of reforming economic and political institutions. In this section, I quickly summarize 3 examples of the institutional critique. The first is given by Alexander Dietz,²⁸ the second by Antonin Broj,²⁹ and the third by Timothy Syme.³⁰ Taken together, these 3 examples give us a good idea of what the institutional critique is about. This is relevant as I mean for the Benthamite critique I sketch out in the next section to be understood as an institutional critique of EA.

- 22 Dietz's version of the critique is based in the importance of collective obligations. He argues that rather than focusing on what individuals should do, EA needs 'to think more about what they collectively should do.'³¹ That is to say, rather than only thinking of itself as being made up of individuals each of which are obligated to do the most good they personally can, EA should also think of itself as a 'collective subject of action' with a collective obligation to do the most good it can as an organization.³² Dietz explains that by not recognizing its collective obligations, EA is able to easily justify its neglect of institutional change, since it is almost always the case that an individual can do more good (at least according to how EA quantifies 'good') by donating to an effective charity than by donating to, or participating in, efforts aimed at reforming institutions that cause a lot of suffering. In other words, even if it's true that as an individual, one does the most good by donating to effective charities, the same doesn't necessarily hold true at the group level. Dietz's charge is thus not that EA is mistaken in any of its core commitments, but that it is incomplete insofar as it does not give enough thought to its collective obligation to do the most good.
- 23 Broi's version of the critique starts with the question: 'If there were an effective systemic intervention available, would the EA movement undertake it?'³³ Broi argues that the answer to this question is 'no,' because of the role played by the law of diminishing marginal returns in cause prioritization for EA. One of EA's key methodological assumptions is that all interventions have diminishing marginal rates of return, whereby initial resources spent on some intervention would be expected to have a relatively high marginal rate of return, which goes down as more resources are directed towards it. As discussed in the previous section, EA understands 'effective' interventions as those which provide the highest expected marginal rate of return. This assumption thus ends with EA focusing its efforts on 'low-hanging fruits' i.e., 'interventions which are initially very effective, but with rapidly diminishing marginal returns.'³⁴ With such interventions, EA can expect the highest marginal rate of return, and move on once these marginal rates are sufficiently diminished such that other interventions can be considered more effective. The problem with this is that systemic interventions are the opposite of low-hanging fruits (i.e. it is unlikely that initial resources directed at reforming the system will yield high marginal returns, and plausible that with more resources directed at such interventions, marginal returns go up rather than down), so even if there were effective systemic interventions available, and even if these interventions would produce the most good in the long run, EA would ignore them in favour of low-hanging fruits.³⁵
- 24 Syme's version of the critique centres on EA's empiricist approach to altruism. He argues that since 'structural change requires qualitative judgements that cannot be easily proved or disproved,'³⁶ EA's narrow prioritization of 'quantifiable data and empiricist methods' that seek to 'ground inductive conclusions about the effectiveness of available interventions' using 'observations and experiments' ('with randomized controlled trials as the "gold standard"') is 'inadequate for structural analysis and thus for taking systemic change seriously.'³⁷ Accordingly, the dilemma Syme presents for EA is between embracing structural analysis on the one hand (which would mean abandoning in large part the empiricist approach that makes EA distinctive, and making controversial judgments about social systems that are 'inevitably controversial and largely immune to empirical demonstration and refutation' thus risking a splintering of EA) and continuing to refrain from making judgments about social systems on the other (which would mean an uncritical acceptance of the status quo in which case EA cannot claim to take systemic change seriously).³⁸
- 25 To sum up: Dietz's critique is targeted at EA's lack of attention to its collective obligations, Broi's at how EA handles decisions about cause prioritization, and Syme's at how EA's empiricist approach to altruism forecloses any serious consideration of alternatives to the status-quo. The Benthamite critique I sketch out in the next section has its own unique character but can be seen as touching on issues raised by each of the 3 critiques summarized above.

