



## Non-Consequentialist Utilitarianism

By/Par | **Nir Eyal**

Harvard University and Harvard Medical School

### ABSTRACT

Ethics 101 students read that utilitarianism is a version of consequentialism. It is not, for the following reason. Utilitarianism says that an act is morally right insofar as it maximizes total utility. Consequentialism says that an act is morally right insofar as it maximizes good consequences. Utilitarians may insist that you maximize total utility, you not thereby maximize good consequences. Such utilitarians would be non-consequentialists. I address replies to this simple argument. The replies center on the definitions of utilitarianism and consequentialism, respectively. Then I provide indications that non-consequentialist utilitarianism is not only a coherent and intriguing notion, it is also an important one. In particular, building on Kenneth Arrow, John Harsanyi and others, we may re-describe John Rawls's social theory as committed both to non-consequentialism and, provocatively but in my view inescapably, to utilitarianism. On this heretical reading, Rawls's central theory may be non-consequentialist utilitarian.

**Keywords:** utilitarianism, consequentialism, Harsanyi, Rawls, Bentham, Kymlicka

### RÉSUMÉ

L'utilitarisme est généralement considéré comme une version du conséquentialisme. Ce n'est pas le cas, pour la raison suivante. L'utilitarisme stipule qu'un acte est moralement juste dans la mesure où il maximise l'utilité totale. Le conséquentialisme dit qu'un acte est moralement juste dans la mesure où il maximise les bonnes conséquences. Les utilitaristes insistent pour que l'utilité totale soit maximisée, même si les bonnes conséquences ne sont pas maximisées. Ces utilitaristes seraient non-conséquentialistes. J'adresse des réponses à cet argument. Les réponses se centrent sur les définitions de l'utilitarisme et du conséquentialisme, respectivement. Ensuite, je fournis des indications que l'utilitarisme non - conséquentialiste n'est pas seulement une notion cohérente et intrigante, mais également importante. En particulier, en s'appuyant sur Kenneth Arrow, John Harsanyi et d'autres, nous pouvons re-décrire la théorie sociale de John Rawls comme étant à la fois non-conséquentialiste et, de manière provocatrice mais à mon avis, inévitable, utilitariste. Sur la base de cette lecture hérétique, la théorie centrale de Rawls peut être considérée comme une forme d'utilitarisme non-conséquentialiste.

**Mots clés:** utilitarisme, conséquentialisme, Harsanyi, Rawls, Bentham, Kymlicka

JEL: B20 ; D46 ; A12

## **INTRODUCTION**

Philosophers often call utilitarianism a version or a form of consequentialism. Kagan discusses the “consequentialist component of utilitarianism” and terms utilitarianism “the best known consequentialist theory” (Kagan 1998, p. 62). “The paradigm case of consequentialism is utilitarianism,” begins Walter Sinnott-Armstrong’s encyclopedia entry on consequentialism (Sinnott-Armstrong 2003). According to Steve Darwall’s introductory remarks on consequentialism, its “most popular form historically has been utilitarianism” (Darwall 2003, p. 3). For Bernard Williams (1973, p. 79), “Any kind of utilitarianism is by definition consequentialist.”<sup>1</sup>

Such statements usually seek to clarify the relatively recent term “consequentialism,” which designates the theory that we should maximize good consequences. Different versions of consequentialism, the clarification goes, differ on what constitutes a good consequence. One version takes utility to be the only good, and it tells us to maximize utility. That version is utilitarianism. Other, less familiar versions of consequentialism take the good to coincide with things other than utility.

This article shows that, under perfectly conventional definitions, utilitarianism is not a version of consequentialism. Although some utilitarian outlooks are consequentialist, others aren’t. Thus, what some of our best moral thinkers have taken to be a trivial definitional truth is quite simply false. The goal of the article is not to defend non-consequentialist utilitarianism. I only show that non-consequentialist utilitarianism is basically logically consistent, as well as an interesting and a fruitful notion. As I argue, recognizing that non-consequentialists can be utilitarians sheds new light on a number of areas in ethics and in political philosophy, including John Rawls’s theory of justice. In particular, a long-standing critique by John Harsanyi, Kenneth Arrow and others shows that Rawls’s social theory is committed to non-consequentialist utilitarianism, with indirect implications for Rawls’s derivative normative principles. I do not make this point at the level of detail appropriate for a proper critique of Rawls; only at the level that shows enough promise to prove that non-consequentialist utilitarianism should not remain a lacuna for ethicists.

Section 1 builds an initial case for the thesis that non-consequentialist utilitarianism is not a contradiction in terms. Sections 2 and 3 defend that thesis from incompatible definitions of utilitarianism and consequentialism. Section 4 illustrates that the notion of non-consequentialist utilitarianism is also fruitful, by showing that many leading ethical theories could endorse non-consequentialist utilitarianism and by suggesting heretically that Rawls’s social theory may be committed to non-consequentialist utilitarianism. A final section traces a recent history of the use of “non-consequentialist utilitarianism,” explaining how the present notion of non-consequentialist utilitarianism differs from Will Kymlicka’s and Daniel Jacobson’s notions.

---

<sup>1</sup> Similar statements abound (Scheffler 1988, p. 2; Scheffler 1994, p. 4; Rawls 1999, p. 26; Frankena 1973, p. 34; Parfit 1986, p. 26; Frey 1984, p. 4; Dancy 1993, p. 167; Pettit 1993, p. viii).

## **1. WHY UTILITARIANISM IS NOT A VERSION OF CONSEQUENTIALISM**

Consider Henry Sidgwick's definition of utilitarianism and Sam Scheffler's definition of consequentialism:

By Utilitarianism is here meant the ethical theory, that the conduct which, under any given circumstances, is objectively right, is that which will produce the greatest amount of happiness on the whole (Sidgwick 1981, p. 411).

Consequentialism provides a very simple *theory of the right*: an act is morally right (or morally permissible) if and only if it produces the best consequences (Scheffler 1988, p. 1).

The following formulations of utilitarianism and of consequentialism seem in line with these conventional definitions:

*Utilitarianism:*

An act is morally right if and only if it maximizes total utility.

*Consequentialism:*

An act is morally right if and only if it maximizes good consequences.

An important clarification should be made immediately. This article uses these working definitions, but little hangs on their wording. For example, we could replace "maximize" in these formulations by "maximize expected", "promote", or "satisfice." We could replace "morally right if and only if" by "morally preferable insofar as." We could replace "act" by "rule," "institution," "regulation," "motive," "disposition," or "global." We could define utility and good consequences indexically: relative to an agent or a moment (I assumed that they are agent-neutral and moment neutral). This would accommodate subjectivist, satisficing, scalar, indirect and global versions of utilitarianism and consequentialism. But such alterations in both definitions would not affect the argument that follows, to the effect that utilitarianism (of some form) is not always consequentialist (in that same form). Put differently, for simplicity's sake, I shall focus on establishing that utilitarianism, understood as maximizing nonindexical act utilitarianism, is not a version of consequentialism, understood as maximizing nonindexical act consequentialism. Parallel arguments would show that scalar rule utilitarianism is not a version of scalar rule consequentialism, and so forth. This makes many quibbles on the precise definitions of utilitarianism and consequentialism irrelevant to the assessment of my basic thesis. It is important, however, that, beyond this, my definitions do not substantially revise classical definitions. They are not stipulative. This helps the argument that follows avoid relying on the dirty trick of stipulating a new meaning for familiar terms, then "discovering" new things about the notions that they traditionally designate. Some definitional issues that remain relevant will be discussed in sections 2 and 3.

