



Project  
**MUSE**<sup>®</sup>

*Today's Research. Tomorrow's Inspiration.*

---

## Henry Sidgwick's Moral Epistemology

Anthony Skelton

Journal of the History of Philosophy, Volume 48, Number 4, October 2010, pp. 491-519 (Article)

Published by The Johns Hopkins University Press



For additional information about this article

<http://muse.jhu.edu/journals/hph/summary/v048/48.4.skelton.html>

# Henry Sidgwick's Moral Epistemology

ANTHONY SKELTON

THE PRECISE NATURE OF HENRY SIDGWICK'S moral epistemology is disputed.<sup>1</sup> In *A Theory of Justice*, John Rawls claims that in *The Methods of Ethics* Sidgwick employs a coherentist version of reflective equilibrium.<sup>2</sup> On this view, Sidgwick's justification of utilitarianism is exclusively a matter of demonstrating that it explicates and coheres with common-sense morality better than its competitors after due deliberation and reflective adjudication. *Pace* Rawls, Peter Singer argues that Sidgwick is committed to a form of intuitionist foundationalism, and therefore relies on self-evident propositions to justify utilitarianism.<sup>3</sup> However, unlike Rawls, Singer is less clear about the role that common-sense morality plays in Sidgwick's episte-

---

<sup>1</sup>The following abbreviations of works by Henry Sidgwick are used throughout:

BE Critical Notice of F. H. Bradley, *Ethical Studies*, *Mind* 1 (1876): 545–49.

EOP *The Elements of Politics*, fourth edition (London: Macmillan, 1919).

EP "The Establishment of Ethical First Principles," *Mind* 4 (1879): 106–11.

FC "Some Fundamental Ethical Controversies," *Mind* 14 (1889): 473–87.

IO "Is the Distinction Between 'Is' and 'Ought' Ultimate and Irreducible?," *Proceedings of the Aristotelian Society* 2 (1891–92): 88–92.

LE *Lectures on the Ethics of T. H. Green, Mr. Herbert Spencer, and J. Martineau*, ed. E. E. Constance Jones (London: Macmillan, 1902).

LK *Lectures on the Philosophy of Kant and Other Philosophical Lectures and Essays*, ed. James Ward (London: Macmillan, 1905).

ME *The Methods of Ethics*, 7th ed. (London: Macmillan, 1907). References to the first, second, third and fourth editions (London: Macmillan, 1874, 1877, 1884, 1890) take the form "ME1," "ME2," "ME3," or "ME4."

PC "Professor Calderwood on Intuitionism in Morals," *Mind* 1 (1876): 563–66.

PE *Practical Ethics: A Collection of Addresses and Essays*, ed. Sissela Bok (Oxford: Oxford University Press, 1998).

SE "Mr. Barratt on 'The Suppression of Egoism,'" *Mind* 2 (1877): 411–12.

U "Utilitarianism," *Utilitas* 12 (2000): 253–60.

VB "Verification of Beliefs," *Contemporary Review* 17 (1871): 582–90.

<sup>2</sup>John Rawls, *A Theory of Justice* [*Theory*] (Cambridge, MA: Harvard University Press, 1971), 51.

<sup>3</sup>Peter Singer, "Sidgwick and Reflective Equilibrium" ["Reflective"], *The Monist* 58 (1974): 490–517.

---

**Anthony Skelton** is Assistant Professor of Philosophy at the University of Western Ontario.

*Journal of the History of Philosophy*, vol. 48, no. 4 (2010) 491–520

mology. At times he appears to believe that self-evident propositions perform all of the epistemological labor and that Sidgwick relies on common-sense morality only as part of an “*ad hominem*” argument addressed to the supporter of common sense morality.”<sup>4</sup> At other times, his view appears to be that, for Sidgwick, “we may be able to support our intuitions to some extent by a comparison with common sense morality.”<sup>5</sup> In the years since the debate between Rawls and Singer first occurred, the second of Singer’s accounts of Sidgwick’s epistemology has in one form or another become the orthodox view. The main purpose of this essay is to challenge this orthodoxy.

In this essay I defend the view that Sidgwick’s moral epistemology is a form of intuitionist foundationalism that grants common-sense morality no evidentiary role. In §1 I outline both the problematic of *The Methods of Ethics* and the main elements of its argument for utilitarianism. In §§2–4 I provide my interpretation of Sidgwick’s moral epistemology. In §§5–8 I refute rival interpretations, including the Rawlsian view that Sidgwick endorses some version of reflective equilibrium and the view that he is committed to some kind of pluralistic epistemology. In §9 I contend with some remaining objections to my view.

## I

The problem that marks ME’s philosophical point of departure is a simple one (ME 5–6). The common person relies on several principles and methods in his or her deliberations about what to do or how best to live. However, the common person gives no thought to how these principles or methods are related to each other: “The unphilosophic man is apt to hold different principles at once, and to apply different methods in more or less confused combination” (ME 6). This leaves her in many cases where she has to choose *sans* knowledge of the uniquely right or rational path to follow. She consequently wonders what is *the* right or rational thing to do given that there are several actions in any one circumstance that she believes to have the characteristic of being right or rational.<sup>6</sup> The common person wonders, that is to say, “What do I have ultimate reason to do?”<sup>7</sup>

For Sidgwick, answering this question requires articulating and defending a method of ethics, or “any rational procedure by which we determine what individual human beings ‘ought’—or what it is ‘right’ for them—to do, or to seek to realise by voluntary action” (ME 1). In attempting to articulate a method of ethics, the moral philosopher, unlike the common man, “seeks unity of principle, and consistency of method” (ME 6). A moral framework of this sort would provide an account of which factor matters most to the morality (or rationality) of an action, rule, institution, and so on, and would therefore be capable of delivering

<sup>4</sup>Singer, “Reflective,” 498.

<sup>5</sup>Ibid., 514; also 517.

<sup>6</sup>Sidgwick equates right and rational; see ME, bk. I, ch. 3, 105, 375.

<sup>7</sup>The unifying theme of ME is an examination of several attempts to answer this question. The unity of the work does not therefore depend on seeing it as a “thorough and systematic examination of a single moral theory ... [namely] the one implicit in common sense morality,” *contra* Janice Daurio, “Sidgwick on Moral Theories and Common Sense Morality,” *History of Philosophy Quarterly* 14 (1997): 425–45, at 426.

uniquely rational dictates (ME vi). A method of ethics would, then, determine the *ultimate*, categorically prescribed reason for action and the *method* required for meeting it (ME 7–8). The philosopher differs from the common man in that the former assumes that there can be at most one (valid) ultimate reason for action (ME 6, 12).

Sidgwick considers three distinct methods of ethics, namely, dogmatic intuitionism, rational egoism, and utilitarianism.<sup>8</sup> In so doing, he does not attempt “any such complete and final solution of the chief ethical difficulties and controversies as would convert this exposition of various methods into the development of a harmonious system” (ME 13). Indeed, he concludes that both rational egoism and utilitarianism appear to be equally plausible yet rival accounts of what we have most reason to do, and that therefore there is a dualism of practical reason (ME 496–509). The nature of this dualism is not of interest to me here. Instead, my concern is with the argument that Sidgwick marshals in favor of the utilitarian side of it, since it provides the clearest indication of the moral epistemology to which he subscribes.<sup>9</sup>

The main features of Sidgwick's argument for utilitarianism are clear. It appears to have three separate parts, each of which plays a distinct role. The defense comprises a negative argument against common-sense morality and its philosophical counterpart, dogmatic intuitionism (the main features of which are found in Book III, chapter xi), an appeal to philosophical intuitions (which takes place primarily in Book III, chapters xiii–xiv), and a proof of the sort that Mill provides (which is supplied in Book IV, chapters ii–iii). An explanation of the structure of the case for utilitarianism and the justificatory architecture to which it appeals (if any) is the chief task of any account of Sidgwick's epistemology. In particular, such an account must demonstrate the role played by common-sense morality in the argument and the underlying account of justification. To provide this I am going to focus in what follows on that which he holds are the epistemological and methodological “considerations which should ... be decisive in determining the adoption of ethical first principles” (ME 14).

## 2

Sidgwick endorses philosophical intuitionism, the view that there exist “one or more principles more absolutely and undeniably true and evident” (ME 102). Philosophical intuitions are non-derivatively justified. More specifically, they are “self-evident principles relating to ‘what ought to be’” (ME 102n1; also 379, 383, among others). Therefore a proper understanding of such propositions is sufficient for acquiring intuitive or immediate knowledge of their content (ME 229). The

---

<sup>8</sup>Dogmatic intuitionism holds that “certain kinds of actions are right and reasonable in themselves, apart from their consequences—or rather with a merely partial consideration of consequences, from which other consequences admitted to be possibly good or bad are definitely excluded” (ME 200). Rational egoism is the view that an agent is rational insofar as an agent seeks to maximize his own happiness. Utilitarianism maintains that an agent acts rightly insofar as she performs that action, out of the range of actions open to her, that maximizes aggregate happiness.

<sup>9</sup>For Sidgwick's commitment to utilitarianism, see FC, EOP, and PE.

intuitions that he arrives at—e.g. “that as 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)—are known “by direct reflection” or “by merely reflecting upon them” (ME 383; also PC 565). In keeping with this, he maintains that we have the cognitive capacity for arriving at warranted beliefs or knowledge directly, i.e. we have the capacity to acquire knowledge on the basis of reflection or rational intuition alone (ME 34, 34n1, 400–1, 462). He likens these principles—which he sometimes refers to as axioms (ME 373, 386)—to the axioms of logic and mathematics (ME 383, 507). The function of the moral philosopher’s principles is to dictate to people “what they ought to think, rather than what they do think: he is expected to transcend Common Sense in his premises, and is allowed a certain divergence from Common Sense in his conclusions” (ME 373). The philosopher’s aim in short is to “enunciate, in full breadth and clearness, those primary intuitions of Reason, by the scientific application of which the common moral thought of mankind may be at once systematised and corrected” (ME 373–74).

