

## Sidgwick's Axioms and Consequentialism

*Robert Shaver*

*Philosophy, University of Manitoba*

[draft—not for citation w/o permission; comments very welcome]

Sidgwick gives four tests for highest certainty.

(1) The terms of the proposition must be clear and precise.

(2) The proposition must be (as far as I can tell by reflection on it) self-evident.

(3) The proposition must be consistent with other propositions I take to be self-evident.

(4) The proposition must not be denied by someone of whom I have no more reason to suspect of error than myself.<sup>1</sup>

---

<sup>1</sup> Henry Sidgwick, The Methods of Ethics (Indianapolis: Hackett, 1981) pp. 338-42.

Subsequent parenthetical references to ME are to this, the seventh, edition of the Methods (1907). Parenthetical references to earlier editions give the edition followed by the page.

The first edition came out in 1874, the second in 1877, the third in 1884, and the fourth in 1890 (all London: Macmillan). Other works by Sidgwick cited are “Mr. Barratt on ‘The

Suppression of Egoism” (B), Mind (o.s.) 2 (1877): 411-12; “A Criticism of the Critical

Philosophy I” (CC), Mind (o.s.) 8 (1883): 69-89; “The Establishment of Ethical First

Principles” (EP), Mind (o.s.) 4 (1879): 106-11; “Some Fundamental Ethical

Controversies” (FC), Mind (o.s.) 14 (1889): 473-87; Lectures on the Ethics of T. H.

Green, Mr. H. Spencer, and J. Martineau (GSM) (London: Macmillan, 1902); Lectures

on the Philosophy of Kant and Other Philosophical Lectures and Essays (LK) (London:

Macmillan, 1905); “Professor Calderwood on Intuitionism in Morals” (PC), Mind (o.s.) 1

(1876): 563-6; “Hedonism and Ultimate Good” (UG), Mind (o.s.) 2 (1877): 27-38.

When he applies these tests to common sense morality, he finds little of highest certainty. In contrast, when he applies these tests to his own axioms—axioms that he claims lead to utilitarianism—he finds these axioms to have highest certainty. The key axioms supporting utilitarianism are

*Personal Irrelevance*: “The good of any one individual is of no more importance, from the point of view (if I may say so) of the Universe, than the good of any other; unless, that is, there are special grounds for believing that more good is likely to be realised in the one case than in the other” (ME 382).

*General Good*: “[A]s a rational being I am bound to aim at good generally,—so far as it is attainable by my efforts,—not merely at a particular part of it” (ME 382).

From these axioms Sidgwick claims to deduce

*Benevolence*: “[E]ach one is morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him” (ME 382).<sup>2</sup>

Some think that this is Sidgwick’s main argument for utilitarianism, and indeed a convincing argument.<sup>3</sup> Others find the main argument instead in a coherentist argument

---

<sup>2</sup> I put aside the issue of whether *benevolence* has the same status as the other axioms. Sidgwick first presents it as deduced from some of them. If this is the only way of arriving at it, it fails test (2). But Sidgwick treats it as having the same status. Perhaps he thinks that it can be arrived at just by understanding it, as well as by being inferred. For one discussion, see J. B. Schneewind, Sidgwick’s Ethics and Victorian Moral Philosophy (Oxford: Clarendon, 1977) pp. 294-5.

that presents utilitarianism as the product of a cleaned-up common sense morality.<sup>4</sup> Oddly, it is fairly rare for anyone to give careful attention to the argument from the axioms.<sup>5</sup> This is especially odd given the accolades the Methods has received—Broad finds it “the best treatise on moral theory that has ever been written,” Parfit “the best book on ethics ever written,” “the book it would be best for everyone interested in ethics to read, remember, and be able to assume that others have read.”<sup>6</sup> The lack of attention may be, in part, because Sidgwick himself admits that the argument does not work against an egoist who refuses to take up the point of view of the universe. But Sidgwick does think the argument works against the common sense moralist. He thinks a common sense moralist must agree with the axioms and so with *benevolence* as one principle, and

---

<sup>3</sup> See Peter Singer, “Sidgwick and Reflective Equilibrium,” The Monist 58 (1974): 490-517, 516, 517; “Ethics and Intuitions,” Journal of Ethics 9 (2005): 331-52, 351.

<sup>4</sup> For this debate, see, in addition to Singer, “Reflective Equilibrium,” Anthony Skelton, “Henry Sidgwick’s Moral Epistemology,” Journal of the History of Philosophy 48 (2010): 491-519; Steven Sverdlik, “Sidgwick’s Methodology,” Journal of the History of Philosophy 23 (1985): 537-53; David Brink, “Common Sense and First Principles in Sidgwick’s Methods,” Social Philosophy and Policy 11 (1994): 179-201; Robert Shaver, Rational Egoism (Cambridge: Cambridge University Press, 1999) pp. 62-71; Bart Shultz, Henry Sidgwick: Eye of the Universe (Cambridge: Cambridge University Press, 2004) ch. 4; David Phillips, Sidgwickian Ethics (Oxford: Oxford University Press, 2011) ch. 3.

<sup>5</sup> I note some of the careful treatments below, mainly in notes.

<sup>6</sup> C. D. Broad, Five Types of Ethical Theory (London: Kegan Paul, 1930) p. 143; Derek Parfit, On What Matters (Oxford: Oxford University Press, 2011) v. 1, pp. xxxiii, xl.

that there are no other axioms that pass his tests and could conflict with these (ME 421). Hence it is worth asking whether Sidgwick has a good argument from the axioms against non-egoists.

The axioms face challenges from two sides.

First, Sidgwick thinks he has no more reason to suspect the egoist of error than himself. By test (4), the egoist must not deny the axioms. But it would seem that an egoist would reject *general good* and *benevolence*.

Second, Sidgwick thinks he must show that the common sense moralist agrees to the axioms. *Personal irrelevance* and *benevolence* seem to say that the only reason for departing from being indifferent to who gets some good, or from pursuing the good of another just as one pursues one's own good, is that more good would be produced. But the common sense moralist will not agree that this is the only reason. The common sense moralist thinks there are many reasons—say I have made a promise or owe gratitude to one party, or one party deserves the good.

In reply to the threat of an egoist's disagreement, I argue, in Section I, that not only *personal irrelevance* but all the axioms above should be read as having as their antecedent "from the point of view of the universe." I reply to the worry that, if so, the axioms become tautologous. I also reply to the worry that, since the egoist now agrees with the axioms, there is no dualism of practical reason.

In reply to the threat of a common sense moralist's disagreement, I argue, in Section II, that each axiom states one duty. None states that it is the only duty. The argument against the common sense moralist concerns not *benevolence*, but whether there are further duties that pass the tests. I reply to the worry that, if so, Sidgwick is

wrong to conclude that the axioms by themselves yield consequentialism. I also reply to the worry that *benevolence* is not even one duty.

In section III, I note what Sidgwick needs to move from *benevolence* to consequentialism. It is especially important for him to be able to reject duties other than *benevolence*, where these duties are understood in effect to be prima facie duties. I raise the worry that here Sidgwick is unfair to common sense morality and argue that he does not have wholly successful replies. In an Appendix, I review some of the textual changes Sidgwick made.

Some preliminaries: I ignore the hedonist side of utilitarianism, as Sidgwick does in Methods III.xiii. The issue is the argument for consequentialism.

I take the axioms presented in III.xiii.3 as canonical. I ignore Sidgwick's sloppy labeling elsewhere of (for example) egoism and utilitarianism as axioms.<sup>7</sup>

In interpreting Sidgwick, my aim is to have the axioms pass his tests. Many interpreters seem not to have this aim. Many (understandably) aim to have the axioms yield consequentialism. But that dooms them to fail the agreement test. The axioms would fare no better than common sense morality. Sidgwick's best hope is to find axioms that pass the tests and then argue that additional considerations support consequentialism. This mimics his strategy of adding, to consequentialism,

---

<sup>7</sup> For discussion, and references to the many commentators who make the sloppiness point, see Schneewind, pp. 304-8; Shaver, pp. 74-77, 79-80. I put aside the issue, noted to me by XXX, of whether there is, or ought to be, an axiom of egoism (or something close to egoism) outside III.xiii. For an argument in favour, see Phillips, pp. 126-133, 138-9; for an argument against, see Shaver, pp. 83-98.

considerations (namely hedonism) that support utilitarianism. This strategy motivates the main distinctive features of my interpretation—stressing that the axioms include reference to the point of view of the universe, that they do not claim to state our only duties, and hence that the main argumentative action consists in moving from the axioms to consequentialism.