My argument to the effect that utilitarianism is not a version of consequentialism is simple. Utilitarianism tells you to maximize utility; consequentialism, to maximize good consequences. Maximizing utility does not definitionally maximize good consequences. You may think that a high level of utility is not always good; that it is only part of the good, which sometimes comes at the expense of other important parts: beauty, excellence, rights fulfillment, desert, equality, species diversity...; that while in actuality total utility coincides completely with good, there is a possible world in which they diverge; or that they coincide in all possible worlds, but only through external necessity, not by definition. If you also think that, when they diverge, the right thing to do is to maximize utility, and not good consequences, then you are a non-consequentialist utilitarian. In other words, a utilitarian who denies that what she should maximize, total utility, is identical to good consequences, is a non-consequentialist utilitarian. Such a utilitarian can be consistent.

Utilitarians can be consequentialists. In fact, one argument for utilitarianism is that we should maximize good consequences and that good and utility coincide, so we should maximize utility. But good and utility do not coincide by definition, and some would say that they do not coincide.

In fact, some arguments for utilitarianism are not consequentialist. One such argument says that we should maximize utility out of contractual obligations to honor a hypothetical social contract to maximize happiness impartially.<sup>2</sup> Another non-consequentialist argument says that we should maximize utility out of equal respect and concern second-personally owed.<sup>3</sup> A third argument says that we should maximize utility out of personal virtues like compassion, care, and solidarity.<sup>4</sup> A final argument says that utilitarianism generates intuitive judgments in enough concrete cases to be endorsed (for example, in the shallow pond case, in the simple trolley case, and in catastrophe cases). Thus, not only is a utilitarian non-consequentialist stance logically coherent; some arguments for utilitarianism do not assume consequentialism.

I shall now illustrate a possible non-consequentialist utilitarian stance. Imagine that justly punishing a certain evil assassin would decrease total utility. The punishment would harm the assassin, without benefiting anyone else. The assassin is now a quadriplegic, whose punishment would lack a preventive effect. Broadly perceived as innocent, her punishment would be seen as a fluke and lack deterrent or consoling effects. A judge realizing all that could at the same time acknowledge the good consequences of punishing the assassin in terms of increased proportion to desert in the world. For an evil assassin would suffer in jail, and that would make the world more just, and in that respect, a little better than a similar world in which the assassin goes free (Temkin 1994, p. 353ff.; Temkin 2003). The judge may further believe that punishing the assassin would generate more good consequences (by increasing proportion to desert in the world) than bad consequences (by decreasing utility in the world). It is true that non-punishment would be better for the assassin. But what is better for an individual is not definitionally a better consequence. A state of affairs in which many other persons suffer is better for me than one in which only I suffer, but it is a worse

---

<sup>2</sup> See cites of Harsanyi, Hare, and others in the subsection on Kymlicka under section 5.

<sup>3</sup> Kymlicka discusses this argument (1988, pp. 176–7).

<sup>4</sup> This might be thought to have been Hume's argument according to non-consequentialist feminist interpreters such as Annette Bayer.

consequence. When consequentialists say that you should maximize good, that's short for, you should maximize agent-neutrally good consequences. Now, despite believing that punishment would maximize good consequences, the judge may set the assassin free, out of utilitarianism: the judge believes that one should always maximize utility, perhaps as a matter of showing care and compassion, and that setting the assassin free would maximize utility. The judge thereby breaches what she sees as the consequentialist recommendation in the name of utilitarianism.

Or imagine a botanist who believes that preserving a certain obscure plant species would maximize diversity; that diversity is good; that in this case maximizing diversity would benefit no one; and that, because utilitarianism is true, it is not her duty to maximize diversity by preserving that plant species. This botanist could also be a non-consequentialist utilitarian.

It has long been recognized that consistent consequentialists can be non-utilitarian. What is special about the judge and the botanist is that they can be consistent non-consequentialist utilitarians. Admittedly, the judge's and the botanist's assumptions that something (proportion to desert and species diversity) is good but that it cannot provide sufficient or indeed any moral reason for action (because it does not maximize utility) may seem odd. As I mentioned, my goal herein is not to defend non-consequentialist utilitarianism, just to show that it is basically logically consistent.<sup>5</sup>

Let me address two objections to my claim that utilitarianism is definitionally compatible with non-consequentialism.

## **2. A DIFFERENT DEFINITION OF UTILITARIANISM?**

Some philosophers view utilitarianism as a version of consequentialism because they define utilitarianism differently than I proposed.

*Utilitarianism, according to a definition alternative to mine:*

1. An act is morally right if and only if it maximizes total utility.
2. The reason for 1 is that
  - 2.1 An act is morally right if and only if it maximizes good consequences.
  - 2.2 Good coincides with total utility.

---

<sup>5</sup> Some readers may argue that the particular oddity of believing that a good thing generates no moral reason for action, is so extreme that no one could be a non-consequentialist utilitarian: that no judge or botanist could hold all the beliefs that I have imputed to them. However, philosophers of arguably sound mind already believe that certain good prospects generate no moral reason for action. Consider Frances Kamm's "irrelevant utilities" (Kamm 1993, ch. 8). For Kamm, in certain situations, we have no moral reason to produce certain utilities. Since she takes this point to undermine consequentialism, she seems to hold that producing these utilities would have maximized good. In principle, Judges and botanists can entertain the beliefs that Kamm entertains. In addition, the judge and the botanist can be non-consequentialist utilitarians even if they believe that good consequences always generate some reason for action, just not sufficient reason to override their reasons to maximize utility.

Note that, in this formulation, 2 is part of the definition, not an exogenous assumption. Utilitarianism is conceived as the conjunction of 1 and a theory about why 1 is the case.

Since 2.1 is identical to consequentialism, people who accept this alternative definition of utilitarianism take utilitarianism to be a version of consequentialism. According to their definition, utilitarianism conjoins consequentialism (2.1) with welfarism about good consequences (1 and 2.2): a detail that distinguishes utilitarianism from other versions of consequentialism. Indeed, this seems to be Derek Parfit's reason for classifying utilitarianism under "C," which is his term for consequentialism:

To apply C, we must ask what makes outcomes better or worse. The simplest answer is given by *Utilitarianism*. This theory combines C with the following claim: the best outcome is the one that gives to people the greatest net sum of benefits minus burdens...There are many other versions of C.<sup>6</sup>

The alternative definition of utilitarianism is offered by many others (Sen 1993, pp. 261, 263, 279; Kagan 1998, pp. 61, 215; Griffin 1992, pp. 119–20; Darwall 2003, pp. 3-5; Dancy 1993, p. 167; Broome forthcoming; Frey 1984, p. 4; Scheffler 1988, p. 2). I, too, have taught that definition as though it were correct. There is something elegant and helpful about how the alternative definition organizes the field. Nevertheless, I now think that the alternative definition is neither the traditional one nor the most helpful one.