A Benthamite Critique of EA

- 26 My Benthamite critique of EA can be broken into two parts. The first part is negative, in that it concludes that EA is mistaken, and can be summarized with the following premises:
- 1.If pleasure is inherently subjective, then abstract utilitarianism is mistaken.
 - 2.If abstract utilitarianism is mistaken, then EA is mistaken.
 - 3.Pleasure is inherently subjective.
- 27 Therefore,
- 4.EA is mistaken.
- 28 Premise 1 is motivated by the thought that if pleasure is inherently subjective, then it is rarely appropriate to frame happiness in terms of abstract fungible units. Instead, we need to keep in mind

that determinations and comparisons of ‘amounts’ of happiness are necessarily rough and need to start with the individual subject. In trying to decide the best course of action, it is often useful to consider measurements of things that we have good reason to think are tightly correlated with happiness, but we should take care not to conflate these things with happiness itself. Premise 2 follows from my claim that abstract utilitarianism is at the heart of EA. Premise 3 is not really argued for in this paper, but it’s what Bentham thought, so if Bentham’s conception of pleasure is right, and the other premises are correct, then EA is mistaken.

29 The second part of the critique is positive, in that it suggests how EA might reform itself in accordance with the critique, and can be summarized with the following premises:

5.If pleasure is inherently subjective, then the most effective means of promoting well-being is with universal enfranchisement.

30 Therefore,

6.EA should concern itself with securing universal enfranchisement insofar as it’s concerned with promoting well-being (from premises 3 and 5).

31 Premise 5 is motivated by Bentham’s view that what is needed for the greatest happiness is a system which genuinely values the material happiness of each individual. The importance of such a system to Bentham can be clearly seen throughout his writings for legal reform. For example, in ‘Emancipate Your Colonies!’, first published in 1793 as an address to the National Convention of France, and later reprinted in 1830 for British audiences, Bentham outlines how denying rights to certain groups of people for the sake of power and profit is antithetical to the utilitarian project of promoting the greatest happiness.³⁹ He sums up the essay saying :

You will, I say, give up your colonies—because you have no right to govern them, because they had rather not be governed by you, because it is against their interest to be governed by you, because you get nothing by governing them, because you can’t keep them, because the expense of trying to keep them would be ruinous, because your constitution would suffer by your keeping them, because your principles forbid your keeping them, and because you would do good to all the world by parting with them.

32 And concludes:

If hatred is your ruling patron, and the gratification of it your first object, you will still grasp your colonies. If the happiness of mankind is your object, and the declaration of rights your guide, you will set them free.—The sooner the better: it costs you but a word: and by that word you cover yourselves with the purest glory.

33 Along similar lines, Bentham observed that the abstract declarations of rights undertaken by wealthy countries like the USA and France did little to lessen the oppression of disenfranchised groups. Indeed, as Carrie Shanafelt explains, Bentham felt that ‘rights discourse threatened to give liberty to the powerful to oppress the victims of their prejudice without limit.’⁴⁰ This is so as :

[T]he only people with the legal means to seek redress for violation of stated rights were those with sufficient education in the law to understand what privileges they could draw upon and the resources to press the courts to defend those privileges. Without thorough reform of the entire legal code to eradicate and reverse statutory prejudice and disenfranchisement, declarations of rights could only serve to facilitate greater violence against those whom the law had weakened through political and economic privation.⁴¹

34 Much of this argument about the dangers of a seemingly benevolent rights discourse removed from meaningful political reform can be applied to EA and abstract utilitarianism. Just as how Bentham warns that abstract declarations of rights can ultimately undermine the ability of disenfranchised groups to seek political and legal redress for discrimination, EA likewise risks doing more harm than good by uncritically working within (and thus legitimizing with the cover of rational cost-effective altruism) institutions that cause much misery.