## I

It seems that egoists would deny *general good* and *benevolence*. *Personal irrelevance*, however, is worded to avoid this. It states what follows from taking up the point of view of the universe. Egoists could agree about what follows, since they do not take up this point of view.<sup>8</sup>

This reply faces two objections: it threatens to make *personal irrelevance* tautologous, and it does not save *general good* or *benevolence*.

Consider the tautology worry first. Sidgwick starts the axioms chapter by cautioning against “sham-axioms” that “appear certain and self-evident because they are substantially tautological” (ME 374-5). He catalogues these in Plato, Aristotle, the Stoics, and Butler, poking fun at “it is right to act rationally,” “it is right that the Lower

---

<sup>8</sup> Broad notes that the egoist can agree with *personal irrelevance* by “remark[ing] that, after all, he is not the Universe, and therefore it is not obvious that he ought to take the Universe’s point of view” (*Five Types* pp. 244-5). One might object that an egoist would say that once she leaves her own point of view, there is nothing she ought or ought not to do. (I owe this objection to XXX.) But an egoist need not say this, and saying it seems odd. The egoist would, for example, have to deny that if I take up the point of view of being your friend, I ought to promote your good.

parts of our nature should be governed by the Higher,” “we ought to give every man his own,” and “live according to Nature” (ME 375-8). When formulating one axiom, he discards an attempt because “it does not clearly avoid tautology” (ME 381). He prefaces his axioms by wanting “self-evident moral principles of real significance” rather than “doctrines that lead us round in a circle” (ME 379). And after giving his axioms, he compares them to one of Euclid’s axioms, “if equals are added to equals the wholes are equal,” which he seems to find synthetic (ME 383; also 507).<sup>9</sup> But one might worry that *personal irrelevance* is really a sham. If the point of view of the universe just is the point of view from which the good of any one is no more important than the good of any other, *personal irrelevance* seems tautologous.

Sidgwick could make two replies.<sup>10</sup>

---

<sup>9</sup> Sidgwick writes that “geometry as commonly accepted does give us...synthetical universals: I know that all triangles inscribed in a semicircle must be right-angled triangles, and I could not obtain this knowledge by mere analysis of the notions of ‘semicircle’ and ‘right-angled triangle’” (LK 45). This does not, however, directly address the axioms. Elsewhere Sidgwick finds Kant guilty of a “hasty extension to Arithmetic of a view originally formed by a consideration of Geometry,” which can be seen in his “famous distinction between ‘analytical’ and ‘synthetical’ propositions” (CC 80). Sidgwick goes on to argue that arithmetic is analytic, perhaps implying that geometry is not.

<sup>10</sup> At one point, Sidgwick suggests that either *personal irrelevance* or *benevolence* is analytic: “that one ought not to prefer one’s own good to the greater good of another...is evolved immediately out of the notion of ‘good’...if this notion is used

First, *personal irrelevance* is best read as having a consequent that makes an “ought” claim: one ought to be indifferent between distributions that result in the same amount of good (though not between distributions that result in different amounts of good). The antecedent of *personal irrelevance* says that I have adopted the point of view of the universe. According to Sidgwick, no “ought” claim is entailed by strictly non-normative claims (ME 98, EP 107-8, B 412). If so, the consequent does not for Sidgwick follow from the antecedent. *Personal irrelevance* is not, then, tautologous in the way that, for example, “it is right to act rationally” is: “I have adopted the point of view of the universe” and “I ought to be indifferent between distributions that result in the same amount of good (though not between distributions that result in different amounts of good)” are not mere verbal equivalents, unlike (for Sidgwick) “right” and “rational” (ME 375).<sup>11</sup>

---

absolutely” (B 411). This cannot be his considered view, unless he takes up the suggestion made in the next note.

<sup>11</sup> This leaves open the status of *personal irrelevance*. (I owe this issue to an anonymous referee.) Sidgwick allows inobvious analyses. He suggests, for example, that “the uninstructed majority of mankind could not define a circle as a figure bounded by a line of which every point is equidistant from the centre: but nevertheless, when the definition is explained to them, they will accept it as expressing the perfect type of that notion of roundness which they have long had in their minds” (ME 353; also 229). *Personal irrelevance* might, then, be analytic, though not like “it is right to act rationally,” or it might be synthetic a priori. Given Sidgwick’s view on “is” and “ought,” I think he clearly favours the latter, but in the context of the tautology objection, this does not



This is not the usual reading of *personal irrelevance*. Many read *personal irrelevance* as making a point only about goodness: the goodness of some good does not vary with mere changes in who possesses the good.<sup>12</sup> But to say that the good of A is no

---

matter. One thing that the debate over “every cube has twelve edges” and the like made clear is that it is hard to tell whether some claims are inobvious analyses or synthetic a priori. The important distinction is between useful and useless claims.

<sup>12</sup> C. D. Broad, “Self and Others,” in Broad’s Critical Essays on Moral Philosophy, ed. David Cheney (London: George Allen & Unwin, 1971) p. 263; Schneewind, pp. 299, 301; Anthony Skelton, “Sidgwick’s Philosophical Intuitions,” Ethics and Politics 10 (2008): 185-209, 190. Broad thinks that an egoist might agree with *personal irrelevance*, even if “point of view of the universe” is cut. The egoist can agree that the good of any one is of no more importance than the good of any other, in the sense that, say, five units of pleasure for another really is more good than four units of pleasure for himself. On this suggestion, Sidgwick’s careful inclusion of “from the point of view of the universe” does no work. The axiom is consistent with egoism because the egoist can hold that there is no duty to produce more good rather than less. This simply moves the problem of passing the agreement test from *personal irrelevance* to *general good*. See Broad, Five Types pp. 243-5; “Certain Features in Moore’s Ethical Doctrines,” in The Philosophy of G. E. Moore, ed. P. A. Schilpp (Lasalle: Open Court, 1968 [3<sup>rd</sup> ed.]) p. 45; “Self and Others” pp. 263-5. Schneewind objects that Broad’s interpretation does not fit Sidgwick’s view of how the debate with the egoist runs. This debate turns on taking up the point of view of the universe, rather than on arguing for a duty to produce more good rather than less (Schneewind p.365n2). But since I shall argue that *general good* should

more important than the good of B seems to say that one ought to be indifferent between them. And to add “unless more good is produced by giving to A” is to add a point about how goods ought to be distributed. The “unless” clause could not be a condition for the truth of the claim that the goodness of some good does not vary merely with who possesses the good. That claim is true even if more good would be produced by giving the good to A. *Personal irrelevance*, then, claims more than that the goodness of some good does not vary with mere changes in who possesses it. It claims that I ought to be indifferent between equal distributions of good, though not between unequal ones.<sup>13</sup>

---

be understood as having “from the point of the view of the universe” as its antecedent, Broad would have an explanation.

<sup>13</sup> For other interpretations (though not explicitly of *personal irrelevance*) that ascribe ought-claims, see Hastings Rashdall, Ethics (London: T. C. and E. C. Jack, 1913) p. 62 and The Theory of Good and Evil (Oxford: Oxford University Press, 1907) v. i p. 185; F. H. Hayward, The Ethical Philosophy of Sidgwick (London: Swan Sonnenschein) p. 145. Rashdall gives an interpretation like Broad’s at Theory pp. 90-1, 147 and in Is Conscience an Emotion? (London: T. Fisher Unwin, 1914) pp. 41-2. At one point, he argues (not as an objection) that the best version of the axiom—“one man’s good is of as much intrinsic worth as the like good of another”—is analytic since “[t]hat which I recognize as having value in one man I must recognize as having the same value in another, provided it is really the same thing that is implied in the assertion that it has value” (Theory p. 147). But he does not ascribe this version to Sidgwick (Theory p. 147n; see note 24 below).

This interpretation, on which *personal irrelevance* makes an “ought” claim, is supported by Sidgwick’s treatment of the parallel axiom,

*Temporal irrelevance*: “the mere difference of priority and posteriority in time is not a reasonable ground for having more regard to the consciousness of one moment than to that of another.”