Accordingly, the philosophical intuitionist is “not necessarily bound” to take “the commonly accepted moral rules ... as the basis on which his own system is constructed” (ME 373). For even if the rules and judgements of common-sense morality were to combine to form a coherent and complete set of rational dictates, this does not constitute a justification of this system, since “the resulting code seems an accidental aggregate of precepts, which stands in need of some rational synthesis” (ME 102). This rational synthesis is provided by philosophical intuitions, which are indispensable to ethical argument. “In all such [moral] reasonings there must be some ultimate premises: which, as they are not known as inferences from other truths, must be known directly—that is, by Intuition” (PC 564; also ME 98).

It is no surprise, then, to find that Sidgwick takes philosophical intuitions to be central to the justification of utilitarianism. Very early on in his thinking he came to the conclusion that “the utilitarian method ... could not ... be made coherent and harmonious without ... [a] fundamental intuition” (ME xvi–xvii). The philosophical intuitions to which Sidgwick commits himself emerge after a long and exhaustive survey and critical evaluation of the main rules of common-sense morality, philosophical renditions of which are accepted by dogmatic intuitionists. At the conclusion of this discussion he argues that “from such regulation of conduct as the Common Sense of mankind really supports, no proposition can be elicited which, when fairly contemplated, even appears to have the characteristic of a scientific axiom” (ME 360). Although common sense cannot offer us any “independent and self-evident rules of morality” (ME 360), it “may still be perfectly adequate to give practical guidance to common people in common circumstances” (ME 361).

The rules of common-sense morality cannot be classed as intuitive truths because they do not satisfy the tests that Sidgwick claims intuitions must meet. It is important to be clear on the nature of the tests on which he relies. He does not hold that intuitions are infallible. He claims that “by calling any affirmation as to the rightness or wrongness of actions ‘intuitive,’ ... I only mean that its truth is apparently known immediately, and not as the result of reasoning” (ME 211; cf. 400). Indeed, “the history of thought makes it only too clear that the human

mind, philosophic and unphilosophic, is liable to affirm as self-evidently true what is afterwards agreed to be false" (EP 108; also PC 565). To eliminate error from our intuitions, he appeals to a set of "characteristics by which self-evident truths are distinguished from mere opinions" (ME 338).<sup>10</sup> Sidgwick relies on four "conditions" for uncovering and eliminating error in our intuitions. (1) The "terms of the proposition must be clear and precise" (ME 338). (2) The "self-evidence of the proposition must be ascertained by careful reflection" (ME 339). (3) "The propositions accepted as self-evident must be mutually consistent" (ME 341). (4) Dissent from the proposition must be absent or (rationally) explained away, for if my judgements are "in direct conflict with a judgement of some other mind, there must be error somewhere" (ME 342). He describes these as "conditions, the complete fulfilment of which would establish a significant proposition, apparently self-evident, in the highest degree of certainty attainable: and which must be approximately realized by the premises of our reasoning in any inquiry, if that reasoning is to lead us cogently to trustworthy conclusions" (ME 338). The basic idea appears to be that the "conditions" for self-evidence are designed to correct for error in our intuitive capacity. They provide us with "some further protection against the possible failure of our faculty of intuition" (EP 109). Intuitions may be corrected by further reflection, by appeals to consistency, by appeals to clarity and precision, and by comparison with those of other minds (EP 109; ME 212).

The first condition demands that the terms of self-evident propositions be clear and precise. We often discover error in our intuitions by getting clearer on either their practical directives or their terms. The second condition demands that the self-evidence of a proposition be ascertained by careful reflection. This includes inquiring into the causes of an intuition, since they may "suggest to the reflective mind sources of error to which a superficial view of it is liable" (ME 212). In addition, reflection needs to be guided by reason, unfettered by prejudice or bias or authority or tradition or "the more subtle and latent effect of these in fashioning our minds to a facile and unquestioning admission of common but unwarranted assumptions" (ME 339; also 403, 211–12). Sidgwick is particularly concerned to ensure that one does not accept a proposition as self-evident simply because the proposition in question is affirmed by common sense or convention or one's preferences (ME 340–41; EP 108; PC 565). The basis for the affirmation needs to lie in the fact that the proposition appears as a requirement of rationality.

The third condition calls for consistency among propositions one has accepted as self-evident (ME 341; EP 109). This condition is explicitly put forward as designed to uncover and correct errors in our intuitive faculty, for "it is obvious that any collision between two intuitions is a proof that there is error in one or the other, or in both" (ME 341; also 213). This test requires not only that one compare one's intuitions in order to ensure their consistency. It also appears to require drawing out the implications of the intuitions to ensure that a derivation

---

<sup>10</sup>It is important not to be confused by Sidgwick's language in his discussion. He describes himself as using these tests to determine the self-evidence of certain propositions. However, this is not quite right. His actual aim is to use the tests to determine the intuitive warrant or truth of the propositions he examines. This is obvious from the way in which he employs the tests.

of one of the intuitions does not conflict with another intuition or derivations from it. Discovering such conflicts comes about only by carefully comparing one's intuitions and their consequences with each other (LK 462–63).

The fourth condition is also designed to reveal error in or cast doubt on the authenticity of an intuition. In cases of disagreement, since “it is implied in the very notion of Truth that it is essentially the same for all minds, the denial by another of a proposition that I have affirmed has a tendency to impair my confidence in its validity” (ME 341).<sup>11</sup> It is clear from what Sidgwick says that confidence should be impaired only in cases where one has no reason to believe that one's own faculty is more likely to be correct than another's. As he puts it, “If I have no more reason to suspect error in the other mind than in my own, reflective comparison between the two judgements necessarily reduces me temporarily to a state of neutrality” (ME 342). In other words, if I do have more reason to believe that I am correct, then disagreement should not impair my confidence. Whatever the case may be, agreement appears to have no evidential value (ME 341). Instead, “the absence of such disagreement must remain an indispensable negative condition of the certainty of our beliefs” (ME 342). This understanding of the fourth condition is confirmed when Sidgwick says at the conclusion of his discussion of the intuitions he accepts that he would “rely less confidently” on them were they contrary to fact rejected by Clarke or Kant (ME 384). He does not say that the agreement provides him with yet more reason to believe or have more confidence in them.

Returning now to the argument against common-sense morality, Sidgwick argues that the problem with the rules of common-sense morality is that

So long as they are left in the state of somewhat vague generalities, as we meet them in ordinary discourse, we are disposed to yield them unquestioning assent. ... But as soon as we attempt to give them the definiteness which science requires, we find that we cannot do this without abandoning the universality of acceptance. (ME 342)

In other words, either the rules of common-sense morality fail the first test, or they fail the fourth test.<sup>12</sup> From this negative argument against common-sense morality emerges Sidgwick's positive thesis that certain philosophical intuitions do meet his conditions, and that these philosophical intuitions “exhibit the truth of” the first principle of utilitarianism (ME 387; also 388). The philosophical intuitions that he embraces include, among others, the claims that “as 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) and that “Happiness (when explained to mean a sum of pleasures) ... [is] the sole ultimate end” (ME 402; also LE 107, 126–31).<sup>13</sup> At the conclusion of his discussion of philosophical intuitionism he

---

<sup>11</sup>How serious Sidgwick is about this test can be questioned, since he scarcely considers the lack of agreement regarding the philosophical intuitions he defends. This is especially true of his intuition regarding the ultimate good, the defense of which appears in ME III, ch. xiv and in LE, chs. vi–viii. Indeed, Sidgwick is rather less than scrupulous when it comes to determining that his intuitions meet his conditions. For more on this, see Anthony Skelton, “Sidgwick's Philosophical Intuitions” [“Sidgwick's”], *Etica & Politica / Ethics & Politics* 10 (2008): 185–209.

<sup>12</sup>He does not devote much space to the version of intuitionism that he calls perceptual intuitionism, but it is clear that he rejects it for similar reasons (ME 100–1, 214).

<sup>13</sup>For a full list of the philosophical intuitions that Sidgwick accepts, see Appendix I. For a full defense of this account of Sidgwick's intuitions, see Skelton, “Sidgwick's,” 188–202.

announces that “the Intuitionist method rigorously applied yields as its final result the doctrine of pure Universalistic Hedonism—which it is convenient to denote by the single word, Utilitarianism” (ME 406–7). In putting forward this argument Sidgwick takes himself to be reconciling utilitarianism with an intuitionist (rationalist) epistemology.<sup>14</sup>

## 3

The negative argument against common-sense morality and the positive argument in favor of utilitarianism allow Sidgwick to conclude that “the only moral intuitions which sound philosophy can accept as ultimately valid are those which at the same time provide the only possible philosophical basis of the Utilitarian creed” (PC 564).<sup>15</sup> The argument to this point appears to rely on the position that justification is at bottom a matter of appeal to self-evident propositions. This position does not simply affirm the view that “so long as any proposition presents itself [to me] as self-evident ... [then there is] no need of anything more” (EP 108). Such a view “does not sufficiently allow for the complexity of our intellectual processes” (EP 109). A more nuanced view is required in light of the fact that one’s own intuitions are not infallible (EP 109; ME 211). In an effort to protect against error in our intuitions we appeal to a set of tests (ME 338; EP 108–9). The tests do not refer to factors that function to amplify the justification of a proposition that is apparently self-evident. Rather, the tests function to aid us in avoiding error; they outline reliable conditions for avoiding or excluding error from intuitive reflection. By reference to careful reflection, clarity and precision, consistency, and engagement with experts, one can correct one’s intuitions. Moral justification is therefore foundationalist in nature: moral theories are justified by appeal to philosophical intuitions.

I take this to be the basic picture of Sidgwick’s epistemology. The question I now want to explore is whether his moral epistemology is one according to which justification is exclusively a matter of appealing to philosophical intuitions, or whether he thinks that the intuitions on which he relies are supported to some degree by reliance on common-sense morality. This epistemology and the related moral argument are not the only resources that he has at his disposal in his quest to argue for utilitarianism. Thus far I have supplied only an account of what I take to be his epistemology. There are, it is important to note, at least two forms of argument for first principles (EP 107). The first form of argument appeals to “some general criteria for distinguishing true first principles (whether ethical or

---

<sup>14</sup>This is what Sidgwick dubs the “wider sense” of intuitionism, according to which the justification of first principles is non-inferential (ME 201, 102n1). He also saw himself improving upon the questionable logic in Mill’s proof for the principle of utility (ME 387–88). See J. S. Mill, *Utilitarianism*, ed. Roger Crisp (Oxford: Oxford University Press, 1998), ch. iv.