35 The Benthamite critique of EA touches on issues important to the 3 examples of the institutional critique discussed in the previous section. As per Dietz, bringing about legal and economic reform would require EA to look beyond the individual and consider its obligations as an increasingly powerful collective. As per Broi, taking the subjective nature of happiness seriously would require moving beyond and away from how EA currently thinks about cause-prioritization and cost-effectiveness. As per Syme, the task of universal enfranchisement would require EA to take structural analysis seriously.

36 That said, the Benthamite critique is interestingly distinct and constructive in that in addition to pointing out how EA is mistaken, it also suggests how EA might remedy these mistakes, as well as specific cause areas for EA to focus on. For example, meaningful enfranchisement would require democratic participation in economic processes, which in turn implies democratic access to the means

of production. Universal enfranchisement would also require ending all kinds of discrimination against marginalized groups, especially those which prevent these groups from participating in important political and economic processes. Thus, promoting more democratic access to the means of production and working towards ending the more pernicious kinds of discrimination are 2 ways in which EA can start to reform itself according to the Benthamite critique. While taking the Benthamite critique seriously requires EA to do away with many of its core assumptions, Bentham offers new principles for EA to organize around. That said, this is probably too strong for most effective altruists to find convincing, in which case they might have to acknowledge that their project is one that runs counter to Bentham's legacy. By placing primary importance on the subjectivity of happiness, Bentham ends up with a politics that is radically democratic and egalitarian. If Bentham is right about the greatest happiness, then EA cannot claim to aim at doing the most good without paying attention to institutional reform.

Bibliographie

- Alexander, Scott, 'Effective Charity—Do Unto Others,' *Effective Altruism* (2013), <https://www.effectivealtruism.org/articles/efficient-charity-do-unto-others>
- Bentham, Jeremy, 'Bentham MSS,' *University College London*, (1800), <http://transcribe-bentham.ucl.ac.uk/td/Metadata:JB/014/008/002>
- Bentham, Jeremy, *The Rationale of Reward* (John and H.L Hunt, 1825)
- Bentham, Jeremy, 'Emancipate Your Colonies!', C. and W. Reymell, *Broadstreet* (1830), https://en.wikisource.org/wiki/Emancipate_your_colonies!
- Bentham, Jeremy, *A Fragment on Government* (Clarendon Press, 1891)
DOI : 10.1017/CBO9781139163675
- Bentham, Jeremy, *Deontology, Together with A Table of the Springs of Action and Article on Utilitarianism*, ed. Ammon Goldworth (Clarendon Press, 1983)
- Bentham, Jeremy, *An Introduction to the Principles of Morals and Legislation*, eds. J.H Burns and H.L.A Hart (Clarendon Press, 1996)
DOI : 10.1093/actrade/9780198205166.book.1
- Bentham, Jeremy, *The Collected Works of Jeremy Bentham: Of Sexual Irregularities, and Other Writings on Sexual Morality*, eds. Phillip Scofield, Catherine Pease-Watkin and Michael Quinn (Oxford University Press, 2014)
- Broi, Antonin, 'Effective Altruism and Systemic Change,' *Utilitas*, 31.1 (2019)
DOI : 10.1017/S0953820819000153
- Côté, Nicolas and Steuwer, Bastian, 'Better Vaguely Right Than Precisely Wrong in Effective Altruism: The Problem of Marginalism,' *Economics and Philosophy* (2018)
- Dietz, Alexander, 'Effective Altruism and Collective Obligations,' *Utilitas* 31.1 (2019)
DOI : 10.1017/S0953820818000158
- Eggleston, Ben, 'Act utilitarianism,' in *The Cambridge Companion to Utilitarianism*, eds. Ben Eggleston and Dale Miller (Cambridge University Press, 2014)
- Feldman, Fred, 'The Good Life: A Defense of Attitudinal Hedonism,' *Philosophy and Phenomenological Research* 65 (2002)
DOI : 10.1111/j.1933-1592.2002.tb00223.x
- Katz, Leonard D., 'Pleasure,' *Stanford Encyclopedia of Philosophy*, (2016), <https://plato.stanford.edu/entries/pleasure/>
- MacAskill, William, 'The Definition of Effective Altruism,' in *Effective Altruism: Philosophical Issues*, eds. Hilary Greaves and Theron Plummer (Oxford University Press, 2019)
- Mckie, John; Richardson, Jeff; Singer, Peter, and Kuhse, Helga, *The Allocation of Health Care Resources* (Routledge, 2016)
- Quinn, Michael, 'Jeremy Bentham on Physical Disability: A Problem for Whom?,' *Review of Disability Studies* 8.4 (2014)
- Ryle, Gilbert, *Dilemmas* (Cambridge University Press, 1964)
- Ryle, Gilbert, *The Concept of Mind* (Routledge, 2009)
DOI : 10.7208/chicago/9780226922652.001.0001
- Shanafelt, Carrie D., 'Jeremy Bentham and the Aesthetics of Sexual Difference,' *The Eighteenth Century* 61.3 (2020)
DOI : 10.1353/ecy.2020.0023
- Shanafelt, Carrie, D., *Uncommon Sense: Jeremy Bentham, Queer Aesthetics, and the Politics of Taste* (University of Virginia Press, 2022)
DOI : 10.2307/j.ctv23wf3cs
- Singer, Peter, *The Most Good You Can Do* (Yale University Press, 2015)
- Syme, Timothy, 'Charity vs. Revolution: Effective Altruism and the Systemic Change Objection,' *Ethical Theory and Moral Practice* 22 (2019)
- Wallis, Wilson D., 'What is Real Pleasure ?,' *The Journal of Philosophy, Psychology, and Scientific Methods*, 16.14 (1919)
DOI : 10.2307/2940523