In alternate formulations and applications, Sidgwick says that “[h]ereafter as such is to be regarded neither more nor less than Now” or “a smaller present good is not to be preferred to a greater future good” or “present pleasure or happiness is reasonably to be foregone with the view of obtaining greater pleasure or happiness hereafter” (ME 381; also 124n1). *Temporal irrelevance* does not say only that the goodness of some good does not vary with mere temporal changes. It also says that one ought to ignore mere changes in temporal position.<sup>14</sup>

The second reply to the tautology charge turns on glossing “point of view of the universe.” When Sidgwick explains how it is reasonable for an egoist to avoid the

---

<sup>14</sup> This interpretation may also be supported by what may be an alternative statement of *personal irrelevance*: “one ought not to prefer one’s own good to the greater good of another...is [a principle] evolved immediately out of the notion of ‘good’ or ‘desirable,’ if this notion is used absolutely; as it then must mean ‘desirable from a universal point of view,’ or ‘what all rational beings, as such, ought to aim at realising’ ” (B 411). One might read this as claiming that from the point of view of the universe, I ought not to produce the smaller good, which is close to *personal irrelevance*. But if, as the final clause suggests, this involves thinking that I ought to bring about the good, the principle may be better identified with *benevolence*.

argument for utilitarianism—and so what it is not to take up the point of view of the universe—he writes that “the distinction between any one individual and any other is real and fundamental, and that consequently ‘I’ am concerned with the quality of my existence as an individual in a sense, fundamentally important, in which I am not concerned with the quality of the existence of other individuals” (ME 498). I take the point of view of the universe to be a point of view from which I have no special concern for myself. This seems to fit Sidgwick’s examples of taking up the point of view of the universe—one does so when one thinks that “the design of the creator of the world is to realise Good” or that “nature designed him to seek his own happiness” (B 411, ME 421). Neither the creator nor nature has special concern for me.<sup>15</sup>

On this reading, *personal irrelevance* is not tautologous. *Personal irrelevance* says not only that I should be indifferent to who has some good, but also that I should depart from indifference when more good would be produced by distributing to one person over another. Presumably this addition is also contingent on adopting the point of view of the universe. (It would be odd to think that the addition lies outside the scope of taking up the point of view of the universe, so that “if I adopt the point of view of the universe, then the good of one is of no more importance than the good of another” is itself conditional on it not being true that more good would be produced by distributing to one person rather than another.) But “I have no special concern for myself” and “I ought to depart from indifference when more good would be produced” are not mere verbal

---

<sup>15</sup> Similarly, Schneewind and Parfit take the point of view of the universe to be an impartial one (Schneewind, p. 299; Parfit, *On What Matters* (Oxford: Oxford University Press, 2011) v. 2, p. 447).

equivalents. However “point of view of the universe” is defined, it seems intended to exclude only egoistic considerations.

Even if one disregards Sidgwick’s “unless...there are special grounds for believing that more good is likely to be realised,” the same reply holds. “I have no special concern for myself” is not a mere verbal variation on “I ought to be indifferent between equal amounts of good.” As the objection considered in the next section indicates, there are possible grounds of preference between equal amounts of good that do not rest on a special concern for myself.

(The tautology worry can also be raised for *general good*, unconnected to adopting the point of view of the universe. One might worry that *general good* is a tautology on the ground that Sidgwick defines “good” as what one has reason to aim at (ME 112, 381, B 411, GSM 331).<sup>16</sup> Since Sidgwick has, one page before, worried that “‘one ought to aim at one’s own good’...does not clearly avoid tautology, since we may define ‘good’ as ‘what one ought to aim at,’” it is odd that he does not raise the worry here (ME 381). But the worry seems misguided in both cases. Sidgwick defines “ultimate good on the whole for me” as “what I should practically desire if my desires were in harmony with reason, assuming my own existence alone to be considered” (ME 112). “I ought to aim at my own good” becomes “I ought to aim at what I ought to desire assuming my own existence alone to be considered.” This is not a tautology. Similarly, Sidgwick defines “‘ultimate good on the whole,’ unqualified by reference to a particular subject” as “what as a rational being I should desire and seek to realize, assuming myself

---

<sup>16</sup> For the point (not presented as an objection), see Phillips, p. 124. Paul Bernays makes the same point as an objection (quoted in Schneewind, pp. 307-8).

to have an equal concern for all existence” (ME 112). Suppose, as is plausible, that “ultimate good on the whole, unqualified by reference to a particular subject” is the same as “general good.” “I ought to aim at the general good” becomes “I ought to aim at what I ought to desire assuming myself to have an equal concern for all existence,” which again is not tautologous.)

Even if the tautology worry is avoided, however, there remains the worry that Sidgwick does not affix “from the point of view of the universe” to *general good* or *benevolence*. But there is evidence indicating exactly this intention.<sup>17</sup>

In the second and third editions, where Sidgwick first introduces *general good*, he writes that “as rational beings we are manifestly bound to aim at good generally, not

---

<sup>17</sup> There is also evidence against it, since Sidgwick sometimes seems to treat *general good* as inconsistent with egoism—although given the agreement test, this must be sloppiness. In the second and third editions, he claims that one can deny aiming at good generally only by “denying that there is any such universal good,” which sounds just like what he supposes the egoist can say to avoid the argument for utilitarianism (ME (2) 355; ME (3) 381). He also contrasts “aiming at happiness generally” with “aiming at one’s own” (ME xii). Later, he writes of the ““sense of the ignobility of Egoism’... which... is the normal emotional concomitant or expression of the moral intuition that the Good of the whole is reasonably to be preferred to the Good of a part” (ME 500). “Intuition” suggests that Sidgwick is referring to one of his axioms. He might be referring to *personal irrelevance*, since that is introduced by noting parts and wholes (ME 382). Or he might (more naturally) be referring to *general good*, since that tells one to prefer good generally to a part of it.

merely at this or that part of it; we can only evade the conviction of this objectivity by denying that there is any such universal good” (ME (2) 355; ME (3) 381). The “we can only evade...,” dropped in later editions, makes being bound to aim at the general good conditional on thinking that “there is any such universal good.” This conceptual claim is dubious—an egoist seems capable of talking about the universal or general good, or of recognising that some states of other people are goods, while not being committed to thinking that he must aim at it or them. Perhaps this explains why Sidgwick cut the addition to *general good* after the third edition, as part of his revising away from talk of what concepts entail.<sup>18</sup> But *general good* could instead note a consequence of adopting the point of view of the universe. An egoist could then agree to it, but refuse to take up this point of view.<sup>19</sup>

---

<sup>18</sup> For evidence of this trend, see the Appendix.

<sup>19</sup> This is similar to Schneewind’s suggestion that *general good* can be read so that an egoist can accept it if it says only that I ought to aim at more good rather than less, where my point of view sets the relevant goods (363). Elsewhere, however, Schneewind gives a different weakened reading of *general good*: it says that “if it is the goodness of some state of affairs which is an agent’s reason for an action, then the goodness of a precisely similar state of affairs is also a reason for action, regardless of the ownership of what is good” (299-300; also 301, 307). *General good* says more than this; it says that there is a reason to promote good states of affairs (wherever they occur.) For a similar point about Schneewind’s treatment of the axioms, see Terence Irwin, The Development of Ethics (Oxford: Oxford University Press, 2009) v. 3, pp. 493-4.

This may be implicit in later editions. Sidgwick begins the axiom with “as a rational being I am bound....” “As a rational being” seems superfluous. But at one point Sidgwick treats “what all rational beings, as such, ought to aim at realising” as equivalent to “desirable from a universal point of view” (B 411).

There are two arguments for reading *benevolence* in the same way.

First, *personal irrelevance* contains as its antecedent “from the point of view of the universe.” I have argued that *general good* should be read as also having this antecedent. Sidgwick deduces *benevolence* from *personal irrelevance* and *general good*. One would think that *benevolence* should have the same antecedent.

Second, although Sidgwick almost always does not include this antecedent in *benevolence*, on occasion something very close to *benevolence* has it: “no doubt it was, from the point of view of the universe, reasonable to prefer the greater good to the lesser, even though the lesser good was the private happiness of the agent;” “I certainly could will it to be a universal law that men should act in such a way as to promote universal happiness; in fact it was the only law that it was perfectly clear to me that I could decisively will, from a universal point of view” (ME xx, xxii).<sup>20</sup>

I conclude that Sidgwick can get the egoist’s assent by making the axioms conditional on taking up the point of view of the universe.

---

<sup>20</sup> Parfit gives, in passing, the same interpretation (On What Matters, v. 2 p. 554). If the principle quoted earlier from “Mr. Barratt” is read as *benevolence*, this is another place in which “from the point of view of the universe” is attached to it (see note 14).



One might object that, since Sidgwick believes there is a dualism of practical reason, an interpretation on which egoists agree to the axioms cannot be right.<sup>21</sup> But, first, Sidgwick does not explicitly say that egoists reject any of the axioms;<sup>22</sup> there is no hint in the chapter on the axioms that anyone disagrees with them; and if Sidgwick did allow disagreement, his crucial claims for the superiority of his axioms over common sense morality (and their status as similar to mathematical axioms) would be false (ME

---

<sup>21</sup> Versions of this objection were raised by XXX, YYY, and an anonymous referee.