Let me first show that, as philosophers traditionally define the term, utilitarianism contains only conjunct 1. In its classical formulations, utilitarianism orders us to promote utility or commends us insofar as we do so. It does not add that we would thereby promote good consequences. To illustrate, note Hutcheson's, Bentham's, Mill's, and (again) Sidgwick's canonical statements of utilitarianism:

An action is best, which procures the greatest happiness for the greatest numbers; and that, worst, which, in like manner, occasions misery (Hutcheson 1971, pp. 177–8).

By the principle of utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency it appears to have to augment or diminish the happiness of the party whose interest is in question (Bentham 1996, pp. 11–2).

The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness (Mill 1990, p. 257).

---

<sup>6</sup> Parfit (1986, p. 26), but see his later concession in support of the thesis that I defend (Parfit 2003, pp. 370–1).

By Utilitarianism is here meant the ethical theory, that the conduct which, under any given circumstances, is objectively right, is that which will produce the greatest amount of happiness on the whole (Sidgwick, *op. cit.*).

These classical definitions identify right conduct with the promotion of happiness or utility. They do not identify it, additionally, with the promotion of good consequences. The notion of agent-neutral good outcomes is not even mentioned. Although happiness and utility are good for the person who has them, that is different from saying that her having them would be a good consequence (see above), and certainly different from saying that what makes her personal good important to promote is that it would be a good consequence (rather than directly that it would promote her personal good).<sup>7</sup>

This point about these definitions does not seek to provide a definitive interpretation of their authors' work in general. Although Bentham expert Douglas Long pointed out evidence from Bentham's letters suggesting that Bentham may have been a non-consequentialist utilitarian,<sup>8</sup> I am not an historian of thought, and remain agnostic on this general question. These thinkers may have elsewhere expressed what we would call consequentialism; they may have identified utilitarianism with it. It remains conceivable that all four thinkers defined utilitarianism infelicitously, doing injustice to their own "real" notions of utilitarianism, which covertly did incorporate conjunct 2 as a definitional matter (that is, not only as a further assumption about coincident properties of utilitarianism).<sup>9</sup> The present point pertains to these classical thinkers' explicit definitions of utilitarianism. It is misleading to break with these classical definitions of utilitarianism without acknowledging that one does so, and especially misleading to call the alternative definition "classical," as some writers have (Scheffler 1994, p. 3).

Some readers may wonder why *not* break with these classical definitions. After all, many terms used to carry different meanings than they do today. And the classical definitions of utilitarianism were themselves only stipulative at some point. But in this case, there exist a number of reasons why we should not break with the classical tradition (or at least why the classical tradition is equally legitimate, which would suffice to show that utilitarianism is

---

<sup>7</sup> That applies even to Bentham, although, as Paul Schofield reminded me in correspondence, Bentham is indifferent between "pleasure", "good", and "happiness" in stating the principle of utility. As Schofield points out, Bentham (1996, p. 12) writes, "By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness (all this in the present case comes to the same thing)..." However, note that "good" herein designates for Bentham only the good of the individual, not good consequences, period. This is evident in Bentham's immediate comment on what the utility or good of a community may mean. Bentham assumes that this must mean "the happiness of the community" or "the interest of the community" (*Ibid.*)—clearly not an impersonal good like, say, the equality of distribution within the community—which Bentham would not have recognized as intrinsically good.

<sup>8</sup> In a conversation.

<sup>9</sup> There is some textual support for this wild-sounding historical hypothesis. For example, " Mill writes: "That the morality of actions depends on the consequences which they tend to produce, is the doctrine of rational persons of all schools; that the good or evil of those consequences is measured solely by pleasure or pain, is all of the doctrine of the school of utility, which is peculiar to it." (Mill 2004, p. 171)

perfectly compatible with non-consequentialism on a (further) legitimate definition, a striking finding in its own right).<sup>10</sup> First, the (admittedly defeasible) presumption is always in favor of keeping existing definitions in place, stipulating new meanings for familiar terms—openly and perhaps with the community’s assent—only when a new definition would represent a major improvement. But incorporating a gloss (condition 2) into a familiar definition does not improve that definition. Usually it is best to keep definitions simple, and to complicate existing definitions only when a simple definition is too crude to pick out a notion discovered to be complex. This is precisely what is not happening here, where the notion of utilitarianism has already been picked out: philosophers have discussed utilitarianism extensively, long before the complex definition was offered. The complex definition does not purport to change either the sense or the reference of ‘utilitarianism’.

Presumably, the reason why many contemporary writers use the complex definition is that they thereby hope to situate utilitarianism in what they assume is its correct relation to other theories. Perhaps they further assume that all classical utilitarians elsewhere expressed clear commitment to consequentialism. But the complex definition compels us to accept these assumptions, on how utilitarianism relates to other theories and on how to interpret classical thinkers, if we are even to refer to utilitarianism. It is better to keep our definitions minimally committed, and to argue for determinate interrelations between notions and for determinate historical exegeses separately. In our case, keeping the definition of utilitarianism minimally committed—maintaining the simple, classical definition, as I do—is compatible with arguing separately that the best justification for utilitarianism is consequentialism.

### **3. A DIFFERENT DEFINITION OF CONSEQUENTIALISM?**

Others take utilitarianism to be a version of consequentialism because they define consequentialism differently than I proposed:

*Consequentialism, according to a definition alternative to mine:*

An act is morally right if and only if it maximizes *a certain kind* of consequence (but not necessarily good consequences).<sup>11</sup>

If we had to use this alternative definition of consequentialism, then utilitarianism might be a version of consequentialism, after all. There is no denial that the level of total utility that an act generates is a certain kind of consequence of the act. Hence, this alternative definition would have made it silly to endorse utilitarianism while opposing consequentialism. It would have probably justified dubbing utilitarianism a version of consequentialism.

---

<sup>10</sup> As Manuel Vargas pointed out to me.

<sup>11</sup> In conversation, Thomas Pogge proposed to me a more radical alternative definition, call it *Consequentialism, according to an alternative\* definition*: An act is morally right if and only if it maximizes *a certain kind of thing* (but not necessarily good consequences). What I say below in reply to the *alternative* definition of consequentialism rules out Pogge’s *alternative\** definition as well.



This alternative definition is in the spirit of definitions by Jonathan Dancy (quoted below), Philip Pettit<sup>12</sup> (who also gives other definitions),<sup>13</sup> and Walter Sinnott-Armstrong (2003). None the less, as I now argue, it does not capture what we normally mean by consequentialism. Conventionally understood, consequentialism tells you to maximize good consequences, not to maximize *whatever* consequence (Moore 1971, p. 148; Anscombe 1958, pp. 9, 12; Scheffler 1994, pp. 1–2; Scheffler 1988, p. 1; Kagan 1989, p. xi, 8; Kagan 1998, p. 61; Sen 1993, p. 260; Freeman 1994, p. 313). If a theory that tells you to maximize something, say, utility or equality, is consequentialist in a non-revisionary sense, then that theory must assume that maximizing it maximizes good consequences.