<sup>15</sup>This claim appears to be in tension with Sidgwick’s remark that his examination of common-sense morality “showed me no clear and self-evident principles except such as were perfectly consistent with Utilitarianism” (ME xxi). This makes it seem as though he find intuitions in common-sense morality. However, this does not seem to be his mature view, since he claims at several places that he finds no intuitions in common-sense morality (see ME 360, 496). What Sidgwick means is that the only intuitions that he finds are those that provide the only rational basis for utilitarianism.



non-ethical) from false ones" (EP 107). It is clear that this is a description of the style of argument and epistemology that I have argued is at work in ME. It starts with an individual's intuitive apprehension and then, using various tests, attempts to determine whether the intuition in question is erroneous (EP 108–9).<sup>16</sup> This argumentative style relies on the application of a general epistemological methodology to the domain of ethics.

The second form of argument is quite different from the first. It aims to make the truth of a certain first principle or self-evident proposition manifest to some particular mind (in part) by using a "rational process connecting it with propositions previously accepted by that mind" (EP 106). It is a dialectical process that attempts to cause one to believe a proposition that one subsequently discovers through non-dialectic means to be self-evident. This form of proof relies on Aristotle's distinction between "logical or natural priority in cognition" and "priority in the knowledge of any particular mind" (EP 106).<sup>17</sup>

This form of argument is intended to work as follows. Suppose Zelda believes it is self-evident that all pain of rational beings is to be avoided. Concerned to show that Zelda is wrong, Amy attempts to demonstrate that the fact that something is not rational is an arbitrary basis for ignoring its pain, since pain is bad no matter what the nature of the being in question. Amy says, "If you think that pain is bad, then why should the fact that a being is rational matter to the badness of its pain? After all, pain is pain: it is bad no matter whose it is! You ought to believe instead (since you think pain is bad) that all pain is to be avoided." Through this sort of process, Amy leads Zelda to the "wider" principle that "all pain is to be avoided" from the narrower principle that "all pain of rational beings is to be avoided." "In this case I shall ultimately regard the wider rule as the [self-evident] principle, and the narrower as a deduction from it; in spite of my having been led by a process of reasoning from the latter to the former" (EP 107). This style of argument is employed at the beginning of ME, Book IV. It is described as "a line of argument which on the one hand allows the validity, to a certain extent, of the maxims already accepted, and on the other hand shows them to be not absolutely valid, but needing to be controlled and completed by some more comprehensive [first] principle" (ME 420; also 422). This argument, Sidgwick says, is addressed to the dogmatic intuitionist and to the rational egoist (ME 420).<sup>18</sup>

A number of interpreters have assumed that at this point in ME Sidgwick begins in earnest to give a coherentist argument in favor of utilitarianism, which relies on granting independent probative status to common-sense morality.<sup>19</sup> In the next sec-

<sup>16</sup>I take it that the search for intuitive truth here is the same as the search for self-evidence, since in EP Sidgwick is trying to demonstrate how we can bring someone to see the intuitive truth of a certain proposition. He explicitly equates the two when he says that through the use of a distinction given to us by Aristotle, "we are thus enabled to see that a proposition may be self-evident, i.e. may be properly cognizable without being viewed in connexion with any other propositions; though in order that its truth be apparent to some particular mind, there is still required some rational process connecting it with propositions previously accepted by that mind" (EP 106).

<sup>17</sup>Sidgwick does not provide a bibliographic reference. See Aristotle's *Prior Analytics*, 71b30–72a5, 72b5–35.

<sup>18</sup>He confirms that this is the argument he employs against rational egoism and dogmatic intuitionism. See EP 107n1. He admits that the argument fails against the rational egoist (ME 420).

<sup>19</sup>David Brink, "Common Sense and First Principles in Sidgwick's *Methods*" ["Common Sense"], *Social Philosophy and Policy* 11 (1994): 179–201, at 182–84; Steven Sverdlik, "Sidgwick's Methodology"

tion I provide an account of the Book IV argument that is consistent with and confirms Sidgwick's commitment to the epistemology that I ascribed to him above. The subsequent sections defend the view against attack and impugn rival interpretations.

It is important to point out that Sidgwick thinks that the above two argument forms are the only two legitimate (rational) forms of establishing first principles (EP 107).<sup>20</sup> He denies, for instance, that one can establish a moral theory by reference to a psychological or physical or theological theory alone. The foundations of ethics are found in ethics and reason itself, not in revelation or biology or metaphysics. Moral principles are not derivable from non-normative propositions, *pace* Mill and some of the other utilitarians and some idealists. As he puts it, "The premises of our reasoning, when strictly stated, must ... be purely ethical: that is, they must contain, implicitly or explicitly, the elementary notion signified by the term 'ought'; otherwise, there is no rational transition possible to a proposition that does affirm 'what ought to be'" (EP 107).<sup>21</sup> Of course, he does think that methodological premises play a role in the discovery of first principles, especially in so far as these premises relate to the truth or falsity of various first principles. Indeed, he describes the first form of argument discussed above as in part the application of an epistemological method to the domain of ethics (EP 107).

## 4

In ME Sidgwick introduces the second form of argument in Book IV, chapter ii where he maintains that his chief aim is to provide a "proof" of the first principle of utilitarianism. At first glance, it appears odd for him to undertake the task of "proving" utilitarianism at this point in the book, since he seems to have already advanced a positive argument in the last two chapters of Book III. In the context of his discussion of the propositions that he thinks are self-evident, he claims that "I arrive, in my search for really clear and certain ethical intuitions, at the fundamental principle of Utilitarianism" (ME 387). He claims that these intuitions provide a "rational basis for the Utilitarian system" (ME 387). At the outset of his discussion of Ultimate Good, he remarks that "if the conclusions of the preceding chapters are to be trusted, it would seem that the practical determination of Right Conduct depends on the determination of Ultimate Good" (ME 391). He goes on to argue that the sole ultimate good is happiness, which consists in pleasure defined as "a feeling which, when experienced by intelligent beings, is at least implicitly apprehended as desirable or—in cases of comparison—preferable" (ME 127; also 131; LE 130).<sup>22</sup> After he arrives at this view, he claims again that

---

["Methodology"], *Journal of the History of Philosophy* 23 (1985): 537–53, at 539–40; J. B. Schneewind, *Sidgwick's Ethics and Victorian Moral Philosophy [Sidgwick's Ethics]* (Oxford: Oxford University Press, 1977), 260–65. I say "in earnest" because Book IV is not the only place to which interpreters have pointed in their attempt to show that Sidgwick is working with (at least in part) a coherentist epistemology. I discuss these passages below.

<sup>20</sup>That is, there is no third *rational* means of convincing someone to endorse a first principle. There may be many non-rational ways of convincing someone. See ME 420n1.

<sup>21</sup>Sidgwick argues for this conclusion in IO and LE.

<sup>22</sup>Sidgwick is less than clear about what he means by pleasure. The passage quoted in the text is just one of the accounts of pleasure he provides. For others, see ME 94; 93, 120–21; 402. He appears to favor the account quoted in the text; see LE 130; ME 398.

“the Intuitionist method rigorously applied yields as its final result the doctrine of pure Universalistic Hedonism—which it is convenient to denote by the single word, Utilitarianism” (ME 406–7).

Yet at the beginning of Book IV, he says that what has been his main concern with respect to utilitarianism is “not how its principle is to be proved to those who do not accept it, but what consequences are logically involved in its acceptance” (ME 418). But has he not clearly been primarily concerned with much more than what is “logically involved” in accepting utilitarianism? He does appear to acknowledge his previous argument in a footnote where he states that “the argument in chap. xiii. [of Book III] only leads to the first principle of Utilitarianism, if it be admitted that Happiness is the only thing ultimately and intrinsically Good or Desirable. I afterwards in chap. xiv. endeavoured to bring Common Sense to this admission” (ME 421n1; also 421, 462, 496–97). He does not connect it with what he takes to be his main aim in Book IV. However, if we interpret Sidgwick’s Book IV argument as akin to the first form of argument outlined in EP, the mysteries disappear. The second argument outlined in EP is the argument that the utilitarian employs to epistemically justify the view’s first principle, while the first argument is addressed primarily to those with whom he disagrees. The first argument described in EP is alone designed to deal with disagreement, while the second argument is concerned with truth or epistemic justification. That the first argument is designed to deal with disagreement and the second is not is confirmed by the fact that Sidgwick states that what the arguments outlined share in common is that each is a “species of the rational process ... by which we are logically led to a conclusion which yet when reached we regard as a first principle” (EP 107). He does not say that each is designed to deal with disagreement in first principles of the sort discussed in EP. This explains why Sidgwick would not employ the epistemic argument to deal with the disagreement he notes in Book IV where it may appear to be useful to him.

This may be too quick. There is a sense in which the intuitive argument is designed to deal with a certain sort of disagreement, namely, disagreement over the existence of intuitions. The second argument discussed in EP appears to be designed to ward off a certain kind of skeptical challenge. Elsewhere Sidgwick says that the method is used to “verify beliefs originally certain, if their certainty be called in question on general grounds” (VB 584; also LK 433–60). This, too, would explain why Sidgwick does not appeal to this argument in Book IV, where intuitions are not being subjected to this sort of skeptical attack.

In addition, there is some evidence that Sidgwick is offering the Book IV argument to help him complete the case in favor of utilitarianism.<sup>23</sup> In earlier editions of the *Methods*, he is clear that he is offering two arguments for utilitarianism. In the first three editions of the *Methods* he says that though the argument of Book III is “effective,” it is “scarcely adapted to produce perfect conviction” (ME1 392; ME2 390; ME3 418). This suggests that he is relying on his previous argument at this stage in the book, and that the Book IV argument is presented to produce “perfect conviction” in favor of the “first principle” of utilitarianism (at least against common-sense morality). Sidgwick claims that the appeal to philosophical intuitions

<sup>23</sup>It may be seen in part as a way of satisfying the fourth test above.

in Book III establishes that the utilitarian first principle is only “*one* moral axiom: it does not prove that it is *sole* or *supreme*” (ME 421; italics in original), and that the negative argument contra common-sense morality and dogmatic intuitionism of Book III helps him move from this claim to the conclusion that the philosophical intuitions that justify utilitarianism are the only intuitions. The only part of the argument that is left to complete in Book IV is the argument that demonstrates the positive relation between utilitarianism and common-sense morality, and to demonstrate this Sidgwick relies on something like the first sort of argument that he outlines in EP. This argument together with the others is supposed to produce “perfect conviction.” In other words, the Book III argument by itself is effective, but it does not produce perfect conviction, and therefore he supplies a proof to deal with any residual disagreement. This again would explain why he does not mention the Book III argument in Book IV.