<sup>22</sup> He does come close for *general good*; see note 17. He also comes close in his sketch of the history of the Methods. There he claims that because of the conflict between egoism and utilitarianism, “I was forced to recognise the need of a fundamental ethical intuition” (ME xviii; also xxi). He seems to identify this intuition as “That a rational agent is bound to aim at Universal Happiness” (ME xxi). (This is inserted in square brackets into Sidgwick’s text, presumably by E. E. Constance Jones, but it seems a correct insertion, given Sidgwick’s reference there to his discussion of More and Clarke.) That intuition, unlike my version of *benevolence*, would conflict with egoism (provided Sidgwick is right that what brings about my happiness is not always what brings about universal happiness). But since Sidgwick also gives, in the same sketch, versions of *benevolence* that do seem to include the condition that one take up the point of view of the universe (ME xx, xxii, quoted above), perhaps he is best read as saying in the sketch that an intuition is needed to choose utilitarianism over egoism, although that intuition is not the unimpeached axiom (my *benevolence*) found in Methods III.XIII. One needs the intuition that one ought to take up the point of view of the universe—an intuition Sidgwick does not claim to pass his tests.

xxii, xxiii, 383, 421, 496, 507, PC 564). What is explicit is that the egoist avoids utilitarianism by refusing to say that “his happiness...is Good, not only for him but from the point of view of the Universe” (ME 420). The egoist “may avoid the proof of Utilitarianism by declining to affirm” that “his own greatest happiness is... a part of Universal Good” (ME 497-8).

Second, there is a gap between the positions which constitute the dualism—egoism and utilitarianism—and the axioms.<sup>23</sup> The relevant axioms include the condition that one takes up the point of view of the universe; they state only one set of duties; and they do not specify the good. (For further gaps, see section III.) This allows for the possibility that all agree to the axioms and there is a dualism between egoism and utilitarianism. Admittedly, Sidgwick is misleading when, in stating the dualism, he writes that we must “admit an ultimate and fundamental contradiction in our apparent intuitions of what is Reasonable in conduct” and that “the apparently intuitive operation of the Practical Reason, manifested in these contradictory judgments, is after all illusory” (ME 508). If *benevolence* includes the condition that one takes up the point of view of the universe, there is no contradiction between it and any egoist axiom (if there is one). But even if one thought that there were unconditional consequentialist and egoist axioms, Sidgwick is misleading here, since even then there is no contradiction—in the sense of guidance toward two actions such that if one is performed, the other cannot be—without adding to the axioms some specification of the good. On my view, similarly, conflict occurs only when more is added, and in particular when it is added that someone does, and another does not, take up the point of view of the universe, and that either choice is

---

<sup>23</sup> See the references to Schneewind and Shaver in note 7.

reasonable. This could be described as a contradiction in our judgments about “what is reasonable in conduct,” and as a failure of practical reason—so perhaps Sidgwick is not so misleading. In any case, his pessimistic conclusion remains.

## II

*Personal irrelevance* says that when I have no special concern for myself, I ought to be indifferent between distributing x to A and x to B, unless more good would be produced by distributing to one of us. The common sense moralist might object that there are reasons for departing from indifference that do not depend on special concern for myself or on producing more good. Perhaps I ought to distribute to A, who is virtuous, rather than to B, who is vicious, even if no more good would be produced. Perhaps I should distribute to A, to whom I have promised the item, or to whom I owe gratitude, rather than to B, even if no more good is produced. *Personal irrelevance* says that I ought to be indifferent between A and B. The same points can be made for *benevolence*.<sup>24</sup>

---

<sup>24</sup> Rashdall raises a related objection, concentrating on cases in which A can attain a “higher” good than B and cases of differences in merit (Theory v. i pp. 234-242, 262-3). He thinks the axiom should read “Every man’s good to count as equal to the like good of every other man,” to allow for distributing more to one with the capacity for higher goods (Theory v. i p. 240). He seems to omit the “unless,” at one point interpreting Sidgwick as holding that “I ought always, so far as possible, to produce an equal distribution of good among all the people who can be affected by my action” (Conscience p. 42).

One could reply that more good comes of distributing things to A. Perhaps (for example) distributing to A encourages virtue and so produces more goods such as virtue and happiness in the long run. Perhaps the deserving getting what they deserve is itself a good. But some will still disagree: I might say that I ought to distribute to A, without thinking that rewarding A produces more good than not doing so; I might think (with Prichard and Carritt) that there is a duty of justice that is independent of any duty to produce more good. The standard reply is that this position has no clear advantage over one which admits goods such as distribution according to virtue. But this is rejected by those who think that, say, I ought to give the deserving what they deserve even when the result is that five other people fail to give the deserving what they deserve.

Considerations of goodness favour my not distributing according to desert here.

There is a different reply to the objection, one which relies more on Sidgwick's text. *Personal irrelevance* claims that one of my duties is to be indifferent between distributing to myself and distributing to another, unless more good would be produced by distributing to one of us. As far as *personal irrelevance* goes, I may have other duties, such as to distribute according to desert or gratitude. (Again, the same reply can be made for *benevolence*. Henceforth I omit this rider.)

This interpretation can be supported in three ways.

First, Sidgwick thinks that the argument from the axioms "only shows the Utilitarian first principle to be one moral axiom: it does not prove that it is sole or supreme. The premises from which the Intuitionist starts commonly include other

formulae held as independent and self-evident” (ME 421).<sup>25</sup> This suggests that *personal irrelevance* is not intended to rule out all reasons for departing from indifference other than the amount of good produced. If it did, it by itself would imply that many other principles are false. One could not, for example, hold that although giving the resource to A produces less good than giving it to B, it should be given to A because A deserves or has been promised it. That Sidgwick does not consider this sort of objection to *personal irrelevance* is evidence that he did not intend it to rule out all reasons for departing from indifference other than bringing about more good.

Similarly, Sidgwick notes elsewhere that common sense moralists agree at least with *benevolence*.

No Intuitionist ever maintained that all our conduct can be ordered rightly without any calculation of its effects on human happiness. On the contrary, this calculation...is expressly inculcated by the maxims of Prudence and Benevolence, as commonly understood (PC 564).

---

<sup>25</sup> One might object that Sidgwick writes that “as addressed to the Intuitionist, this reasoning only shows....” Perhaps Sidgwick does suppose he has shown that there are no reasons for departing from indifference other than the amount of good produced. He grants here that there might be other duties only in the course of an *ad hominem* argument addressed to the intuitionist. (I owe this objection to XXX.) But—putting aside the vexed issue of whether Sidgwick’s argument is just *ad hominem*—in the texts given below, where there is no suggestion of a concession made just for the sake of argument, Sidgwick again claims that the common sense moralist would agree with *benevolence*.

The judgment of good and ill is...ethical. And this is admitted by Common Sense. In the recognition of the virtues of Prudence and Benevolence it is allowed to be (1) right...that a man should prefer greater to lesser good in his own case, and (2) virtuous to promote good generally for others....And the only difference between Common Sense and Utilitarianism as regards the outward act is, that in the view of Common Sense these rules are but rules among others, while in the view of Utilitarianism, the rule of promoting the greatest good is supreme (GSM 331-2).

Second, Sidgwick claims that Clarke and Kant agree with his axioms (ME 384-6). This may be true if *personal irrelevance* states one duty. It is false (at least for Kant) if *personal irrelevance* rules out all reasons for departing from indifference other than the amount of good.<sup>26</sup>

Third, in the first edition discussion of (what becomes) *personal irrelevance*, Sidgwick does not say that the only reason for departing from indifference is that more good would be produced. The emphasis is on denying egoism. A “rational being...cannot regard the satisfaction of his own personal desires as intrinsically more desirable (as being his own) than the satisfaction of the desires of any other person” (ME (1) 363; also (1) 360, (1) 364). The chapter concludes by noting that “we have only evolved the suppression of Egoism” (ME (1) 366). In later editions (and elsewhere),

---

<sup>26</sup> Irwin makes this point, though as a criticism of Sidgwick (pp. 499, 501, 504-5, 510, 512-13, 516, 518).

when Sidgwick glosses what is self-evident, the gloss similarly does not make the claim that the only reason for departing from indifference is that more good would be produced:

I ought not to prefer my own lesser good to the greater good of another (ME 383; also PC 565, B 411, FC 483).

[T]he good...of any one individual must as such be an object of rational aim to any other reasonable individual no less than his own similar good (ME 385).

[One is to seek] others' good no less than one's own (ME 392).

Universal Benevolence...sets before each man the happiness of all others as an object of pursuit no less worthy than his own (ME 496; also UG 31).

I ought to sacrifice my own happiness, if by doing so I can increase the happiness of others to a greater extent than I diminish my own (FC 483).

[M]y own good is no more to be regarded than the good of another (FC 484).

Sidgwick glosses the axioms chapter as giving an argument “addressed to Egoism” (ME 420).