One illustration that defining consequentialism in terms of *good* consequences, as most contemporary philosophers and I do, dovetails with a shared latent notion of consequentialism better than these alternative definition does, is as follows. Even the few authors who proffer the alternative definitions of consequentialism inadvertently slip back to my conventional definition. For example, Dancy’s definition of consequentialism resembles the alternative definition. It makes no mention of the goodness or the value of consequences: “Consequentialism claims that we assess the moral worth of an action by appeal to its consequences—to the difference it makes to the world that the action was done” (Dancy 1993, p. 167). However, having given that definition, Dancy immediately explains that consequentialism is “flexible about what is to count as a “consequence”...So a consequentialist need not deny the existence of *value* in an action that is an expression of a deeply felt personal commitment...a consequentialist may adopt a theory of *value* under which the world is a *better* place for having such expressions of personal commitment going on in it” (Ibid; italics added). Contrary to first appearances, then, Dancy presupposes that, for consequentialists, the consequences to be maximized are only those laden with value, those under which the world is a better place. Dancy may toy with a new definition of consequentialism (in the first excerpt), but our shared notion of consequentialism surfaces in his wording (in the second excerpt).

In sum, since the alternative definition does not capture our notion of consequentialism, and certainly not standard definitions of the term, utilitarianism is compatible with non-consequentialism.

---

<sup>12</sup> Pettit (1993, p. viii) writes, “Roughly speaking, consequentialism is the theory that the way to tell whether a particular choice is the right choice for an agent to have made is to look at the relevant consequences of the decision: to look at the relevant effects of the decision on the world”

<sup>13</sup> Pettit sometimes defines consequentialism differently, in terms of promoting valuable things (in general), not good consequences. See, for example, Pettit (1991, p. 231): “Consequentialism is the view that whatever values an individual or institutional agent adopts, the proper response to those values is to promote them. The agent should honor the values only so far as honoring them is part of promoting them, or is necessary in order to promote them”. See also Pettit & Smith (2000, p. 121). This definition of consequentialism focuses on values not specifically on good consequences. Hence, *Consequentialism, according to an alternative\*\* definition*: An act is morally right if and only if it maximizes value (but not necessarily good consequences). Utilitarianism is not a version of consequentialism on this definition either; for one may hold that utility must always be maximized as a matter of personal virtue, although utility is not always of value. In addition, it is not clear how literally Pettit would like us to take this definition. Pettit & Smith (2000) equate it with Parfit’s (quoted above), which explicitly defines consequentialism in terms of good consequences. Thus, Pettit may tacitly assume that the only valuable things are good consequences.

#### **4. MIGHT RAWLS BE A NON-CONSEQUENTIALIST UTILITARIAN?**

Discovering logical space that was locked between pivotal theories like utilitarianism and consequentialism can be interesting for a number of reasons. First, insights about the interrelations of important theories deepen and sharpen our understanding of these theories, and of the field to which they are central. Second, if utilitarianism can be non-consequentialist, then it is invalid to argue for consequentialism as some have done: by independently establishing utilitarianism, then saying or assuming that consequentialism logically follows.<sup>14</sup> Third, discovering such logical space may enable new (political) moralities to emerge. As one example, revisionary Dworkinians may one day argue that the state must display the sovereign virtue of equal concern, rather than maximize good outcomes, and that equal concern translates, not into the liberal-egalitarian politics that Dworkin hopes to motivate, but into assigning each citizen an equal vote in a utilitarian calculus. And a future variety of Humeans and feminists may argue that in order to treat people with sympathy and care, we must maximize actual people's happiness—even when doing so does not maximize good consequences, for example, when it does not maximize abstract, luck-egalitarian equality.

Setting all that to one side, one paramount theory may already be utilitarian and non-consequentialist. That theory does not overtly espouse non-consequentialist utilitarianism (which would have foreshadowed the present article), but I shall make the preliminary case that its commitment to that position is unavoidable. I now attempt a re-positioning of Rawls that makes the core of his philosophy both non-consequentialist and utilitarian—of course, in a way of which he was unaware, and would doubtless protest. This re-positioning is grounded in a critique of Rawls made by Kenneth Arrow and others and largely ignored by committed Rawlsians.

I shall keep my discussion condensed and pitch it primarily at readers who are already familiar with Rawls's work. Even this preliminary case may however suffice for illustrating that Rawls *might* be a non-consequentialist utilitarian, and hence, that the logical space for non-consequentialist utilitarianism is important enough that it should not be defined away. I ask advance pardon from Rawls scholars for pronouncing on such a fundamental aspect of Rawls's work over the span of just a few pages. One justification is that establishing that non-consequentialist utilitarianism is an important theory—which is my ultimate goal in this section—does not require motivating very high credence that Rawls was committed to that theory; only some credence. Another justification is that an exegetical text on this matter (coming from this writer) might not be very interesting to readers of this journal.

---

<sup>14</sup> The Stanford Encyclopedia entry on consequentialism commits this common fallacy. The author Walter Sinnott-Armstrong writes, "Consequentialists also might be supported by *deductive* arguments from abstract moral intuitions," but the author's only example of such an argument is Sidgwick's argument from abstract moral intuitions to *utilitarianism* (Sinnott-Armstrong 2003; original italics). The author's consequent arguments for consequentialism are Mill's, Hare's, and Harsanyi's respective arguments for *utilitarianism*. Throughout, he seems to assume that consequentialism would logically follow from utilitarianism, and we have seen that this is not the case.

That Rawls is a non-consequentialist about the basic institutions of society is straightforward. For Rawls, basic social institutions should not maximize either good consequences or their prospect; they should treat citizens with justice, as free and equal. Only in this way may these institutions honor fair decisions between citizens' conflicting claims: decisions that hypothetical representatives on behalf of these citizens, standing as equals behind a veil of ignorance, would have freely made. The reason to honor their decisions is not a need to maximize good. Rather, political institutions must honor fair decisions because justice is the first virtue of political institutions, and it demands compliance with fair decisions. Thus, political institutions must honor these decisions even when doing so does not bring about the agent-neutrally best consequences or prospects: best, for example, in terms of improving animal welfare and species diversity,<sup>15</sup> or human excellence.<sup>16</sup> Rawls, we can safely conclude, is a basic-institution non-consequentialist.

Far more interestingly, I now add, there are reasons to think that Rawls may be committed to the position that basic institutions should maximize expected social utility. If so, then Rawls may turn out to be a basic-institution utilitarian. That would not necessarily commit Rawls to non-consequentialist utilitarianism regarding personal conduct and regarding global institutions. But as I proceed to arguing, there is a preliminary case to the effect that non-consequentialist utilitarianism is inescapable for Rawls regarding his main focus, society's basic institutions.

I am aware, of course, that any association of Rawls with a utilitarian social theory may seem doomed. Rawls could not have been more emphatic in opposing his theory to utilitarianism. Nevertheless, a number of astute observers have made a compelling case that the commitments of Rawlsian social theory are utilitarian (Harsanyi 1953; Harsanyi 1955, p. 316; Arrow 1973; Hare 1973; Barry 1996, pp. 52ff.; Pogge 1995; Pogge 2004). Some of these observers are utilitarian and some are anti-utilitarian; some are Rawlsian and some are anti-Rawlsian. These observers' shared interpretation and its implications merit our close consideration. Let me repeat their basic point, offer very brief new defenses from potential rebuttals, and connect these observers' point to the main theme of the article.