Is the proof in Book IV and the epistemology that it presupposes consistent with the epistemology presupposed by the argument for utilitarianism in Book III? Is it the case that Sidgwick is confused, and provides us with two proofs with two different and conflicting epistemologies?<sup>24</sup> Or is the Book IV proof a continuation of the case for utilitarianism and consistent with the epistemology that he has already provided in Book III? That is, are there two arguments with only one epistemology? Or, are there two arguments with two epistemologies? I argue for the former option.

Sidgwick employs his Book IV argument or proof in an effort to solve a specific problem that he thinks besets any individual who subscribes to an ethical first principle or self-evident proposition.<sup>25</sup> He says that these principles cannot be proved, if by “proved” we mean “a process which exhibits the principle in question as an inference from premises upon which it remains dependent for its certainty; for these premises, and not the inference drawn from them, would then be the real first principles” (ME 419; also EP 106). Yet there must be some way to convince people doubting the principles in question to accept them; we must provide rivals with “rational inducements for accepting them” (EP 106). Without any such inducements we are left without any resources to appeal to in cases of disagreement about first principles. In other words, because there is a conflict of “opinion as to first principles, we can hardly refuse to give reasons for taking our side in the conflict: as rational beings conversing with other rationals it seems absurd that we should not be able to explain to each other why we accept one first principle rather than another” (EP 106; also ME 420).

At the same time, we do not want to undermine the first principle’s status by deriving it from another proposition or set of propositions. The trick then is to provide some “rational inducements” to accept a principle without deriving the principle from other propositions or “premises” (EP 106; ME 419). Hence the rational inducements that he is going to provide are not the sort that exhibits the

---

<sup>24</sup>Brink (“Common Sense,” 191–92) suggests that Sidgwick relies on two argumentative strategies to establish utilitarianism, but wonders why, if the philosophical intuitions justify utilitarianism, Sidgwick needs to appeal to coherence with common-sense morality. He suggests that Sidgwick is confused. What I go on to say dispels the confusion.

<sup>25</sup>The argument is addressed mainly to the advocate of common-sense morality.

principles in question “as dependent for their certainty on the acceptance by the mind of certain other truths” (EP 106). Sidgwick describes this kind of “proof” as “clearly different from ordinary proof” (ME 420). It supposes that the rational inducements are valid but that they do not “prove [in the ordinary sense] the first principle which they (to use Mill’s phrase) ‘determine the mind’ to accept” (EP 106). The proof relies on the validity of some propositions that are used to lead one to accept, but do not justify, a first principle.<sup>26</sup>

Now it seems perfectly clear from this picture that Sidgwick is trying not to contradict the foundationalist epistemology that is at work in Book III and elsewhere. He is working with the idea that we can use certain “reasons” which are “valid” to “determine the mind to accept” a first principle, but not at the same time to epistemically justify that principle, since this would leave him with an “inconsistency” in his epistemology (EP 106). The argument has rational but not deductive force. An inconsistency would arise if the “valid” reasons were used as justification, because the reasons that he is using to lead one to accept a first principle are justified *only* by appeal to the same principle to which they are supposed to lead one (EP 107). To appeal to these reasons to justify the first principle would be circular. It is important for Sidgwick to keep this problem in the front of his mind, for he says that we can lead one who believes that “all pain of human and rational beings is to be avoided” to the claim that “all pain is to be avoided” even though the only epistemic basis of the former is found in the latter (EP 106–7).

He makes this maneuver with respect to common-sense morality as well. He accepts some of the maxims of common sense as valid, so as to lead its proponents to the first principle of utilitarianism. But by “validity” he does not mean that they have warrant apart from the first principle. This is no surprise, since both within and without ME he claims that the rules and judgements of common-sense morality receive their justification from utilitarianism alone (ME 496–97; PE 73).<sup>27</sup> Sidgwick is also trying to avoid a second form of “inconsistency.” It would be inconsistent to argue that a certain proposition is a first principle but then proceed to argue for it using premises logically prior to it. This would undermine the status of the first principle (hence generating an “inconsistency” by seeming to assume that the same principle is simultaneously superordinate and subordinate).

What, then, is the best way to interpret the proof? In light of the foregoing, it seems that the proof is best understood as an *ad hominem* argument in favor of utilitarianism.<sup>28</sup> It tries to show that, given what the proponents of common-sense morality and rational egoism already believe, they ought to believe utilitarianism. This argument is especially effective against the proponents of common-sense

---

<sup>26</sup>Aristotle also allows for the demonstration of a principle in this sense. He calls it demonstration in a qualified sense. See *Posterior Analytics*, 72b30–35.

<sup>27</sup>More on this point below.

<sup>28</sup>This is one of the views put forward in Singer, “Reflective,” and it is suggested by Terence Irwin in “Aristotle’s Methods of Ethics,” in *Studies in Aristotle*, ed. Dominic J. O’Meara. (Washington: Catholic University of America Press, 1981), 193–223, at 201–4. My account of the compatibility of the two arguments for utilitarianism is more consistent than Singer’s, since at no point do I claim that common-sense morality is “an important check on our intuitions of self-evidence about ethical first principles” (509). It is more robust than both, since it provides textual evidence.

morality. According to Sidgwick, utilitarianism makes up for many of the deficiencies found in common moral thought. It explains and justifies many of its most important requirements.<sup>29</sup> It provides a principle for clarifying, synthesizing, and “binding the unconnected and occasionally conflicting principles of common moral reasoning into a complete and harmonious system” (ME 422). Because proponents of common-sense morality accept its most important elements, but in addition desire a justification for their principles, a method for adjudicating conflicts, and, finally, a “complete and harmonious system,” they have rational inducements to accept utilitarianism.<sup>30</sup>

But the argument follows only in so far as the proponents desire such features of their moral position. Were they to deny that completeness is a virtue of a theory or that conflicts always admit of a unique answer, they would have no reason to accept utilitarianism. Indeed, Sidgwick admits that the *ad hominem* style of argument is completely impotent against the rational egoist, so long as the egoist denies that his own happiness is good “not only *for him* but from the point of view of the Universe” (ME 420; italics in original). The *ad hominem* proof works on the egoist only in so far as he claims that his happiness is agent-neutrally valuable.<sup>31</sup>

It should be clear from what I have said thus far that a coherentist reading of Sidgwick's proof is implausible. If the proof is coherentist in nature, then there is no need to be careful about showing that the first principle is not derived from other propositions in the context of it. If he really is giving a coherentist justification of utilitarianism, then he would be keen to show that the theory really was derived from the rules and judgements of common-sense morality. But this he does not do: instead, he tries to do exactly the opposite. This reading of the Book IV proof also confirms my view that the tests for self-evidence function to correct for

<sup>29</sup>Sidgwick is clear that the fit is not perfect. See ME 425–26, 460–67.

<sup>30</sup>Sidgwick also says that part of this proof relies on the fact that common sense appeals to utilitarianism for “further development of its system” (ME 422). This is not a plausible argument, since this could be a case in which common sense is appealing to one of its own principles, namely, benevolence. Although this system may well be coherent, it is still without a “rational synthesis” (ME 102). This rational synthesis is provided by utilitarianism (ME 423–59), a fact that provides a “rational inducement” for believing it.

<sup>31</sup>Brink thinks that the sort of interpretation advanced here “does not explain why the examination of common-sense morality is necessary to the identification of first principles” (Brink, “Common Sense,” 194). This is untrue. The examination is required not only to demonstrate the deficiencies in common-sense morality but also as a distinct heuristic device aimed to aid in the sort of reflection that ultimately leads one to certain first principles of morality. Sidgwick notes that most of us are habitually disposed to believe the main dictates of common-sense morality. But this is not enough to establish that these beliefs are rational. Indeed, we are apt to confuse what we believe by force of habit and rational cognition (ME 212, 339, 341, 383). Consequently, we need to learn to separate “opinions and sentiments” from rational “cognition” (PC 565). Only “reflection” on the morality of common sense can lead us to make this important distinction. As Sidgwick puts it, “I hold that here, as in other departments of thought, the primitive spontaneous processes [common-sense intuitions] of the mind are mixed with error, which is only to be removed gradually by comprehensive reflection upon the results of these processes. Through such a course of reflection I have endeavoured to lead my readers in chaps. 2–10 of Book III. of my treatise: in the hope that after they have gone through it they may find their original apprehension of the self-evidence of moral maxims importantly modified” (PC 565). In other words, the analysis of common moral thought is to lead readers from dogmatic, unreflective acceptance of common moral thought to a truly rational morality. In addition, the analysis of common-sense morality can be used as the basis of an *ad hominem* argument.

erroneous judgements of self-evidence rather than to amplify their justification. If appeal is made to common sense in this context, then we would again have an inconsistency on our hands.

## 5

A number of interpreters reject my view. In *A Theory of Justice*, Rawls suggests that Sidgwick is a proponent of coherentist reflective equilibrium. He holds that “justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view”; it is not a matter of appealing to “self-evident premises or conditions on principles.”<sup>32</sup> Coherence is sought between considered judgements and candidate principles. There is a “definite if limited class of facts against which conjectured principles can be checked, namely, our considered judgements in reflective equilibrium.”<sup>33</sup> Conflicts that arise between considered judgements and principles are reflectively reconciled by either renouncing the considered conviction or by modifying the principle.<sup>34</sup> From these contentions we get the view that the justification of a moral framework is exclusively a matter of demonstrating that it explicates and coheres with common-sense morality better than its competitors after due deliberation and reflective adjudication. But, as we have seen, Sidgwick does not think that justification is exclusively a matter of coherence with common-sense morality (ME 373) nor that coherence is sufficient for justification (ME 102), and he does hold that justification is in part a matter of appealing to self-evident principles. Therefore, Sidgwick is not a proponent of this form of reflective equilibrium.