Why, then, does Sidgwick include the “unless” clause, given that he does not intend *personal irrelevance* to rule out departing from indifference on grounds such as desert or fidelity?<sup>27</sup> I think he is working toward stating the main duty concerned with

---

<sup>27</sup> Some commentators, such as Schneewind, simply leave off the “unless” clause (294, 296). When he first quotes *personal irrelevance*, he adds—not as part of the axiom—that “[a]s in the case of [temporal irrelevance], quantitative differences...are discounted here” (294). Irwin (see below), A. R. Lacey (“Sidgwick’s Ethical Maxims,” Philosophy 34

promoting the good. He is denying, in one way, a conclusion one might be tempted to draw from the claim of equal importance, just as he denies a conclusion one might be tempted to draw from *temporal irrelevance*. There he writes that “[i]t is not, of course, meant that the good of the present may not reasonably be preferred to that of the future on account of its greater certainty: or again, that a week ten years hence may not be more important to us than a week now, through an increase in our means or capacities of happiness” (ME 381).<sup>28</sup>

---

(1959): 217-228, 218) and Skelton (“Intuitions,” 187) include the “unless.” Phillips (121), James Seth (“The Ethical System of Henry Sidgwick,” *Mind* 10 (1901): 172-187, 182), Broad (“Self and Others” p. 262), Parfit (*On What Matters* v. 2, p. 447), Rashdall (see note 13) and James D. Wallace (*Moral Relevance and Moral Conflict* (Ithaca: Cornell University Press, 1988) p. 42) do not. A similar split occurs on the “except” clause in *benevolence*.

<sup>28</sup> Irwin notes the importance of the “unless” addition (498-9). (He seems to be the only one to do so.) He treats it as a momentous shift between the first edition and later ones, a shift to an argument that takes maximising the good to be the fundamental axiom. (Skelton gives some textual evidence for thinking Sidgwick finds maximising self-evident; see “Intuitions” 188, 193-4.) Irwin objects that Sidgwick could not expect agreement to taking maximising the good to be the fundamental axiom (504, 510, 512-13), and gives various textual objections to doing so (504, 507-8). This, along with the textual evidence I give (from editions other than the first), favours a reading on which *personal irrelevance* is just one duty. Irwin may want his reading because he supposes that Sidgwick must prove consequentialism directly from the axioms, rather than from a



Sidgwick is not denying that there might be other reasons for not being indifferent. These would be found in other duties. Perhaps common sense thinks I ought to favour (say) my parents, when that does not bring about the greatest good, but it does so only in virtue of duties other than *benevolence*, such as gratitude (ME 248, 259).<sup>29</sup> In the case of duties of parents to children, Sidgwick suggests (and rejects) a “special

---

combination of the axioms and an attack on any alternative axioms. He objects to a similar combination suggestion by Schneewind (307, 308) that “[f]ailure to find any other universal right-making property does not show that maximizing utility is a viable candidate; perhaps we ought to conclude that we cannot find any single right-making property” (498n30). But this objection is avoided if a case is first made for maximizing utility, and then a case is made against any other property (see III below).

<sup>29</sup> Ross writes that it is on the prima facie duty of gratitude that “our special duty to parents and friends in the main depends” (W. D. Ross, Foundations of Ethics (Oxford: Clarendon, 1939) p. 76). He links the duty of parents to children to promising (Foundations, p. 273). Broad may think that special relationships are sufficient to ground duties, without reference to further duties. He writes that common sense “holds that...special relationships are the ultimate and sufficient ground for...specially urgent claims on one’s beneficence” (“Self and Others” p. 280). “[P]rima facie the special urgency of the claims of certain others upon one’s beneficence seem to be founded directly on certain special relationships of those others to oneself” (“Self and Others” p. 269; also 265). Sidgwick may disagree with Broad; this would explain why he so brusquely finds aiming at the happiness of one’s family, as an ultimate end, arbitrary (ME 10). But even if Broad is correct, this does not affect agreement on *benevolence*.

duty...aris[ing] from the fact that I have brought them into being” (ME 347).

*Benevolence* does not rule out these duties.

One might object that there is no such duty to promote the good. Consider cases familiar from the anti-utilitarian literature, such as Scanlon’s World Cup case: A is trapped in a TV transmission unit, suffering painful shocks; A can be removed, but only by stopping transmission for fifteen minutes; since a World Cup match is on, with millions of viewers, more good would be produced by leaving A in until the game ends in an hour; doing so is clearly wrong.<sup>30</sup>

Sidgwick can reply, however, that *personal irrelevance* does not say that favouring the viewers is my duty all things considered. *Personal irrelevance* says that it is one duty, one thing I have reason to do. This is hard to deny. Even if the loss of viewing has little significance compared to the pain of the shocks, the loss of viewing still is a loss, something there is some reason to avoid. It is much more plausible to think that the reason to allow continued viewing is outweighed than it is to think that, although the viewing is a good, it is a good there is no reason to bring about (especially, but not only, on a view of the good as what there is reason to bring about).

Sidgwick could note that his later critics, speaking on behalf of common sense morality, agree. Ross writes that “it seems self-evident that if there are things that are intrinsically good, it is prima facie a duty to bring them into existence rather than not to do so, and to bring as much of them into existence as possible.” To say that “there is no

---

<sup>30</sup> Thomas Scanlon, What We Owe to Each Other (Cambridge, MA: Harvard University Press, 1998) p. 235.

prima facie duty to produce as much that is good as we can” is “plainly false.”<sup>31</sup> Broad writes that “on any view, one important prima facie duty is to produce and conserve good and to avert and diminish evil.”<sup>32</sup>

One might instead object that “the duty of Benevolence as recognised by common sense seems to fall somewhat short” of what *benevolence* requires. Common sense favours aiming at “the good of a limited number of human beings... generally in proportion to the closeness of their connexion with him” (ME 382; also 431). Sidgwick also notes that one might worry that the ordinary duty of benevolence does not rule out my sacrificing more than another gains, and that it requires only a small sacrifice for a large gain (ME 252-3, 348-9, 431-2). One might add that perhaps benevolence is better captured by giving priority to the worse-off.

These worries are best answered by claiming that the common sense moralist accepts *benevolence* as one duty—as Ross and Broad do—whether or not *benevolence* should be described as the duty of benevolence. But I admit that this is not what Sidgwick says. In reply to the impartiality worry, Sidgwick writes that “practically each man, even with a view to universal Good, ought chiefly to concern himself with promoting the good of a limited number of human beings, and that generally in proportion to the closeness of their connexion with him” (ME 382; also ME 327, 431-9, UG 31-2). He then gives

---

<sup>31</sup> W. D. Ross, The Right and the Good (Oxford: Clarendon, 1930) pp. 24, 25; also 27, 39, 58, Foundations pp. 271, 273.

<sup>32</sup> Broad, “Self and Others” p. 264; also Five Types p. 221.

*Plain man*: “a ‘plain man’...if his conscience were fairly brought to consider the hypothetical question, whether it would be morally right for him to seek his own happiness on any occasion if it involved a certain sacrifice of the greater happiness of some other human being,—without any counterbalancing gain to any one else,—would answer unhesitatingly in the negative” (ME 382)

to confirm that when these practical considerations are absent, common sense agrees with *benevolence*.

If Sidgwick’s point is that the prevalence of partial duties does not show the absence of an impartial duty, he is correct, and backing up what I take to be the best answer. But if his point is that the plain man accepts *benevolence* as an all things considered duty, Sidgwick is saying something both unnecessary to his argument and probably false.<sup>33</sup> Even if *plain man* is true, it does not follow that common sense forbids me to favour the smaller good of, say, my children or my parents, over the greater good of some strangers. Sidgwick makes his work easier by choosing oneself as the part.<sup>34</sup>

---

<sup>33</sup> Parfit thinks *plain man* is “simply false” (On What Matters v. 1, p. 453).

<sup>34</sup> Thanks to XXX for this point. Sidgwick may cheat in a further way. The plain man is asked whether it is “morally right” to refuse to sacrifice his good for the sake of the greater good of another. (Similarly, *benevolence* concerns what one is “morally bound” to do.) One might agree with *plain man* when thinking of morality as by definition ruling out special concern for oneself. But this is not Sidgwick’s view: although he expresses doubt about whether egoism counts as a moral theory, he does not make anything hang on this taxonomic point (ME 119). His concern is with what it is most reasonable to do.

I have argued that *benevolence* states just one duty. One might try to give Sidgwick a more ambitious view. There are arguments one could give for moving from adopting the point of view of the universe, understood in a particular way, to being concerned only with amounts of good, all things considered.