Notes

- 1 MacAskill, William, 'The Definition of Effective Altruism,' in *Effective Altruism: Philosophical Issues*, eds. Hilary Greaves and Theron Plummer (Oxford University Press, 2019), p. 14
- 2 Katz, Leonard D., 'Pleasure,' *Stanford Encyclopedia of Philosophy* (2016), <https://plato.stanford.edu/entries/pleasure/>
- 3 Ryle, Gilbert, *Dilemmas* (Cambridge University Press, 1964), p. 58
- 4 Ryle, Gilbert, *The Concept of Mind* (Routledge, 2009), p. 93
- 5 Wallis, Wilson D., 'What is Real Pleasure?,' *The Journal of Philosophy, Psychology, and Scientific Methods*, 16: 14 (1919), p. 384
- 6 Feldman, Fred, 'The Good Life: A Defense of Attitudinal Hedonism,' *Philosophy and Phenomenological Research*, 65 (2002), p. 607
- 7 Bentham, Jeremy, 'Bentham MSS,' *University College London*, xiv. 8 (1800), <http://transcribe-bentham.ucl.ac.uk/td/Metadata:JB/014/008/002>
- 8 Shanafelt, Carrie, *Uncommon Sense: Jeremy Bentham, Queer Aesthetics, and the Politics of Taste* (University of Virginia Press: 2022), p. 12
- 9 Bentham, Jeremy, *Deontology, Together with A Table of the Springs of Action and Article on Utilitarianism*, ed. Ammon Goldworth (Clarendon Press: 1983)
- 10 Bentham, Jeremy, *A Fragment on Government* (Clarendon Press: 1891), p. 93
- 11 Bentham, Jeremy, *An Introduction to the Principles of Morals and Legislation*, eds. J.H Burns and H.L.A Hart (Clarendon Press: 1996), p. 11
- 12 Eggleston, Ben, 'Act utilitarianism,' in *The Cambridge Companion to Utilitarianism*, eds. Ben Eggleston and Dale Miller (Cambridge University Press: 2014), p. 128
- 13 Bentham, J, *IMPL*, p. 38
- 14 Bentham, J, *IPML*, p. 39
- 15 Bentham, J, *IPML*, p. 39
- 16 Bentham, J, *IPML*, p. 39
- 17 Bentham, J, *Deontology*, pp. 309-310
- 18 Mckie, John; Richardson, Jeff; Singer, Peter, and Kuhse, Helga, *The Allocation of Health Care Resources* (Routledge: 2016)
- 19 Quinn, Michael, 'Jeremy Bentham on Physical Disability: A Problem for Whom?,' *Review of Disability Studies*, 8:4 (2014), p. 1
- 20 Quinn, M, 'Jeremy Bentham on Physical Disability,' p. 1
- 21 Shanafelt, Carrie D., 'Jeremy Bentham and the Aesthetics of Sexual Difference,' *The Eighteenth Century*, 61:3 (2020)
- 22 Bentham, Jeremy, *The Collected Works of Jeremy Bentham: Of Sexual Irregularities, and Other Writings on Sexual Morality*, eds Phillip Scofield, Catherine Pease-Watkin and Michael Quinn (Oxford University Press: 2014), p. 4