He may, however, espouse a different version of the view. Although Rawls appears to endorse coherentist reflective equilibrium in *A Theory of Justice*, in “The Independence of Moral Theory” he seems to articulate a different version of the position, which I will call methodological wide reflective equilibrium.<sup>35</sup> According to this account, reflective equilibrium is not and does not presuppose a specific epistemology, e.g. coherentism. The idea of “constructing a correct theory of right and wrong, that is, a systematic account of what we regard as objective moral truths” is put aside when trying to reach wide reflective equilibrium.<sup>36</sup> The method is used to investigate an aspect of human psychology; we are trying to figure out which moral conception (i.e. which structured arrangement of the right, the good and the morally good) people accept given their considered moral judgements at

<sup>32</sup>Rawls, *Theory*, 21. The context makes it clear that he is referring to epistemic justification. This is the most common interpretation of reflective equilibrium in Rawls. See Norman Daniels, “Wide Reflective Equilibrium and Theory Acceptance in Ethics,” *Journal of Philosophy* 76 (1979): 256–82; and Kai Nielson, “Reflective Equilibrium,” in *Blackwell Encyclopedia of Management: Business Ethics*, 2<sup>nd</sup> ed., ed. P. H. Werhane and R. E. Freeman (London: Blackwell, 2005), 449–52.

<sup>33</sup>Rawls, *Theory*, 51.

<sup>34</sup>Ibid., 20–21.

<sup>35</sup>Rawls, “The Independence of Moral Theory” [“Independence”], in *John Rawls: Collected Papers*, ed. Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), 286–302. The suggestion that Sidgwick is a proponent of this version of reflective equilibrium is found in Bart Schultz, “Introduction: Henry Sidgwick Today,” in *Essays on Henry Sidgwick*, ed. Bart Schultz (Cambridge: Cambridge University Press, 1992), 1–61.

<sup>36</sup>Rawls, “Independence,” 288.

all levels of generality. The purpose of the methodology is to discover the moral "conception, or plurality of conceptions, that would survive the rational consideration of all feasible conceptions and all reasonable arguments for them."<sup>37</sup> In arriving at a moral conception or numerous conceptions, the problems of moral truth (and, presumably, of justification) and of metaphysics and of theories of meaning are bracketed and put off until further work on moral conceptions is completed. In particular, providing an account of moral justification and of moral truth has to wait until "we have a deeper understanding of moral conceptions."<sup>38</sup> This methodology is intended to be consistent with a variety of different views in the philosophy of mind and language and, most important of all, moral epistemology. Rawls maintains that "one's moral conception may turn out to be based on self-evident first principles. The procedure of reflective equilibrium does not, by itself, exclude this possibility, however unlikely it may be."<sup>39</sup> The requirements of methodological wide reflective equilibrium are simply that if one assumes that some principles are self-evident, one makes revisions to other beliefs "with conviction and confidence, and continues to affirm these principles when it comes to accepting their consequences in practice."<sup>40</sup>

There is some evidence that this is Sidgwick's view. He remarks at one point that he is neutral as regards the various methods he discusses. His aim in writing ME is no more than to "eliminate or reduce ... [the] indefiniteness and confusion," which "lurks in the fundamental notions of our common practical reasonings" (ME 13; also vi; SE 411). However, this is contradicted by a great number of statements that Sidgwick makes elsewhere about the aim of ME. As we have seen, he attempts to reconcile utilitarianism with an intuitionist epistemology. In doing so he sees himself grounding utilitarianism on a "fundamental intuition" (ME xxi; also xix). He claims that in ME "Rational Utilitarianism" is the view "to which I ultimately endeavour to lead my reader" (PC 564), and that the appeal to the philosophical intuitions that he embraces functions to "exhibit the truth of ... [utilitarianism's] first principle" (ME 387).

He speaks in several of the editions of the *Methods* about attempting to find truths. In ME1 he says that one of his aims is to apply criteria that will help him distinguish "knowledge" from mere opinion (ME1 317). In ME2 he says that his aim is to distinguish "certain truths" from mere opinions (ME2 313). In ME3 he modifies the language to claiming that he aims to distinguish "self-evident truths ... from mere opinions" (ME3 339; also ME 338). He further claims that he finds no intuitive truths in common-sense morality (ME 360). He rejects several different views of ultimate goodness, and he argues vigorously for the proposition that happiness, hedonistically construed, is the sole ultimate good (ME 89–95, 398–406; LE 126–31). This seems to suggest that Sidgwick does not, *contra* the requirements of methodological wide reflective equilibrium, "suspend consideration" of moral truth or "put aside the idea of constructing a correct theory of right and wrong."<sup>41</sup>

<sup>37</sup>Rawls, "Independence," 289.

<sup>38</sup>Ibid., 288.

<sup>39</sup>Ibid.

<sup>40</sup>Ibid., 289.

<sup>41</sup>Ibid., 288.



Of course this account of the reflective equilibrium methodology allows for the possibility that one's moral conception will end up being based on self-evident principles. But the way in which Sidgwick arrives at the idea that utilitarianism is based on self-evident truths and the way in which those employing methodological wide reflective equilibrium do so are distinct. He does not first work out the utilitarian theory or moral conception and then show that it is based on self-evident principles. Instead, he goes on the search for self-evident principles, and he argues that the intuitions that he accepts provide, in some fashion, a rational basis for utilitarianism (ME 388, 406–7).

Two further points are worth noting. Sidgwick appears to rely on a particular view of personal identity in his argument for rational egoism and against utilitarianism (ME 498; FC 484). He argues that this conception of personal identity—"that the distinction between any one individual and any other is real and fundamental"—provides the basis for rational egoism (FC 484; cf. ME 418–19), and therefore plays a pivotal role in generating the dualism of practical reason. He is then not of the view that "every [moral] conception employs a criterion of identity that recognizes the results of the philosophy of mind, [and that] each may specialize its criterion to fit the requirements of a particular moral order and conception of the person."<sup>42</sup> His view appears to be that an account of personal identity can play a role in undermining or supporting a moral view, and that it is not true that "the variations among the criteria [of personal identity] are not antecedent to moral theory but explained by it."<sup>43</sup>

Second, Sidgwick argues that solving the problem of free will is of limited importance to ethics (ME 66; FC 474–79). In his discussion he maintains that the determinist is unable to employ the same accounts of the concepts "merit," "demerit," and "responsibility" as the proponent of freedom of the will (ME 71–72). In response, he opines that the determinist can simply give to these terms a "signification which is not only clear and definite, but, from an [*sic*] utilitarian point of view, the only suitable meaning" (ME 71). This suggests that he is willing to adjust moral terminology in light of what he regards as metaphysical reality, though it is important to note that in this case he also believes that he has a strong moral argument for the same conclusion about these concepts (ME 284n1, 290–93).

In sum, Sidgwick does not espouse either of the two versions of Rawlsian reflective equilibrium which have been attributed to him. Singer is therefore correct that Sidgwick does not appeal to coherence with common-sense morality as "the ultimate test of the validity of utilitarianism."<sup>44</sup> It is therefore incorrect that Sidgwick suspends consideration of truth and justification in working out his methods or conceptions of ethics, and he does think that metaphysical views can force us to rearrange our moral views.

---

<sup>42</sup>Rawls, "Independence," 297.

<sup>43</sup>Ibid., 296.

<sup>44</sup>Singer, "Reflective," 499. Singer provides different arguments for this conclusion. His arguments rely in part on a faulty account of reflective equilibrium.

## 6

Roger Crisp, Robert Shaver and Bart Schultz argue that Sidgwick's epistemology contains elements of both foundationalism and coherentism.<sup>45</sup> Broadly speaking, on this view Sidgwick holds that epistemic justification in ethics is a matter of appealing to both coherence with the propositions of common-sense morality and self-evident propositions. There are some differences between the views. Crisp maintains that failure to capture and explain a vast majority of the truths of common-sense morality casts the truth of a self-evident proposition into doubt, while agreement with common-sense morality, which includes the views of the many and the wise, increases the self-evidence and hence the reliability of a directly justified proposition.<sup>46</sup> Shaver argues that full justification of a moral proposition occurs when that proposition achieves the status of a higher certainty.<sup>47</sup> Schultz concurs.<sup>48</sup> According to this view, a self-evident proposition is not the highest form of certainty. Instead, the highest form of certainty is achieved when self-evident propositions are supported by common-sense moral judgements, both one's own and those of others. Shaver describes the view as follows: "I believe some self-evident proposition *p* on the basis of seeing its self-evidence and seeing that it agrees with common-sense morality. If I have no reason to trust common-sense morality other than noting *p*, seeing the agreement with common-sense morality should not increase my confidence in *p*. But where there is independent reason for believing in common-sense morality, agreement with it increases my confidence in *p*."<sup>49</sup>

## 7

Crisp's view is not acceptable. First, Sidgwick clearly employs the intuitions to correct and systematise common-sense morality (ME 373, 383). Crisp's understanding of Sidgwick's epistemology contradicts this. On his account, the content of a self-evident proposition is constrained by common-sense morality, since any such proposition that is seriously inconsistent with common-sense morality is rejected. This being the case, any acceptable self-evident proposition (by this criterion) will be largely consistent with common-sense morality and therefore have no or nearly no critical or reformatory role to play *vis-à-vis* common-sense morality. There is no robust critical or reformatory role for the self-evident propositions to play, since the only elements of common-sense morality that remain after arriving at an acceptable self-evident proposition will be trivial or unimportant aspects of it. Crisp does say that some aspects of common-sense morality that conflict with a

---

<sup>45</sup>Roger Crisp, "Griffin's Pessimism," in *Well-Being and Morality*, ed. Roger Crisp and Brad Hooker (Oxford: Oxford University Press, 2000), 115–28, at 116–21; Roger Crisp, "Sidgwick and the Boundaries of Intuitionism" ["Boundaries"], in *Ethical Intuitionism: Re-evaluations*, ed. Philip Stratton-Lake (Oxford: Oxford University Press, 2002), 56–75, at 67–72; Robert Shaver, *Rational Egoism [Rational]* (Cambridge: Cambridge University Press, 1999); and Bart Schultz, *Henry Sidgwick: Eye of the Universe [Eye]* (Cambridge: Cambridge University Press, 2004).