First, in “Mr. Barratt on ‘The Suppression of Egoism,’” Sidgwick writes that “one ought not to prefer one’s own good to the greater good of another...is [a principle] evolved immediately out of the notion of ‘good’ or ‘desirable,’ if this notion is used absolutely, as it then must mean ‘desirable from a universal point of view,’ or ‘what all rational beings, as such, ought to aim at realising’” (B 411). One might gloss “desirable from a universal point of view” as requiring that everyone desire some one state of affairs. This admits only agent-neutral reasons. One might then argue that all reasons other than a concern with the amount of good are agent-relative. For example, deontological constraints add reasons in addition to the amount of good, but many have argued that they are agent-relative. I see no evidence for this in Sidgwick, however, and thinking of deontological constraints as agent-relative is controversial.<sup>35</sup>

Second, the same passage from “Mr. Barratt” identifies a “universal point of view” with what “all rational beings as such, ought to aim at realising.” *Personal irrelevance* thus claims that qua rational being I ought to be indifferent between different distributions that produce the same amount of good, and more generally, that only amounts of good are relevant to judging between distributions. All I know about myself when I consider myself merely as a rational being is that I act on what I have most reason

---

<sup>35</sup> For a sample discussion, see F. M. Kamm, *Morality, Mortality* (Oxford: Oxford University Press, 1996) v. 2 ch. 9.

to do. The good is what it is reasonable to aim at (ME 112, 381, GSM 331). So I should be concerned only with amounts of good.

If this is Sidgwick's thought, there is a familiar objection. He would be "consequentialising"—the good expands to cover anything one ought to aim at, so there are no considerations other than the amount of good. But this simply relocates the debate between (traditional) consequentialists and their opponents. And Sidgwick himself seems to deny this taxonomy. When he repeats his opening view that moral theories take either the right or the good as fundamental, he notes that taking the right as fundamental has been defeated because he has shown that most duties "contain an implicit subordination" to axioms concerning the good and that only the axioms I have discussed, all of which but *consistency* concern the good, pass his tests (ME 391). The consequentialising tactic would give a much quicker victory.

Putting aside these more ambitious readings, Sidgwick's argument for *benevolence* goes as follows:

(1) From a point of view from which I have no special concern for myself, one thing I ought to do is be indifferent between distributing  $x$  to A and  $x$  to B, unless I could produce more good by one distribution or the other (*personal irrelevance*).

(2) From a point of view from which I have no special concern for myself, one thing I ought to do is aim at the good of everyone (*general good*).

Therefore (3) From a point of view from which I have no special concern for myself, one thing I ought to do is aim at any other's good as much as at my own, unless I could produce more good by one distribution or the other (*benevolence*).

This reconstruction gives *personal irrelevance* and *general good* separate and necessary roles. *Personal irrelevance* claims that I ought to be indifferent between equal distributions of good to myself and A, as well as between equal distributions to A and B. It also says that, when more good can be produced by not being indifferent, I need not be indifferent. But it does not tell me that I ought to aim at the good of others. *General good* tells me to aim at the good of others, excluding no one, but does not forbid giving someone or some group special weight. (One might think that this is ruled out by adopting a point of view from which I have no special concern for myself. But I might well favour those close to me without special concern for myself.)

One might object that my version of *benevolence* is far weaker than Sidgwick intends.<sup>36</sup> In the next section, I note the many things Sidgwick needs to argue for to move from *benevolence* to the consequentialism he wants. Most obviously, he needs to argue in favour of taking up a point of view from which I have no special concern for myself, in favour of maximising, and in favour of thinking that *benevolence* is the only duty. But—so the objection goes—Sidgwick claims that he has established his consequentialism very soon after stating the axioms. He writes that “I find that I arrive, in my search for really clear and certain ethical intuitions, at the fundamental principle of utilitarianism” (ME 387). “Utilitarianism is thus presented as the final form into which Intuitionism tends to pass, when the demand for really self-evident first principles is rigorously pressed” (ME 388). The “Intuitionist method rigorously applied yields as its final result...Utilitarianism” (ME 406-7). Elsewhere he writes that “considerations of

---

<sup>36</sup> I owe this objection to XXX and an anonymous referee.

intuitionism in its most philosophical form... would have led me at once to utilitarianism” (PC 564).

These claims cannot, however, claim so much. Again, Sidgwick cannot say that consequentialism passes the agreement test. He has just noted that Clarke and Kant agree with his axioms (ME 384-6); if he thinks he has established consequentialism, he cannot even claim the agreement of those he has just cited as agreeing. And again, Sidgwick later notes that the argument from the axioms “only shows the Utilitarian first principle to be one moral axiom: it does not prove that it is sole or supreme” (ME 421). Fortunately, there is some indication that Sidgwick is aware that he has established something more modest. Directly before saying that he has arrived at “the fundamental principle of utilitarianism,” he writes that “the axiom of Rational Benevolence is... required as a rational basis for the Utilitarian system.” This is then glossed as “[a]ccordingly, I find that I arrive...” (ME 387). *Benevolence*, on my reading, is plausibly required as a basis for utilitarianism. Later Sidgwick glosses his argument as showing that “such abstract moral principles as we can admit to be really self-evident... seem required to furnish a rational basis for such a [utilitarian] system” (ME 496; also ME xxii-xxiii).

Utilitarianism “rest[s]” on the intuitions (UG 31; also ME xxi). It has *benevolence* as its “first principle” (ME 505). As for the claim that Intuitionism “tends to pass” into utilitarianism, the “tends” suggests that there is some slack between accepting the axioms and adopting utilitarianism, just as my interpretation predicts. (Perhaps, with more strain, “yields as its final result” also permits some slack.) And the claim that consideration of Intuitionism “at once” leads to utilitarianism is followed by this explanation: “because I hold that the only moral intuitions which sound philosophy can accept as ultimately valid



are those which...provide the only possible philosophical basis of the Utilitarian creed” (PC 564).

### III

To reach the maximising, agent-neutral consequentialism Sidgwick wants from *benevolence*, he has at least these tasks.

(i) Sidgwick’s strategy is to show that the common sense moralist must agree to *benevolence*, as the utilitarian does, and then argue that common sense morality contains nothing on a par with *benevolence* (ME 421).<sup>37</sup> If so, he needs to drop “from the point of view from which I have no special concern for myself” from *benevolence*. For utilitarians believe not just that if I have no special concern for myself, I ought to aim at any other’s good as much as at my own (etc.). They believe I have this duty, period.

I think Sidgwick assumes that the common sense moralist adopts the point of view from which one has no special concern for oneself. He may claim more than this—that this is on reflection the only point of view common sense adopts. But provided the common sense moralist always takes this point of view to be relevant, she agrees that one thing I ought to aim at is any other’s good as much as my own. Even in cases in which she thinks this “ought” is outweighed by, say, a claim that I ought to favour my family

---

<sup>37</sup> This fits Sidgwick’s account of his thought: “I found the axiom I required for my Utilitarianism...[But] why not further? The orthodox moralists such as Whewell...said there was a whole intelligible system of intuitions...But the result of the examination [of common sense morality] was to bring out with fresh force and vividness the difference between the maxims of Common Sense Morality...and the intuitions which I had already attained” (ME xxi-xxii).

over the greater good of a stranger, she thinks there is something to be outweighed.

Critics often argue that Sidgwick slides from the claim that benevolence is one duty to the claim that benevolence is the only duty. But if Sidgwick (here) only wants benevolence as one duty, and the critics agree with that claim, there is no objection.<sup>38</sup>

(ii) *Benevolence* says that if more good would be produced by distributing to A rather than to myself, I am not required to aim at A's good as much as at my own. It does not say that when more good would be produced by distributing to A, I ought to distribute to A. Presumably Sidgwick takes that to be understood.

(iii) *Benevolence* concerns myself and another, but *personal irrelevance*, from which *benevolence* is deduced, concerns anyone. Presumably Sidgwick intends the wider application. Although *benevolence* notes only the egoist case, the objection he considers to *benevolence* is that common sense thinks one ought to favour friends and family (ME 382). This understands *benevolence* as more than anti-egoist.