As these observers have argued, the parties in a Rawlsian original position will vote for results that maximize expected social utility. The observers' argument is simple and striking. Ignorant of the determinate social position of the individual person she represents, each representative must vote for the arrangement that maximizes expected social utility. Why? Because this is the arrangement most likely to promote the represented individual's interests. It gives her the best prospect. Hence, all parties vote for maximal expected social utility. Honoring the Rawlsian representatives' vote coincides fully with honoring a utilitarian duty to maximize expected social utility. QED.

---

<sup>15</sup> For that reason, Rawls (1999, p. 15) writes, justice as fairness "would seem to include only our relationships with other persons and to leave out of account how we are to conduct ourselves toward animals and the rest of nature." Rawls immediately concedes that such conduct also matters; it simply lies outside the scope of his book on justice. However, the book famously starts with the claim that for political institutions, justice is the first priority. Carruthers (1992, ch. 5) argues that in Rawls's contractualist framework, animals must lack intrinsic moral status.

<sup>16</sup> Rawls (1999, pp. 289f.) accepts that human excellence may be intrinsically good, and that maximizing it is not what his contractualist system prescribes. But he thinks that we should follow contractualist prescriptions.

Historically, and with a few notable exceptions, (due to Barry, Tim Scanlon, and a few others) what the majority of Rawlsian writers did about the observers' striking argument was to ignore it. Nevertheless, Rawlsians could have attempted many rebuttals. I shall offer the briefest defense against three--just detailed enough to show that the point continues to merit our attention. Note, though, that while some of the observers took their interpretation to show that Rawls's politics lacks a sound foundation (e.g., Harsanyi 1975, esp. pp. 596–7), my own point is far more modest. I argue that, whether or not Rawls's derivative political principles have a sound foundation, and whichever concrete form these principles take, there is a preliminary case for describing Rawls's political system as utilitarian, leaving much of it intact, and in place.

Let me look, then, at three rebuttals that Rawlsians might attempt to the observers' argument, corresponding to three components of Rawls's system: what we may call non-sacrifice, the higher-order interests, and publicity.

#### 4.1. *Non-sacrifice*

First, it may seem as though Rawlsian parties vote against a system that "sacrifices" the good of individuals on the altar of maximal expected social utility. Famously, the parties elect maximin, not maximum; and they protect individual liberties and rights as bulwarks against collective interests. In these ways, Rawlsians may point out, Rawls's political system clearly parts ways with utilitarian systems.

However, the parties choose maximin and individual liberty primarily in the light of risk-aversion or uncertainty-aversion, either of which is *aversion*; apparently, Rawls assumes that such second-order aversions are stronger than the first-order plans that they override, so much so that the parties refuse to maximize the prospect of fulfilling persons' first-order plans. It is more important to the parties to fulfill to the maximum these same persons' second-order aversions. But that just means that, in voting against the sacrifice of anyone's plans, basic liberties and basic income on the altar of social interests (against a sacrifice that, on a superficial level, would have maximized expected social utility), the parties do in fact maximize *properly understood* expected social utility. Why? Because fully-fleshed out, surely social utility must take second-order aversions into account. In that important sense, the parties make utilitarian choices, after all. Their refusal to sacrifice individuals and their dearest commitments is rooted in the strongest aversions of the individuals they represent. That these aversions are second-order is beside the point. What matters is that the refusal to sacrifice individuals is not rooted in a truly non-utilitarian consideration (say, that God forbade sacrificing any of Her children).

Some readers may respond that the relevant aversions characterize the parties, not the people they represent. The parties, they may add, do not maximize the expected utility of those people by refusing to sacrifice any. Risk-aversion, for example, characterizes the parties not represented people. It stems from the parties' own fiduciary obligations. Because it is not a second-order interest of represented people, it does not maximize their properly-understood expected utility. Despite this response, Rawls defends the aforementioned non-sacrificial policies based precisely on the preferences of represented people. For example, the main point of basic liberty and its priority is the overwhelming importance of core religious beliefs to those who hold them (Rawls 1993, pp. 310–2), and not directly any importance of these beliefs (or of the freedom to hold them) to the *parties*. They elect to uphold and prioritize

religious freedom because that might turn out to matter to the people their job it is to represent. Indeed, Rawls's understanding of the parties' task seems justified. For any parties to remain faithful representatives to people they represent, they must act only in the light of objective interests, desires and aversions, including risk-aversion, that these people have. The parties should not serve their own interests, desires and aversions, say, their own desire that represented people adhere to a certain faith, or their aversion to putting them at risk. An ideal representative respects the wishes and promotes the good of the people she represents; she does not *qua* good representative promote exogenous ideas on how much and when to respect their wishes and promote their good—or she would not have acted purely as their representative. In the absence of an independent *argument* to the effect that maximin, say, is always the prudent policy, a faithful representative will introduce maximin only as dictated by the represented people's own second-order aversions, interests, and requests. There is no general fiduciary duty to select maximin for the sake of persons whom one protects or represents. A good oncologist does not necessarily prescribe highly disagreeable chemotherapy *whenever* it would minimize her patient's risk of bad death by a fraction.<sup>17</sup> To invariably maximize as this patient's representative, the oncologist would need to be granting a peculiar request or desire on the part of that patient.

Admittedly, another reason that Rawlsian parties espouse maximin and individual liberty is that these principles express respect for citizens as free and equal, and thus maintain the “social bases of self-respect.” But even this Kantian-ringing reason can be plausibly seen as utilitarian. In Rawls's system, self-respect is protected because it is a primary good—only in virtue of its instrumental value for the effective pursuit of plans (and not, say, because protecting it is otherwise a duty toward oneself).<sup>18</sup> The promotion of the effective pursuit of plans may, of course, constitute a utilitarian goal.

#### 4.2. *The higher-order interests*

Second, Rawlsians might point out that, in Rawls's system, the parties assign special importance to the “higher-order” interests of the individuals they represent: their interests in preserving their two moral powers—the power to form, pursue, and revise plans rationally, and the power to treat others' pursuits with justice, reasonably. The emphasis on the higher-order interests, Rawlsians could argue, refutes the observers' interpretation. For it shows that the parties are not the maximizing egoists that the interpretation allegedly takes them to be. As Rawls put a related point, since the high-order interests “are taken to specify people's needs as reasonable and rational, the parties' aims are not egoistic but entirely fitting and proper” (Rawls 1993, p. 106).

My response is that, for Rawls, the parties prioritize the higher-order interests precisely because higher-order interests are especially strong *interests*, and fulfilling them is especially good for the persons whom the parties represent, on Rawls's so-called full conception of the

<sup>17</sup> Space constraints do not allow me to address the uncertainty-aversion interpretation of Rawls, defended e.g. in Hurley 2003. But see Eyal 2005.