<sup>46</sup>Crisp, "Boundaries," 68–69, 72. Crisp holds that the credibility of a self-evident proposition is impugned only if it is "seriously inconsistent" with common-sense morality.

<sup>47</sup>Shaver, *Rational*, 64–68.

<sup>48</sup>Schultz, *Eye*, 190, 200–4.

<sup>49</sup>Shaver, *Rational*, 70–71.

resulting self-evident principle will be rejected. But this is not the sort of critical role that Sidgwick envisages for self-evident intuitions. Crisp's view of self-evident principles does not therefore make proper sense of Sidgwick's employment of philosophical intuitions.

In reply, Crisp might insist that this argument is a *non-sequitur*, for a self-evident proposition that is not seriously inconsistent with common-sense morality may still be in conflict with non-trivial or important aspects of common-sense morality.<sup>50</sup> This reply relies on a certain view of what constitutes a serious inconsistency. A self-evident proposition may be both not seriously inconsistent with common-sense morality and inconsistent with non-trivial aspects of common-sense morality. This is a plausible reply. But it still does not capture the sense in which for Sidgwick the philosophical intuitions are relied on to determine which aspects of common-sense morality are acceptable. If the view is that elements of common-sense morality dictate (in part) which philosophical intuitions to accept, then there are some elements of common-sense morality that are not deemed acceptable purely by reference to philosophical intuitions. But this is not what Sidgwick appears to envisage as the role played by the philosophical intuitions, which are sought to supply rational justification. This reply also seems in conflict with the general theme of Crisp's position. If the position is that a self-evident proposition ought not to be seriously inconsistent with the views of the many and the wise, should not the fact that a self-evident proposition is inconsistent with important aspects of common-sense morality count against it?

Second, Crisp thinks that Sidgwick appeals to the views of the wise and the many for epistemic support.<sup>51</sup> However, Sidgwick repeatedly rejects the idea that he is interested in the views of the many. In a discussion of intuitionism, he says that in the *Methods* he prefers the "moral thought of the reflective few to that of the unreflective many" and that the thought of the unreflective is "crude" (PC 564). In addition, Sidgwick's definition of common-sense morality makes it clear whose views he is concerned with. He defines it as follows: the positive morality of a community when it is regarded as "a body of moral truth, warranted to be such by the *consensus* of mankind—or at least of that portion of mankind which combines adequate intellectual enlightenment with a serious concern for morality" (ME 215; italics in original; also PE 22–23).<sup>52</sup>

Third, Crisp thinks that Sidgwick is an exponent of the idea that there are degrees of self-evidence. The support of the beliefs of the few and the many can increase the self-evidence of a proposition and hence its reliability. There is scant evidence for this view of self-evidence in Sidgwick's writing. He usually speaks as

<sup>50</sup>I owe this reply to an anonymous referee.

<sup>51</sup>Crisp, "Boundaries," 69.

<sup>52</sup>In his defense of the claim that the sole ultimate good is happiness, hedonistically construed, Sidgwick appeals to both intuition and "the ordinary judgements of mankind" (ME 400). An anonymous referee suggested that this is a case in which Sidgwick is appealing to the views of the many. But this is rather less than clear. He states that this argument refers to the "common sense of mankind," and he uses the phrase "common sense" throughout his discussion (ME 401). If we accept Sidgwick's account of common-sense morality just given, then it seems clear that he is concerned with the views of the "portion of mankind which combines adequate intellectual enlightenment with a serious concern for morality" (ME 215). Therefore, he is not obviously appealing to the views of the many.

if a proposition is either self-evident or in need of rational justification (ME 383; PC 565; ME2 349–50). He draws a strict dichotomy between the intuitive and the non-intuitive.<sup>53</sup> For example, he thinks that if it is in fact required of us that we seek general or aggregate happiness, then “it must either be immediately known to be true—and therefore, we may say, a moral intuition—or be inferred ultimately from premises which include at least one such moral intuition” (ME 98). There is no place where Sidgwick says that a proposition has a high degree of self-evidence or that it has some but not maximal self-evidence.

Shaver's view that Sidgwick is after higher certainties and not simply self-evident propositions turns in part on a specific understanding of the tests or conditions (discussed above) that Sidgwick applies to self-evident propositions (ME 338–42). He holds that in appealing to the third and fourth conditions Sidgwick is in part appealing to common-sense morality to enhance the justification of the intuition or putatively self-evident proposition in question. The idea is that in virtue of finding consistency between our own intuitions and our common-sense moral beliefs and the views of competent others we have more reason to believe an intuition or self-evident truth. This position holds that Sidgwick is appealing to external criteria to enhance the epistemic justification of a self-evident proposition. He is not simply trying to avoid erroneous intuitions.

But in light of what I have said above about the conditions, this cannot be correct. First, in the third condition Sidgwick calls for consistency between the propositions one finds self-evident upon reflection, not coherence between one's intuitions. Were he a proponent of the mixed view, then he would be sure to place emphasis here on coherence, rather than mere consistency. He does not appear to adopt the implausible position that mere consistency between propositions enhances their epistemic credibility. Second, he calls for consistency among intuitions, not coherence between one's intuitions and one's common-sense moral beliefs.<sup>54</sup> Third, he characterizes the fourth test as a negative condition on the certainty of our beliefs. He maintains only that conflict would reduce his confidence in his intuitions and not that agreement provides further epistemic reason to believe a proposition or specify conditions regarding reliable belief formation.

These points challenge Shaver's interpretation, but they are not sufficient to refute it. He supports his account of the tests in ME by comparing them with the three methods of verification that Sidgwick employs in his epistemological writings, namely, intuitive, discursive, and social verification (LK 461–67). He argues that these methods of verification are intended to aid in “excluding error” from “the kind of knowledge which is to be taken as the ultimately valid basis of all else that is commonly taken for knowledge” (LK 461). Intuitive verification is essentially analogous to the first two tests in ME (LK 461). It is described as a test applied to “judgements [thought to be] intuitively certain” in “contemplating the beliefs by themselves” (VB 584). He treats this test as an application of a “wider rule,” namely: “Assure yourself of the self-evidence of what appears self-evident, by careful examination” (LK 461). He notes here as elsewhere that this is not sufficient

<sup>53</sup>This is clearer in earlier editions of the *Methods*. See ME1 317, ME2 313–14.

<sup>54</sup>For this point, see Shaver, *Rational*, 65. For more on this, see below.

protection from error: “We cannot unhesitatingly grant that no false intuition ever appears clear and distinct to one who is not misled by it” (VB 585).

The way to remedy this problem is to employ more tests. The discursive test requires “contemplating the belief that appears intuitively certain along with other beliefs which may possibly be found to conflict with it” (LK 462).<sup>55</sup> This test is important because “it provides against the kind of error which the conflict of beliefs reveals” (LK 463). This test is similar to the third test in ME that calls for consistency between intuitions. The social or ecumenical test resembles the fourth test in ME. It aims to eliminate error through comparing our own intuitions with the intuitions of other scrupulous thinkers or experts, for we often discover error when we find that our beliefs “conflict with the beliefs of other persons competent to judge” (LK 465).

In his epistemological writings, Sidgwick suggests that one “case” of discursive verification is inductive verification (LK 463; VB 590). Shaver argues that this forces us to reinterpret Sidgwick as attempting to inferentially justify intuitions in the third test in ME. He argues that we should understand the third test in ME as requiring inductive verification of self-evident propositions by reference to common-sense morality. A self-evident proposition so verified has the status of a higher certainty. There are three problems with this suggestion. First, at times Sidgwick asks only for consistency between self-evident propositions (LK 461, 462–63, 464). At other times he calls for consistency between “beliefs relating to the same subject” (LK 465). Shaver makes a lot of this discrepancy. He exaggerates it. Nearly all of Sidgwick’s remarks call for consistency between intuitions (LK 461, 462–63, 464). *A fortiori*, the reinterpretation ascribes to Sidgwick an implausible view. The test as Shaver reinterprets it is too indiscriminate. Not all beliefs relating to the same subject are to be appealed to, and Sidgwick is (rightly) hesitant to accept our unfiltered beliefs about a certain subject matter in thinking about justification (LK 430). He is especially hesitant to do so in ethics (ME 374). After all, what justification could one have for relying on all beliefs relating to a certain subject, whether examined or unexamined?

Second, there is another explanation of the discrepancies between the epistemological writings and ME, namely, that the tests in ME are similar to but not completely coincident with the tests in Sidgwick’s epistemological writings. There is a strong possibility that the tests function differently across different domains of inquiry. Sidgwick refers to the tests as having a “special application to Ethics” (LK 464n1), and his references to ethics in his epistemological writings confirm my understanding of the tests above (LK 464, 462). According to this account, inductive verification is only one “case” of discursive verification (LK 463).<sup>56</sup> This view would allow that sometimes the epistemology aims at self-evidence and sometimes at higher certainties, and so explains the differences between ME and the epistemological writings.

---

<sup>55</sup>It is clear from the context that we are to compare only intuitions with each other. This test demands “carefully grouping the intuitions that we see to be related” (LK 462; also 463). More on this below.

<sup>56</sup>Sidgwick here suggests that there is more than one variety of discursive verification. Not all discursive verification is inductive verification.

Third, Shaver's argument for the reinterpretation is based on a mistaken inference. Shaver moves from Sidgwick's claim that intuitive verification of an apparently self-evident proposition is not enough to protect us from error to the claim that Sidgwick is no longer aiming for self-evident propositions but rather higher certainties. True, he thinks that solitary careful reflection may not eliminate error from our intuitions, making the felt need for more conditions palpable. But his point is that we need more than careful reflection to protect us from errors in our judgements about what is self-evident, not that we need more than self-evidence. We need to assure ourselves of the self-evidence of a proposition in ways other than by solitary careful reflection. The other ways include aiming for greater clarity, comparing intuitions alongside each other in order to eliminate inconsistencies, and comparing one's own results with those of others who have contemplated the proposition in question. As he says, "Careful reflection ... may indeed be much aided by communication with other minds" (ME 212; also 28). So, although Sidgwick appeals to more than solitary careful reflection to save him from errors in affirmations of self-evidence, it does not follow that he thinks that self-evidence is not the highest form of knowledge.