- 23 Bentham, Jeremy, *The Rationale of Reward* (John and H.L Hunt: 1825), p. 208
- 24 Singer, Peter, *The Most Good You Can Do* (Yale University Press: 2015)
- 25 Alexander, Scott, 'Effective Charity—Do Unto Others,' *Effective Altruism* (2013), <https://www.effectivealtruism.org/articles/efficient-charity-do-unto-others>
- 26 GiveWell, *Cost-Effectiveness*, Last Updated: July 2023, <https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness>
- 27 Côté, Nicolas and Steuwer, Bastian, 'Better Vaguely Right Than Precisely Wrong in Effective Altruism: The Problem of Marginalism,' *Economics and Philosophy*, (2018), p. 2
- 28 Dietz, Alexander, 'Effective Altruism and Collective Obligations,' *Utilitas*, 31:1 (2019)
- 29 Broi, Antonin, 'Effective Altruism and Systemic Change,' *Utilitas*, 31: 1 (2019)
- 30 Syme, Timothy, 'Charity vs. Revolution: Effective Altruism and the Systemic Change Objection,' *Ethical Theory and Moral Practice*, 22 (2019)
- 31 Dietz, A, 'Effective Altruism and Collective Obligations,' p. 113
- 32 Dietz, A, 'Effective Altruism and Collective Obligations,' p. 113
- 33 Broi, A, 'Effective Altruism and Systemic Change,' p. 271
- 34 Broi, A, 'Effective Altruism and Systemic Change,' p. 272
- 35 Broi, A, 'Effective Altruism and Systemic Change,' p. 272
- 36 Syme, T, 'Charity vs. Revolution,' p. 119
- 37 Syme, T, 'Charity vs. Revolution,' p. 100
- 38 Syme, T, 'Charity vs. Revolution,' p. 119
- 39 Bentham, Jeremy, 'Emancipate Your Colonies!' *C and W. Reymell, Broad Street* (1830), https://en.wikisource.org/wiki/Emancipate_your_colonies!#:~:text=I%20say%20again%2C%20emancipate%20your,a%20new%20path%20to%20glo
- 40 Shanafelt, C, *Uncommon Sense*, p. 88
- 41 Shanafelt, C, *Uncommon Sense*, p. 88

Pour citer cet article

Référence électronique

Jeremias Koh, « Bentham and Effective Altruism », *Revue d'études benthamiennes* [En ligne], 25 | 2024, mis en ligne le 30 août 2024, consulté le 02 novembre 2024. URL : <http://journals.openedition.org/etudes-benthamiennes/11702> ; DOI : <https://doi.org/10.4000/129mo>

Auteur

Jeremias Koh
National University of Singapore

Droits d'auteur

Le texte et les autres éléments (illustrations, fichiers annexes importés), sont « Tous droits réservés », sauf mention contraire.