<sup>18</sup> For Rawls (1999, p. 386), without self-respect “...we cannot ... continue in our endeavors” and that is what makes it “...clear why self-respect is a primary good...” See also Rawls (1993, p. 203): a Rawlsian “political society is a *good* for citizens in that it secures for them the... social bases of self-respect” (italics added), and *Ibid.*, p. 318). For a similar interpretation see Stark (1998, p. 1), Eyal 2005.

good of these persons (Rawls 1993, pp. 106–7, 202–3; Freeman 1994, p. 346). It is only because the “overall aim [of the parties] is to fulfill their responsibility and to do the best they can to advance the determinate good of the persons they represent” (Rawls 1993, p. 307; see also pp. 74–7, 105–6) that the parties focus chiefly on these relatively strong interests. Thus, when the parties prioritize the higher-order interests over other interests, the result of their choice remains the maximization of expected social utility.

### *4.3. Publicity*

A final potential objection is that Rawls’s publicity condition rules out utilitarianism. Roughly, the publicity condition states that all justified collective policy could be publicized without loss in efficiency. For example, the publicity condition rules out policies that require deceit and misinformation. It may seem as though Rawls’s championship of the publicity condition makes him a clear non-utilitarian, for among the policies that the condition rules out, say, for involving deceit, are ones that otherwise maximize utility.<sup>19</sup>

Nevertheless, in Rawls’s own system, the publicity condition, far from being anti-utilitarian, is potentially the handmaiden of utilitarianism. It may serve a deeper utilitarian goal: maximizing utility over time. The parties value publicity precisely because it enhances long-term stability, which they value because it enables citizens to reap the fruits of a well-ordered political system over time (e.g. Rawls 1999, §69, §82). If the observers are right that the fruits of that system are utilitarian then publicity has clear utilitarian value in the long run. Admittedly, full publicity would undermine some policies that utilitarians would have recommended if they overlooked the utilitarian value of publicity. But that just shows that utilitarianism does not recommend all policies that it initially appears to recommend.

In sum, the decisions of the parties in the original position may at first appear non-utilitarian, but in sophisticated ways, these decisions may well turn out to maximize expected utility. Insofar as the derivation of Rawlsian principles for the state is sound, sound utilitarianism may generate the same principles for the state.

Rawlsians should stop ignoring the observers’ critique. For one thing, that critique can be interpreted—contentiously but, further discussion may show, inescapably—to mean that Rawls’s principal philosophical contribution, his social theory, occupies the logical space that I have charted—a space for non-consequentialist utilitarianism.

While nothing that I have said logically entails that any part of Rawls’s derivative politics is unjustified (here my purport may differ from that of some of the observers cited), my proposed re-positioning of Rawls is not merely a re-description. If proven right, it could threaten tenets of his politics indirectly. For example, on a number of political issues, Rawls is explicitly or implicitly an absolutist. In a standard interpretation, Rawls assigns nearly or completely absolute (in his terms, lexicographical) priority to the interests of the worst-off (over those of other citizens); to the preservation of the two moral powers (over lower-order interests); to some principles of justice (over others), and arguably also to the social bases of self-respect (over other primary goods). If the observers’ interpretation is shown to be right, such that Rawls’s theory is indeed non-consequentialist utilitarian, then these absolute

---

<sup>19</sup> For a recent defense of the publicity condition, see Hadfield and Macedo (forthcoming 2011). While I tend to accept Gerald Cohen’s and Marc Schroeder’s recent critiques of the publicity condition, my argument does not rest on its critique.

priorities become difficult to sustain. If, for example, the higher-order interests command priority only insofar as they comprise citizens' strongest desires, then surely their priority should be charitably re-interpreted as less than (nearly) absolute. It is unrealistic to assume that every citizen—each of your fellow citizens, say—ascribes (nearly) absolute priority to her higher-order interests over all her other plans of life—that no one would take a small probability of minor harm to her higher-order interests over guaranteed total failure on her other plans of life. Nor is there any reason to assume that idealized citizens who prioritize justice over their own pursuits must as such give the highest priority to keeping these priorities in the future. They may prefer to pursue and accomplish their plans or justice now over optimally fulfilling their so-called higher-order interests in maintaining their moral powers intact for later. Similarly, even if Rawls is right that self-respect is “perhaps the most important primary good”, it is unrealistic to assume that any tiny amount of social support for self-respect is more indispensable to the successful pursuit of each and every citizen's plan than are enormous amounts of any other social primary good. Re-describing Rawls as a non-consequentialist utilitarian, far from being merely semantic, could force Rawlsians to abandon the lexical structure of some of their political commitments. That alone, I believe, shows that non-consequentialist utilitarianism is a fruitful and an important notion.

## **5. “NON-CONSEQUENTIALIST UTILITARIANISM”: A RECENT HISTORY**

A number of philosophers, and, most recently, Will Kymlicka and Daniel Jacobson, have labeled certain moralities “non-consequentialist utilitarian” or “deontological utilitarian.” Let me state briefly why the moralities that they discussed are *not* non-consequentialist utilitarian in the sense I have expounded.

### *5.1. Kymlicka*

In an ingenious article, Will Kymlicka describes a utilitarian outlook that he calls “deontological:”

On one interpretation utilitarianism is...a moral theory because it purports to treat people as equal, with equal concern and respect. It does so by counting everyone for one, and no one for more than one...The problem, on this interpretation of utilitarianism, is how to treat distinct people fairly. The standard solution is to give each person's interests equal weight. Each person's life matters equally, from the moral point of view, and hence each person's interests deserve equal consideration... If we decide how to act on this basis, then...[m]aximization occurs, but as a by-product of a decision-procedure that is intended to aggregate people's preferences fairly (Kymlicka 1988, pp. 177–8).

Kymlicka holds important utilitarian theories to be deontological in his sense:

...it is the concern with equal consideration that clearly underlies Bentham's argument and is explicitly affirmed by recent utilitarians such as John Harsanyi and James Griffin. And while this is not his preferred method, R. M. Hare too

claims that one could defend utilitarianism by reference to a foundation premise of equal consideration (pp. 176–7).<sup>20</sup>

It may seem as though, on Kymlicka's interpretation, the thinkers he cites—Bentham, Harsanyi, Griffin and Hare—are non-consequentialist utilitarians in my sense. But his interpretation would not imply that. Kymlicka's interpretation implicitly takes consequentialism to mean something different than it means to me, and closer to:

*Consequentialism, according to a second alternative definition:*

An act is morally right if and only if *and because* it maximizes good consequences.

This second alternative definition, which Kymlicka seems to use, differs from the one that I use above by incorporating “and because.” It assumes that, for consequentialists, it is the contribution to the best outcomes that *makes* acts right. Put differently, acts are right *in virtue* of maximizing good (not the other way around, say). Therefore, as Kymlicka interprets these utilitarian thinkers, they could still accept that right acts always maximize good—they would have to deny only that this is what makes these acts right. In other words, these thinkers could remain consequentialists in my sense. They affirm deontology only in the different sense of insisting that maximization “occurs, but as a by-product” (by that Kymlicka may mean, not as the necessary motivation, or not as the justifying ground of the act). By Kymlicka's lights, these thinkers are deontologists, because for them, maximization of good is not what makes actions right, although it does for them coincide with right action. But since according to these thinkers as Kymlicka understands them, maximization of good and right action do coincide, then, by my lights, these thinkers remain consequentialists.