## 8

Crucial to the view that Sidgwick has a mixed epistemology is the claim that he grants common sense independent probative status.<sup>57</sup> Crisp, Shaver, and Schultz depend on the idea that Sidgwick thinks that the claims of common sense have "imperfect certitude" or "initial authority."<sup>58</sup> But what is the basis of common-sense morality's built-in epistemic authority?<sup>59</sup> Does common sense provide us with intuitions that are not as reliable as philosophical intuitions but which are "adequate for (most) practical purposes"?<sup>60</sup>

There is good evidence for thinking that Sidgwick does not hold that common-sense morality possesses initial authority. He submits that the rules and norms of common sense require a "rational synthesis" (ME 102): "However true they may be—[they] are not self-evident to me; they present themselves as propositions requiring *rational justification* of some kind" (ME 383; italics added). When the justification of common-sense morality is at issue, he appeals to other principles for justification (ME 343). "It is still more widely held by Utilitarian thinkers that such rules [i.e. the rules of common-sense morality], however they may originate, are *only valid* so far as their observance is conducive to the general happiness ... and this is the view that I myself take of the Utilitarian principle" (ME 8; italics added; also PE 73). Were common sense to have some independent justification itself, then it would be false to say that a fully coherent arrangement of common-

<sup>57</sup>For the same view, see Brink, "Common Sense," 183, 184, 193, 194; John Skorupski, "Three Methods and A Dualism" ["Three"], in *Henry Sidgwick*, ed. Ross Harrison (Oxford: Oxford University Press, 2001), 61–81, at 64; Steven Sverdlik, "Methodology," 540, 541, 544–45, 548, 552; J. B. Schneewind, *Sidgwick's Ethics*, 260–65, 283; and Singer, "Reflective," 507, 514, 517.

<sup>58</sup>Crisp, "Boundaries," 68; Shaver, *Rational*, 70; Schultz, *Eye*, 202.

<sup>59</sup>The following is taken from Anthony Skelton, "Schultz's Sidgwick," *Utilitas* 19 (2007): 91–103.

<sup>60</sup>Skorupski, "Three," 64; also 65.

sense moral norms is no more than “an accidental aggregate of precepts” (ME 102),<sup>61</sup> and that his examination reveals that “time-honoured virtues seem to be fitly explained as special manifestations of impartial benevolence under various circumstances of human life, or else as habits and dispositions indispensable to the maintenance of prudent or beneficent behaviour under the seductive force of various non-rational impulses” (ME 496–97). In his critical notice of F. H. Bradley’s *Ethical Studies*, he remarks that if the authority of common sense is found to be “*falsus in uno*, it must be at least *fallibilis in omnibus*: and thus we have still to seek for some criterion of the validity of its dictates” (BE 548).

Is the morality presented in Book III, chapters iv–x required support for whatever moral theory one arrives at?<sup>62</sup> Sidgwick rejects this: “It should be borne in mind throughout the discussion carried on in this and the next six chapters that what we are primarily endeavouring to ascertain is not true morality but the morality of Common Sense: so that if any moral proposition is admitted to be paradoxical, the admission excludes it—not as being necessarily false, but as being not what Common Sense holds” (ME 263n.). The issue of truth enters the picture only when he endeavors to determine whether the rules of common-sense morality “possess the characteristics by which self-evident truths are distinguished from mere opinions” (ME 338).<sup>63</sup> Here he says, “Before [referring to Book III, chapters iv–x], our primary aim was to ascertain impartially what the deliverances of Common Sense actually are: we have now to ask how far these enunciations can claim to be classed as Intuitive Truths” (ME 343). This seems to suggest that it is at this point only that he is concerned to determine the truth of the edicts of common-sense morality. Since common sense fails to offer us any axioms, he claims that at best it may be “adequate to give practical guidance to common people in common circumstances” (ME 361). He does not at this time suggest that common sense offers us imperfect certitudes. The implication appears to be that edicts of common-sense morality are mere opinions, not partial truths. More explicitly, he holds that moral propositions must either be “immediately known to be true—and therefore, we may say, a moral intuition—or be inferred ultimately

---

<sup>61</sup>In the same place that Sidgwick makes this statement he remarks that philosophical intuitionism accepts “the morality of common sense as in the main *sound* ... [while still attempting] to find for it a philosophic basis which it does not itself offer” (ME 102; italics added). Is not this clearly a case where Sidgwick grants common-sense morality initial authority? Not necessarily, for he might be here equating being sound with being useful as a practical guide. He seems to do just this at various places. See ME xxi, 361.

<sup>62</sup>Schneewind, *Sidgwick’s Ethics*, 193. Schneewind claims that for Sidgwick the rules of common-sense morality possess the same degree of epistemic credibility as self-evident propositions, even though each possesses a different degree of moral validity (263–64). This seems undermined by the fact that Sidgwick thinks that the rules of common-sense morality require rational justification (ME 383) and by the fact that the main elements of common-sense morality “seem to be fitly explained as special manifestations of impartial benevolence under various circumstances of human life, or else as habits and dispositions indispensable to the maintenance of prudent or beneficent behaviour under the seductive force of various non-rational impulses” (ME 496–97).

<sup>63</sup>In ME1 Sidgwick says that his aim is to distinguish “knowledge” from opinion (ME1 317), and in ME2 he says that his aim is to distinguish “certain truths” from mere opinions (ME2 313). This seems to suggest that the fact that the avowals of common-sense morality fail to satisfy the conditions of scientific axioms entails that they are mere opinions, and hence lack epistemic justification of any sort.

from premises which include at least one such moral intuition" (ME 98; also PC 564). This seems to suggest that the source of justification is the philosophical intuitions themselves and therefore that common sense does not have independent epistemic credibility.

Finally, it is not obvious what on Sidgwick's view could provide for the epistemic credibility of common-sense morality. It is not some sort of reflection.<sup>64</sup> He finds that upon reflection the dictates of common sense require rational justification (ME 383; PC 565).<sup>65</sup> It has been suggested that the fact that a dictate is part of common-sense morality is enough to provide it with some epistemic credibility.<sup>66</sup> This is implausible. It would entail granting probative value to a proposition on no other ground than that it is what people happen (unreflectively) to believe. This does not look to be a plausible basis for belief, and there is no evidence that Sidgwick holds this view. Schultz offers that common-sense morality has initial credibility because "it represents the experience of many generations, experience suggesting some presumptive evolutionary success."<sup>67</sup> It is far from clear that this is Sidgwick's view, and Schultz provides no textual evidence to support this suggestion. Sidgwick does speak of the "general presumption which evolution afforded that moral sentiments and opinions would point to conduct conducive to general happiness" (ME xxi).<sup>68</sup> But he takes this to be a reason in favor of using aspects of common-sense morality for practical guidance, not for thinking that common-sense morality possesses imperfect certitude.

It is well known that Sidgwick relies on some of the rules of common-sense morality as a decision procedure in his indirect utilitarianism (ME 413, 460–74). When he finishes his discussion of common-sense morality, he remarks of it that it "may still be perfectly adequate to give practical guidance to common people in common circumstances" despite not offering us intuitions (ME 361; also 466). John Skorupski argues that Sidgwick is here suggesting that common-sense morality yields "intuitive knowledge, though through lack of precision it cannot qualify as 'scientific' or 'philosophical' knowledge."<sup>69</sup> But he says here only that

---

<sup>64</sup>Contra Shaver, *Rational*, 73.

<sup>65</sup>In private correspondence, Shaver has suggested to me that though common-sense morality requires rational justification, this is consistent with it possessing some kind initial epistemic warrant. But if common-sense morality lacks rational justification, what sort of justification does it possess? What sort of epistemic justification is it able to confer, if not rational justification or justification of the sort that supplies reasons that enhance the chances to an apposite degree that a certain proposition is true?

<sup>66</sup>Crisp, "Boundaries," 69.

<sup>67</sup>Schultz, *Eye*, 202.

<sup>68</sup>Shaver also contends that Sidgwick has an "evolutionary argument for the *approximate* reliability of common-sense morality" (Shaver, *Rational*, 70; italics in original). Sidgwick says that "so far as any moral habit or sentiment was unfavorable to the preservation of the social organism, it would be a disadvantage in the struggle for existence, and would therefore tend to perish with the community that adhered to it" (ME 465). But it is unclear whether Sidgwick's point here is only that this fact is a sign of the approximate reliability of common sense as a means to general happiness rather than epistemic reliability. (He seems to mean the former; see ME xxi.) He does raise this issue in the context of a discussion of what decision procedure utilitarianism should rely on. He goes on to suggest, in addition, that utilitarians should (if they were not held up by certain paradoxes) "make a thorough revision of these rules" (ME 467). His acceptance then of common sense turns out to be based on pragmatic rather than epistemic reasons (ME 467–74, xxi).

<sup>69</sup>Skorupski, "Three," 65.



common-sense morality “may” still be perfectly adequate for practice, so his claim can at most be that it “may” provide us with adequate knowledge. The only stage at which he says that indeed common sense should be our decision procedure is after he has established that it has its “basis” in utilitarianism and that it is useful because its main tenets are “unconsciously” or “inchoately and imperfectly” utilitarian (ME 424, 427; also 454, 463). This suggests that for Sidgwick common-sense morality *can* only serve as a practical guide, in so far as it conduces to the main goal of utilitarianism (ME 8, 460–61, 467). Finally, he does not appear to draw a distinction between unscientific intuitive knowledge and philosophical or scientific intuitive knowledge.