(iv) Sidgwick must argue for maximising. *Benevolence* says that my attitude toward another's good must be the same as my attitude toward my own good. If I think I

---

<sup>38</sup> For example, Broad writes that "on any view, one important prima facie duty is to produce and conserve good and to avert and diminish evil," but reads *general good* as saying that the "only legitimate ground for devoting more of one's beneficent activities to one person...rather than to another...is that by doing so one will promote more good or avert more evil" ("Self and Others" p. 264). For complaints like Broad's, see Skelton, "Utilitarian Practical Ethics," in Proceedings of the Second World Congress on Henry Sidgwick, ed. Placido Bucolo, Roger Crisp, and Bart Schultz (Catania: Cooperativa Universitaria Editrice Catanese di Magistero, 2011) p. 608 and Phillips, pp. 125-6.

ought to maximise my own good, maximising is the correct attitude toward the good of each person. Sidgwick assumes, I think, that maximisation is the correct attitude to take toward one's own good. He does not consider satisficing, or, say, preferring slightly less total good with the best parts toward the end to slightly more total good with the best parts toward the start.<sup>39</sup>

Sidgwick does not, however, understand "I ought to aim at any other's good as much as at my own" as telling me to maximise the good of each person in the sense that I ought to put each person at a level of good such that there is no possible world in which he or she has more good. In most cases, doing so is impossible. (Say I can save A or save B and C.) This impossibility is what makes it often false that I ought to maximise the good of each person in this sense—not, as *benevolence* claims, that I can produce more good by not maximising the good of each. Indeed, if I can maximise the good of each in this sense, producing more good by not doing so is impossible (for a fixed population). I think, then, that Sidgwick understands "I ought to aim at any other's good as much as at my own" as telling me to maximise the equal good of everyone.<sup>40</sup>

It is worth stressing the role of equality. Later Sidgwick claims that when equal amounts of good could be produced, one should favour an equal distribution over an unequal one (ME 417). Critics object that there is no ground in maximisation or utilitarianism for favouring an equal distribution, and that there is no reason to think that

---

<sup>39</sup> For the latter, see, for example, Broad, *Five Types* pp. 225-6; Irwin pp. 510-12.

<sup>40</sup> Irwin suggests this as what the first edition supports (495). Gerald Postema has interpreted Bentham as endorsing a version of this principle. See, for example, "Bentham's Equality-Sensitive Utilitarianism," *Utilitas* 10 (1998): 144-58.

an independent principle of equal distribution need always be subordinate to maximising the good.<sup>41</sup> But the equality claim follows from *benevolence*: I can depart from maximising the equal good of each only if a greater total amount of good can be produced by doing so. Sidgwick's utilitarianism has a concern for equal distribution built-in. Structurally, it resembles Rawls's difference principle, although the condition for departing from equality is of course different.<sup>42</sup>

(v) Sidgwick must argue that one thing I ought to do is not to maximise the equal good of each when I could produce more good by not doing so. The argument for this parallels the treatment of the World Cup case.

(vi) Most importantly, Sidgwick must argue that there are no other axioms that pass the tests. This is the job of Methods III.xi.<sup>43</sup>

---

<sup>41</sup> Ross, Foundations pp. 71-2 (though he is willing to accept a consequentialist reply, making certain distributions themselves goods, on 73-4); Wallace pp. 45-8; Irwin pp. 513-16.

<sup>42</sup> Thanks to XXX for pointing out the structural similarity.

<sup>43</sup> After stating some of the axioms, Rashdall gives this reconstruction of the plan of the Methods: "(1) That there are the two above-mentioned formal intuitive principles of Morality; (2) That there are no others: hence the necessity for the elaborate examination of 'Common Sense' Morality" ("Professor Sidgwick's Utilitarianism," Mind (o.s.) 10 (1885): 200-226, 202). Rashdall does not specify whether eliminating other principles is needed because the two principles are, in effect, *prima facie* duties, or because they state duties all things considered that must be shown to meet the agreement test. Since Rashdall does not mention agreement, the former is more likely.

Many have charged that Sidgwick is unfair: he holds other proposed axioms to a standard higher than that he applies to his own.<sup>44</sup> My reconstruction makes this charge clearly justified—in some cases. Sidgwick cannot complain that, say, the axiom “One ought to keep one’s promises” is not clear and precise because it fails to tell me what to

---

<sup>44</sup> For discussion of the unfairness charge, see Roger Crisp, “Sidgwick and the Boundaries of Intuitionism,” in Ethical Intuitionism, ed. Philip Stratton-Lake (Oxford: Clarendon, 2002) pp. 74-5; Skelton “Intuitions” 205-7 and “Utilitarian Practical Ethics” p. 608; Phillips ch. 4; Hallvard Lillehammer, “Methods of Ethics and the Descent of Man: Darwin and Sidgwick on Ethics and Evolution,” Biology and Philosophy 25 (2010): 361-78, 371; Shaver, “Sidgwick’s Axioms and Common Sense Morality,” in Proceedings of the Second World Congress. Earlier commentators were more impressed by the negative examination of common sense morality. Hayward writes that “Sidgwick’s discussion of this “common sense” doctrine is admitted by all critics to be extremely able, to be, in fact, the most irrefutable part of the book. The student... will thus come to see the weakness of popular intuitionism, and the necessity for a sounder ethical theory.... Most of the third book requires no commentary whatever; by common consent its chapters are so lucid, and the conclusions they embody are, for the most part, so indisputable, that even to point out their merits would be to gild refined gold” (viii, xviii). Rashdall writes that the “strongest part of Sidgwick’s great work consists in its analysis of common-sense Morality. The loose statements of Intuitionists as to the clearness, certainty, adequacy, and self-evidence of the ordinarily received rules of conduct have never been subjected to so searching, so exhaustive, and so illuminating an examination. That task has been done once for all” (Theory v. i p. 83).

do when keeping a promise conflicts with some other duty. *Benevolence* fails in the same way. Nor can Sidgwick argue that “One ought to keep one’s promises” fails the agreement test because there are cases in which, given other duties, we think we ought to break the promise. The common sense moralist can say that there is still a reason to keep the promise, just as Sidgwick can reply to standard objections to consequentialism that there is still a reason to bring about the most good.

Reviewing Sidgwick’s examination of common sense morality is a task for a paper devoted to the unfairness charge. Here I give some examples—some good for Sidgwick, some not so much.

Sidgwick notes that there is controversy about lying to invalids (ME 316, 355). The common sense moralist could reply that there is always a duty to tell invalids the truth, though it can be outweighed by a duty of benevolence. But Sidgwick also worries that it is not clear whether “it is our actual affirmation as understood by the recipient which we are bound to make correspondent with fact...or whatever inferences we foresee that he is likely to draw from this, or both” (ME 355; see 316-17). This seems to be a problem internal to the duty. But it might be met by copying the strategy used to defend *benevolence*. One might claim that according to common sense, there is always a reason to say what is true and always a reason to make the inferences the recipient draws from what we say result in truths.<sup>45</sup> There is not one duty with an internal problem, but rather two unproblematic duties.

---

<sup>45</sup> Broad suggests that it is fitting to tell the truth, but also fitting to avoid false beliefs, so one’s duty is determined by whichever is most fitting (*Five Types* p. 219). Ross instead

Sidgwick's treatment of "one ought to distribute according to desert" does better (where this is understood as a rival to producing the good rather than an instance of doing so). Sidgwick objects that

the question...recurs: whether...we are to apportion the reward to the effort made, or to the results attained. For it may be said that the actual utility of any service must depend much upon favourable circumstances and fortunate accidents, not due to any desert of the agent: or again, may be due to powers and skills which were connate, or have been developed by favourable conditions of life, or by good education, and why should we reward him for these?...[Going to moral excellences does not help, since] it may still be said that good actions are due entirely, or to a great extent, to good dispositions and habits, and that these are partly inherited and partly due to the care of parents and teachers; so that in rewarding these we are rewarding the results of natural and accidental advantages, and it is unreasonable to distinguish these from others, such as skill and knowledge, and to say that it is even ideally just to reward the one and not the other. Shall we say, then, that the reward should be proportionate to the amount of voluntary effort for a good end? But Determinists will say that even this is ultimately the effect of causes extraneous to the man's self...[T]here seems to be no justice in making A happier than B, merely because circumstances beyond his control have first made him better...[I]n any case it does not seem possible to separate in practice that part

---

argues that veracity requires saying what is true, though utility may favour avoiding inferences to false beliefs (Foundations pp. 81-2).

of a man's achievement which is due strictly to his free choice from that part which is due to the original gift of nature and to favouring circumstances (ME 283-5).<sup>46</sup> These objections do not rely on noting conflict with other duties. Nor can one plausibly retreat to thinking that there remains a reason for distributing according to desert. The objections note a problem internal to the duty.

In the case of promising, some of Sidgwick's worries can be seen as worries for thinking that there is a reason to keep a promise. He notes that it is controversial whether one is bound by a promise one makes given that a false statement, or suggestion of one, or concealment, or a falsehood shared by the promiser and promisee, was one consideration in favour of making it. It is also controversial whether changed circumstances, or changed beliefs about them, release the promiser. As a special case of this, Sidgwick considers that, after making a promise, I come to see that my keeping it will injure the promisee; the promisee, falsely believing that keeping the promise is harmless, refuses to release me from the promise. Suppose I come to see that the food I have promised is poisoned (ME 306-8, 353-4). In these cases, many will not think that there is a reason for keeping, with the various complications weakening this reason or suggesting an outweighing duty. Many will think there is no reason.