Admittedly, what I have called the second alternative definition of consequentialism, which Kymlicka seems to use, is in many ways a good definition,<sup>21</sup> and some recent definitions of consequentialism incorporate similar “and because” clauses.<sup>22</sup> However, it is not the shorter,

---

<sup>20</sup> Kymlicka cites Bentham (1996, pp. 26–7), Harsanyi (1955, pp. 315–6), Griffin (1986, pp. 167–70, 208–15, 239–42, 295–301), and Hare (1984, pp. 106–12). Kymlicka himself does not endorse utilitarianism, not even on that “deontological” interpretation.

<sup>21</sup> For example, consider the philosophical position that right acts are right because they are virtuous and that, unbeknownst to agents, God rewards right action by making the entire world instantaneously better, something that God never does otherwise. Intuitively, that position could be non-consequentialist, although it entails that all and only right acts maximize good consequences. What seems to be missing is the claim that these acts are right because they maximize good consequences: that they maximize good not merely as a “side effect” of being right.

<sup>22</sup> See, e.g., “consequentialism is an agent-neutral theory that takes all actions to be permissible (or impermissible) purely in virtue of the value of their resultant states of affairs” (Portmore 2005, p. 95; italics added). See also Kagan's definition of foundational (as opposed to factorial) consequentialism (Kagan 1998, pp. 212ff.), and Parfit's definition of consequentialism as (fundamentally) the claim that “There is one *ultimate* moral aim: that outcomes be as good as possible” (Parfit 1986, p. 24; italics added). Conceivably, earlier definitions did not incorporate the “and because” clause only since appreciation for the importance of such clauses is a recent development in ethics. As Iwao Hirose suggested to me, such appreciation became standard only during recent debates on whether right acts are right because they are not reasonably rejectable, or they are not reasonably rejectable because they are right.



classical definition that we started out with. Therefore, Kymlicka has indicated no theory, historical or possible, that is non-consequentialist utilitarian in the specific sense that I started out with. Because on my definition of consequentialism, more thinkers turn out to be consequentialist and fewer, non-consequentialist, my own claim that some utilitarians are non-consequentialist is more ambitious.

There is also a fallacy in Kymlicka's argument. Consistently applied, Kymlicka's "and because" addition would have prevented Bentham, Harsanyi, Hare, and the other thinkers whom Kymlicka calls "deontological utilitarians" from being utilitarian—either consequentialist or deontological. Why? Because Kymlicka seems to determine ethicists' positions by examining what these ethicists see as the justifying ground of action. Therefore, a consistent Kymlicka would also have to use:

*Utilitarianism, according to a second alternative definition:*

An act is morally right if and only if *and because* it maximizes total utility.

The problem for Kymlicka is that on *that* definition, Bentham, Hare, and their peers are not utilitarian either. On his own interpretation, they deem acts right not *because* these acts maximize collectives' total utility, but *because* these acts give each individual equal consideration. "Maximization of utility", these thinkers hold in his interpretation, "occurs, but as a by-product". Being non-utilitarian, these thinkers could not be non-consequentialist utilitarian. So a consistent Kymlicka would not have made Kymlicka's actual claim.

## 5.2. Jacobson

In a 2003 article, Daniel Jacobson argues that "the inclusiveness of Mill's conception of utilitarianism belies the common tendency to read his work with certain developments of modern consequentialism too much in mind." Mill's work, Jacobson argues, is so "ecumenical" and non-committal that at points, Mill could be read as non-consequentialist (Jacobson 2003, esp. pp. 8, 14–5). However, little in Jacobson's early argument forces us to ascribe to Mill a determinate position that is both non-consequentialist and utilitarian. The present article is more ambitious, indicating a single position that is both utilitarian and non-consequentialist. Put differently, if utilitarianism excluded non-consequentialism, in the way that greenness excludes redness, then the early Jacobson would be saying that a certain apple is red and green in different parts, and the present author would be saying more ambitiously that apples can be red and green all over.

Jacobson's 2008 article "Utilitarianism without consequentialism" is ambitious:

there is no paradox involved in claiming that there is logical space for a utilitarian theory that rejects consequentialism, and there is considerable evidence for ascribing such a view to that most renowned, though not most orthodox, utilitarian, John Stuart Mill (Jacobson 2008, p. 191).

Jacobson's interpretation is now that "In Mill's view, morality does not treat everyone's happiness in exactly the same way (as deontic impartiality demands) even though everyone's happiness is of equal value (as axiological impartiality requires)." (p. 190) For Jacobson, deontic impartiality is the claim that "everyone's happiness counts in exactly the same way, when it comes to evaluating acts of right and wrong" (p.168). Axiological impartiality is the

claim that “everyone’s happiness, if equal in quantity and quality, is equally valuable” (Ibid). Consequently, for Jacobson, on Mill’s view, non-consequentialist utilitarianism follows. My discussion differs from Jacobson’s in not revolving around the notions of impartiality and agent-relative duties and options. Furthermore, my definition of utilitarianism differs from Jacobson’s in viewing utilitarianism as a theory of the right, not also as an axiological theory of the good. My definition may be more in line with Mill’s, for whom “actions are right in proportion as they tend to promote happiness” (*italics added*).

## **CONCLUSION**

Although a utilitarian view can be consequentialist, it need not be consequentialist, not by definition. We can coherently hold that right acts maximize utility without holding that they maximize good outcomes. That insight does not rest on significant departure from classical definitions. In particular, the most influential political theory of our time, Rawls’s theory of the well-ordered society, has utilitarian and non-consequentialist commitments, which question the validity of its political output.<sup>23</sup>

## **REFERENCES**

- Anscombe, E. (1958). ‘Modern Moral Philosophy’. *Philosophy*, 33: 1-19.
- Arrow, K. J. (1973). ‘Some Ordinalist-Utilitarian Notes on Rawls’s Theory of Justice’. *Journal of Philosophy*, 70 (9): 245–63.
- Barry, B. (1996). *Justice and impartiality*. New York: Oxford UP.
- Bentham, J. (1996). *An Introduction to the Principles of Morals and Legislation*. Ed. by J. H. Burns & H. L. A. Hart. Oxford: Oxford UP.
- Broome, J. (1995). *Weighing Goods*. Oxford: Blackwell.
- Broome, J. (1999). *Ethics out of Economics*. Cambridge: Cambridge UP.
- Broome, J. (forthcoming). ‘Equality versus priority: a useful distinction’, in D. Wikler et al (eds.), *Health, Well Being, Justice: Ethical Issues in Health Resource Allocation*.
- Carruthers, P. (1992). *The Animals Issue: Morality in Practice*. Cambridge: Cambridge UP.
- Dancy, J. (1993). *Moral Reasons*. Oxford: Blackwell
- Darwall, S. (2003). ‘Introduction’, in Darwall (ed.), *Consequentialism*. Oxford: Blackwell, p.1-8.
- Darwall, S. (ed.) (2003). *Consequentialism*. Oxford: Blackwell.

---

<sup>23</sup> For helpful comments, I am grateful to Sarah Conly, Guy Fletcher, Iwao Hirose, Douglas Long, Tim Mulgan, Thomas Pogge, Doug Portmore, Leah Price, Alex Rajczi, Philip Schofield, Drew Schroeder, Manuel Vargas, Drew Schroeder, participants of the 10<sup>th</sup> International Society for Utilitarian Studies conference in Berkeley, participants of the 6<sup>th</sup> European Conference for Analytic Philosophy in Krakow, participants of the Deux Siècles d’Utilitarisme conference in Rennes, and anonymous referees.

- Eyal, N. (2005a). “‘Perhaps the Most Important Primary Good’: Self-Respect and Rawls’s Principles of Justice’. *Politics, Philosophy, and Economics* 4 (2):195-219.
- Eyal, N. (2005b). ‘Review of Susan L. Hurley, *Justice, Luck, and Knowledge*’. *Economics and Philosophy*, 21:164-71.
- Feldman, F. (1997). *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*. Cambridge: Cambridge UP.
- Frankena, W. (1973). *Ethics*, second edition. Englewood Cliffs: Prentice Hall.
- Freeman, S. (1994). ‘Utilitarianism, Deontology, and the Priority of Right’. *Philosophy and Public Affairs*, 23 (4): 313–49.
- Frey, R. G. (1984). ‘Introduction. Utilitarianism and Persons,’ in Frey (ed.), *Utility and Rights*. Minneapolis: University of Minnesota Press, p.1-19.
- Frey, R. G. (ed.) (1984). *Utility and Rights*. Minneapolis: University of Minnesota Press.
- Griffin, J. (1986). *Wellbeing*. Oxford: Oxford UP
- Griffin, J. (1992). ‘The Human Good and the Ambitions of Consequentialism’. *Social philosophy and Policy*, 9: 118–32.
- Hadfield, G. K. and Macedo, S. (Forthcoming 2012). ‘Rational Reasonableness: Toward a Positive Theory of Public Reason’. *Law & Ethics of Human Rights*, 4(3).
- Hare, R. M. (1973). ‘A Review of Rawls’ *A Theory of Justice*, parts I & II’. *The Philosophical Quarterly*, 23 (91, 92): 144–55, 241–52.
- Hare, R.M. (1984). ‘Rights, Utility and Universalization: Reply to J. L. Mackie’, in Frey (ed.), *Utility and Rights*. Minneapolis: University of Minnesota Press, p.106–20.
- Harsanyi, J. C. (1953). ‘Cardinal Utility in Welfare Economics and in the Theory of Risk-taking’. *The Journal of Political Economy*, 61 (5): 434–5.
- Harsanyi, J.C. (1955). ‘Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility’. *The Journal of Political Economy*, 63 (4): 309–21.
- Harsanyi, J.C. (1975). ‘Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls’s Theory’. *American Political Science Review*, 69 (2): 594–606.
- Hurley, S.L. (2003). *Justice, Luck and Knowledge*. Oxford: Oxford University Press.
- Hutcheson, F. (1971). ‘Inquiry into the Origins of our Ideas of Beauty and Virtue’, in his *Two Treatises*, second edition. New York: Garland Publishing.
- Jacobson, D. (2003). ‘J.S. Mill and the Diversity of Utilitarianism’. *Philosophers’ Imprint*, 3 (2): 1-18.
- Jacobson, D. (2008). ‘Utilitarianism without Consequentialism: The Case of John Stuart Mill’. *Philosophical Review*, 117 (2): 159–91.
- Kagan, S. (1989). *The Limits of Morality*. Oxford: Clarendon Press.
- Kagan, S. 1998). *Normative Ethics*. Boulder, Colorado: Westview.
- Kamm, F. M. (1993). *Morality, Mortality*, Vol. I. New York: Oxford UP.
- Kymlicka, W. (1988). ‘Rawls on Teleology and Deontology’. *Philosophy and Public Affairs*, 17 (3): 173–90.

- Lyons, D. (1965). *Forms and Limits of Utilitarianism*. Oxford: Clarendon Press.
- Mill, J. S. (1990). 'On Utilitarianism', in his *Utilitarianism/On Liberty/Essay on Bentham*. Ed. by M. Warnock. Glasgow: Fontana Press, p.251–321.
- Mill, J.S. (2004). 'Bentham', in A. Ryan (ed.), *J. S. Mill and Jeremy Bentham, Utilitarianism and Other Essays*. London: Penguin, p.132–76.
- Moore, G. E. (1971). *Principia Ethica*. Cambridge: Cambridge UP.
- Parfit, D. (1986). *Reasons and Persons*. Oxford: Clarendon Press.
- Parfit, D. (2003). 'Justifiability to Each Person'. *Ratio*, 16 (4): 368–90.
- Pettit, P. (1991). 'Consequentialism' in P. Singer (ed.), *A Companion to Ethics*. Oxford: Blackwell, p.230–40.
- Pettit, P. (1993). 'Introduction', in P. Pettit (ed.), *Consequentialism*. Aldershot: Dartmouth Publishing Company, p.viii–xix.
- Pettit, P. & Smith, M. (2000). 'Global Consequentialism', in B. Hooker et al (eds), *Morality, Rules, and Consequences*. Edinburgh: Edinburgh UP, p.121–33.
- Pogge, T. (1995). 'Three Problems with Contractualist-Consequentialist Ways of Assessing Social Institutions'. *Social Philosophy and Policy*, 12 (2): 241–66.
- Pogge, T. (2004). 'Equal Liberty for All?' *Trans.* by D. Tulodziecki. *Midwest Studies in Philosophy*, 28: 266–81.
- Portmore, D. W. (2005). 'Combining Teleological Theories with Evaluator Relativism: A Promising Result'. *Pacific Philosophical Quarterly*, 86 (1): 95-113.
- Rawls, J. (1999). *A Theory of Justice*, revised edition. Cambridge, MA: Harvard UP.
- Scheffler, S. (1988). 'Introduction', in S. Scheffler (ed.), *Consequentialism and its Critics*. Oxford: Oxford UP: 1-13.
- Scheffler, S. (1994). *The Rejection of Consequentialism*, revised edition. Oxford: Oxford UP.
- Sen, A. K. (1993). *Utilitarianism and Welfarism*. Reprinted in Darwall (ed., 2003): 259–85.
- Sidgwick, H. (1981). [1907]: *Methods of Ethics*, seventh edition. Indianapolis: Hackett.
- Sinnott-Armstrong, W. (2003). 'Consequentialism'. *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. Available online at <http://plato.stanford.edu/entries/consequentialism/>.
- Stark, C. A. (1998). 'Self-Respect'. *Routledge Encyclopaedia of Philosophy*. Available online for subscribers.
- Temkin, L. (1994). 'Weighing Goods: Some Questions and Comments'. *Philosophy and Public Affairs*, 23 (4): 350–80.
- Temkin, L. (2003). 'Egalitarianism Defended'. *Ethics*, 113 (2003): 764–82.
- Williams, B. (1973). 'A Critique of Utilitarianism', in Smart & Williams (eds), *Utilitarianism—For and Against*. Cambridge: Cambridge UP (1973), p.77-150.