Ross agrees. When Ross considers various cases of changed circumstances and falsehood, he claims there is no prima facie duty, since the prima facie duty is conditional on various conditions that are now not met.<sup>47</sup> He does not, then, hold the simple view

---

<sup>46</sup> This appears in the first edition of the Methods, before Mill's similar argument was published in Chapters on Socialism but after Chapters was written.

<sup>47</sup> Ross, Foundations pp. 94-97, 110-11.



that there is always a prima facie duty to keep promises. In one case of changed circumstances—a promise to pay a pauper £100, with the pauper getting rich before the payment is due—he thinks one might say either that one is bound, or that one is not bound because an implicit condition was not met.<sup>48</sup> When he notes a case in which keeping the promise turns out to be unexpectedly costly to either the promiser or the promisee (as in the poison case), he does think one is bound, though the duty may be outweighed.<sup>49</sup> But elsewhere he argues that for me to be bound, I must think that my keeping of the promise produces some good to the promisee.<sup>50</sup> Of course it is tricky to distinguish between a weak reason and no reason.<sup>51</sup> Ross suggests that feeling “compunction,” and that “it is our duty to make up somehow to the promisee for the breaking of the promise,” is a sign of an outweighed, rather than absent, reason.<sup>52</sup> But I think these tests leave it controversial whether there is a reason to keep the poison promise, just as it is controversial whether there is a reason to keep the promise to the pauper.

---

<sup>48</sup> Ross, Foundations pp. 97-8.

<sup>49</sup> Ross, Foundations pp. 109, 111-12.

<sup>50</sup> Ross, Right p. 162; “The Nature of Morally Good Action,” Proceedings of the Aristotelian Society 29 (1928-9): 251-74, 267-8.

<sup>51</sup> There are brief discussions of this issue in the literature on particularism. See, for example, Russ Shafer-Landau, “Moral Rules,” Ethics 107 (1997): 584-611, 590, 592-4 and Brad Hooker, “Moral Particularism: Wrong and Bad,” in Moral Particularism (Oxford: Clarendon, 2000), ed. Hooker and Margaret Olivia Little, pp. 7-11.

<sup>52</sup> Ross, Right p. 28.

Ross might insist that there is an unproblematic prima facie duty to keep some promises. Sidgwick himself writes that

it appears that a clear consensus can only be claimed for the principle that a promise...is binding, if a number of conditions are fulfilled: viz. if the promiser has a clear belief as to the sense in which it was understood by the promisee, and if the latter is still in a position to grant release from it, but unwilling to do so, if it was not obtained by force or fraud, if it does not conflict with definite prior obligations, if we do not believe that its fulfilment will be harmful to the promisee, or will inflict a disproportionate sacrifice on the promiser, and if circumstances have not materially changed since it was made (ME 310-11; also 353).

Perhaps Sidgwick can reply by asking what separates these promises from others.

Sidgwick's answer is that there are clear consequentialist rationales for keeping these promises, but not for others (ME 354, 443-4). It seems harder to give an explanation of what makes these promises special that is internal to the duty of fidelity.<sup>53</sup>

So far, I have considered the unfairness charge by looking at common sense morality. One might also press the charge by looking for internal problems with *benevolence*. The obvious worry concerns "more good." Sidgwick intends "more total good" (ME 415-16). But he has no argument for choosing total over average (or some other alternative), or at least no more argument than a common sense moralist might have

---

<sup>53</sup> Of course there may be such an explanation. For one non-consequentialist account of some of these conditions, see Judith Jarvis Thomson, The Realm of Rights (Cambridge, MA: Harvard University Press, 1990) pp. 310-20.

for, say, coming to a particular verdict about whether desert should be proportioned to effort or result.<sup>54</sup> Nor does it seem plausible to argue that there is a duty to maximise the total and a duty to maximise the average.<sup>55</sup> In lieu of a decisive argument, Sidgwick might reply that whatever internal problems *benevolence* has are problems for common sense morality as well as for consequentialism, since common sense morality accepts *benevolence*. This form of the unfairness charge, then, might be met by noting that common sense morality faces extra internal problems—though this is hardly the sort of contrast Sidgwick has in mind when he contrasts, for example, “I ought to keep my promises” with his axioms that “present themselves as self-evident; as much (e.g.) as the mathematical axiom that ‘if equals be added to equals the wholes are equal’” (ME 383).

\*\*\*

Stepping back from the details, Sidgwick’s argument is simple. Consider Ross’s list of prima facie duties. Four are presented as instances of the duty to promote the good. Three—gratitude, recompense, fidelity to promises—are not. Sidgwick’s axioms support the duty to promote the good. Since they do not say that this is the only duty, Ross, and common sense moralists, would agree. Perhaps this is why Sidgwick was so careless in presenting the axioms—the only people he thought might reject them were

---

<sup>54</sup> For this point, see Phillips p. 103 and, especially, Wallace pp. 41-4. Broad thinks a utilitarian must favour maximising the total, and that that “would quite certainly be wrong” (Five Types p. 250). Ross, citing Broad, agrees, though again accepts that taking certain distributions as good avoids the problem (Foundations pp. 70-1, 73-4).

<sup>55</sup> See, for example, some of the problems for maximizing the average in Parfit, Reasons and Persons (Oxford: Oxford University Press, 1984) pp. 420-422.

egoists, and here he was careful (sometimes) to word the axioms so that egoists would agree. The issue between Sidgwick and common sense morality is not the axioms, but the much longer portion of the Methods that argues, sometimes unfairly, that the tests rule out any additional axioms. Those who favour the argument for consequentialism from the axioms must direct their energies to showing that Sidgwick is fair after all, by showing that he raises decisive internal problems for all prima facie duties other than the duty to promote the good.

### Appendix

In the first edition of the Methods, in a discussion of what becomes *personal irrelevance*, Sidgwick writes that we

are supposed to judge that there is something intrinsically desirable—some result which it would be reasonable for each individual to seek for himself, if he considered himself alone....[But] the Good of any one individual cannot be more intrinsically desirable, because it is his, than the equal Good of any other individual. So that our notion of Ultimate Good, at the realization of which it is evidently reasonable to aim, must include the Good of every one on the same ground that it includes that of any one (ME (1) 360).

Sidgwick sums up what has been shown: “we have...evolved the supression of Egoism, the necessary universality of view, which is implied in the mere form of the objective judgment ‘that an end is good’ just as it is in the judgment ‘that an action is right’” (ME (1) 366). Later he notes that the egoist can escape by denying that “his happiness...is objectively desirable or Good” (ME (1) 391).

In the second edition, *personal irrelevance* is stated as it is in the final edition, except that “as a part of universal good” occupies the place later taken by “from the point of view of the universe” (ME (2) 355). Sidgwick adds *general good*: “as rational beings we are manifestly bound to aim at good generally, not merely at this or that part of it; we can only evade the conviction of this objectivity by denying that there is any such universal good” (ME (2) 355). He also adds *benevolence* and claims that a “plain man” would find it wrong to choose his own lesser happiness over the greater happiness of another. The egoist escapes by denying that his good is good “absolutely” (ME (2) 389).

In the third edition, “from the point of view of the universe” appears in *personal irrelevance* (ME (3) 381). The egoist escapes by refusing to speak of good “taken universally” (ME (3) 417).

In the fourth edition, Sidgwick drops the claim, attached to *general good*, about evading the conviction of objectivity by denying that there is a universal good (ME (4) 382). The egoist escapes by avoiding “the point of view of the Universe,” as in the final edition (ME (4) 417).

It is noteworthy that in the earlier editions, Sidgwick argues that using a certain concept or concepts—“intrinsically desirable,” “ultimate good,” “universal good,” “objective good,” “good” or “desirable” “used absolutely”—commits one to thinking that the good of one is no more important than the good of another.

For some of the concepts listed, Sidgwick’s argument seems too ambitious. I can speak of my happiness as “intrinsically desirable” without being committed to thinking that my happiness is no more important than the happiness of another. I seem committed

only to thinking that I ought to desire my happiness as an end rather than merely as a means.

Sidgwick also sometimes puts his point in an ambiguous way. When he writes, for example, of “denying that there is...universal good,” he might be denying the existence of a concept, or of instances of the concept, or of use of the concept.<sup>56</sup> The last is best.

Perhaps for these reasons, in later editions Sidgwick stresses not what is entailed by certain concepts, but instead what is entailed by taking up a certain point of view. On this view, the egoist uses the same concepts as everyone else, but refuses to take up the “point of view of the universe.”

Robert Shaver

Philosophy, University of Manitoba

---

<sup>56</sup> I think this ambiguity slightly mars Phillips’ treatment of the axioms. He thinks Sidgwick’s argument turns on the claim that “there is a general reason to suppose that there is such a thing as universal goodness. And Sidgwick, so far as I can see, supplies no reason” (125). It is not the existence of the concept or instances of it that can be denied, but rather that one must employ it.