Sidgwick has of course emphasized the importance of using common-sense morality (at least provisionally) in practical ethics (ME 361, Book IV, chapters iii–v; PE 3–30). But it appears that he has different standards for what can be appealed to in practice than he has for what can be appealed to for philosophical justification (EP 109). He says, for example, that general agreement or consent is “practically the only evidence upon which the greater part of mankind can rely ... [but that] a proposition accepted as true upon this ground alone has, of course, neither self-evidence nor demonstrative evidence for the mind that so accepts it” (ME 341). Thus such a proposition lacks epistemic justification. Indeed, most of the remarks he makes about reliance on common sense refer to practical not epistemic considerations. This is certainly the case with the claims that he makes about reliance on common sense in PE, where he explicitly tells us that his concern is “not knowledge but action” (PE 5). The philosopher should always “study with reverent and patient care the Morality of Common Sense” (PE 22), and the philosopher’s *conclusions* should be “aided, checked, and controlled by the moral judgements of persons with less philosophy but more special experience” (PE 20). Common sense provides indispensable information for the application of a “philosophical” account of morality, information which “finds no place in any statement of facts or reasoned forecast of consequences that ... [common sense] could furnish; it is only represented in ... [its] judgements as to what ought to be done and aimed at” (PE 21).<sup>70</sup> But whatever is “represented” in these judgements, it appears to be no more than the information necessary for applying moral views.<sup>71</sup>

<sup>70</sup>For more on Sidgwick’s practical ethics, see Anthony Skelton, “Henry Sidgwick’s Practical Ethics: A Defense,” *Utilitas* 18 (2006): 199–217.

<sup>71</sup>It should be noted that Sidgwick was not always completely clear on the epistemic status of common-sense morality in the various versions of the *Methods*. In a very early paper entitled “Utilitarianism,” he writes that utilitarianism supports the rules of common-sense morality by providing “a further justification of them, besides the intuitive recognition of their stringency” (U 259). This passage appears with minor changes again in the first two editions of the *Methods* (ME1 393; ME2 390). In the third edition of the *Methods*, he claims again that utilitarianism gives common sense “a further justification,” suggesting that common sense has some other justification; however, he drops the claim about intuitive recognition (ME3 418). In the fourth edition of the *Methods*, Sidgwick eliminates the claim that utilitarianism provides common-sense morality with “further justification”; instead, he claims that “Utilitarianism sustains the general validity of the current moral judgements, and thus supplements the defects which reflection finds in the intuitive recognition of their stringency” (ME4 419). It seems clear that Sidgwick wanted to resist the idea that common-sense morality has some sort of initial epistemic warrant. If his view of common-sense morality is what commentators say it is, then surely this is where he should and would have set the record straight. Instead, what we find is Sidgwick moving away from the idea that common-sense morality possesses some probative status to the idea that it possesses none. Why else would he remove the “further justification” claim from subsequent editions?

A number of other passages in ME appear to cast doubt on my interpretation. Many have been used to buttress the rival interpretations discussed above. These passages also serve as the basis for Steven Sverdlik's argument that Rawls's interpretation is "the least dissatisfying understanding of Sidgwick's approach [to establishing first principles]."<sup>72</sup> Unlike Rawls, Sverdlik does provide some detailed textual evidence to support his claim that the best interpretation of Sidgwick sees him committed to a coherentist version of reflective equilibrium. Two passages are relevant. Sverdlik, like many others, thinks that the proof for utilitarianism that is provided in Book IV is coherentist in nature. I have already dealt with the passages in Book IV that this interpretation relies on.

Sverdlik relies on one other passage from Book III in which Sidgwick states that

we conceive it as the aim of a philosopher, as such, to do somewhat more than define and formulate the common moral opinions of mankind. His function is to tell men what they ought to think, rather than what they do think: he is expected to transcend Common Sense in his premises, and is allowed a certain divergence from Common Sense in his conclusions. It is true that the limits of this deviation are firmly, though indefinitely, fixed: the truth of a philosopher's premises will always be tested by the acceptability of his conclusions: if in any important point he be found in flagrant conflict with common opinion, his method is likely to be declared invalid. (ME 373)

Certain phrases in this passage can be used to support the claim that Sidgwick appeals to coherence with common-sense morality in his epistemology. He speaks of the truth of a philosopher's premises as being "tested" by appeal to common morality and of a method as being "likely" to be declared "invalid," if it is in "flagrant" conflict with common moral opinion. Singer argues that the appeal to coherence here is an illusion in light of the fact that Sidgwick did not think to reject rational egoism as a plausible position about what we have most reason to do.<sup>73</sup> The claim is that, since rational egoism conflicts to a high degree with common-sense morality, were Sidgwick relying on common sense to test moral theories, he would surely have rejected rational egoism on this basis. This he did not do, *ergo* he was not relying on common sense to test moral theories or philosophical intuitions.

This seems a hasty conclusion to draw. Sidgwick remarks that "the limits of this deviation [from common sense] are firmly, though indefinitely, fixed," and that "if in any important point ... [the philosopher] be found in flagrant conflict with common opinion, his method is likely to be declared invalid" (ME 373). Since the point at which deviation has gone too far is not clearly fixed, he might have thought that egoism was within the bounds of appropriate deviation from common sense. Therefore, rational egoism would not be considered to be in "flagrant conflict" with it.

Yet, this passage does not really support a pure coherentist or mixed view. One of Sidgwick's concerns is to remove arbitrary distinctions from ethics. At the outset

---

<sup>72</sup>Sverdlik, "Methodology," 538. Note that, officially, Sverdlik's view is that Sidgwick's epistemology is confused. His position is that if he had to choose, he would choose the Rawlsian interpretation. See Sverdlik, "Methodology," 548.

<sup>73</sup>Singer, "Reflective," 506.

of ME he argues that any view according to which one ought to aim only at the happiness of one's family-members or fellow countrymen or race as opposed to happiness in general is at bottom arbitrarily limited. "It is, of course, possible to adopt an end intermediate between the two [i.e. individual and general happiness], and to aim at the happiness of some limited portion of mankind, such as one's family or nation or race: but any such limitation seems arbitrary, and probably few would maintain it to be reasonable *per se*" (ME 10; also PE 39). He thinks, for example, that it would be "arbitrary and unreasonable" to take the promotion of human happiness as an end, while excluding non-human animal happiness (ME 414; also EP 106–7).<sup>74</sup> Indeed, he rejects all distinctions which are arbitrary or, what comes to the same thing, positions "for which no sufficient reason can be given" (ME 267–68). Returning now to the main issue, it might be thought that Sidgwick chooses utilitarianism over egoism on the grounds that it captures more of the norms of common sense than egoism.<sup>75</sup> He might have chosen the former over the latter because he thinks that the former deviates less from common sense than the latter. But, as Sidgwick rightly notes, the precise point at which deviation from common sense becomes "flagrant" is unclear and subject to dispute. Therefore, unless there can be a precise agreement regarding when "flagrant" conflict occurs, any line will, it seems, be unfairly drawn (because drawn without good reason) and hence arbitrarily exclude one or other of the positions without rational explanation. For either view might dissent from the line of demarcation, leaving no clear, non-arbitrary and non-question-begging way of resolving the dispute between the views using this method. Any demarcation line will appear "arbitrary and without foundation in reason" (EP 107). The dispute will need to be adjudicated on some other grounds.

As should be obvious by now, I think the adjudication between the two views is provided for at least in part by appeal to the intuitions that Sidgwick provides.<sup>76</sup> Of course, as I have already pointed out, this is not the only strategy that Sidgwick relies on to bring advocates of rival views over to utilitarianism (ME 420; EP 106–7).<sup>77</sup> He might use other *ad hominem* arguments, making reference to utilitarianism's clarity, comprehensiveness, and ability to solve the problem of decision that plagues the ordinary person's reasoning, and so on.

How should we deal with his other remarks in the passage that Sverdlik quotes? Sidgwick states that a philosopher's premises or method are "likely" to be declared invalid in a case of conflict, not that they ought to be or that they should be. This latter and stronger reading would be required to make the mixed or coherentist interpretation obvious. It is equally likely here that Sidgwick (given what I have

---

<sup>74</sup>Shaver thinks that "animals provide one of the very few places where he [Sidgwick] explicitly corrects, on utilitarian grounds, what some take to be common sense" (Shaver, *Rational*, 92). This is not obvious. Sidgwick notes that "Common Sense is disposed to regard this [the idea that animal pain does not matter intrinsically] as a hard-hearted paradox, and to hold with Bentham that the pain of animals is *per se* to be avoided" (ME 241). This may well be one instance in which Sidgwick exaggerated the agreement between utilitarianism and common sense.

<sup>75</sup>This seems to be Shaver's strategy; see Shaver, *Rational*, 107–9.

<sup>76</sup>Or, at the very least the adjudication should be provided for by appeal to our intuitive or reflective capacity.

<sup>77</sup>Of course, such lines of argument are open to other views as well, most notably rational egoism.

already said above) is merely putting out a word of caution for philosophers in their quest to implement their account of the most basic requirements of reason. The problem is that it is possible that philosophers might lose “general assent”—a fear which he describes as not unfounded—if “they set before themselves and their readers too rigid a standard of scientific precision” (ME 374).<sup>78</sup> This is presumably why Sidgwick is often quick to short-circuit disagreement as soon as it might arise (e.g. ME 382). But the loss of this general assent is not something that should constrain philosophers in their quest for axioms; rather, it should give them pause in attempting to apply in practice what they think is the basis of duty, that is, their method of ethics. After all, to refrain from articulating the basic demands of reason on the grounds that they may be rejected by common sense, he thinks, is unpalatable, since it leaves one to “implicitly accept the morality of Common Sense *en bloc*, ignoring its defects” (ME 374). His fears about rejection are no doubt in part why he thought that it would be prudent not to try to implement a direct form of utilitarianism (ME 468–69). Since it is likely to be rejected on the grounds that it is far too demanding, it is important to demonstrate how utilitarianism could be employed in practice in a palatable way, namely, in the context of an indirect utilitarianism. I think that this interpretation fares better than its rivals, since a) it does not ascribe an incoherent epistemology to Sidgwick, and b) it does not ascribe to him an arbitrary view regarding the point at which deviation from common sense is flagrant and is detrimental to any moral view.

Sverdlik thinks that the Rawlsian interpretation of Sidgwick's epistemology is the least problematic. His argument turns on worries associated with interpreting Sidgwick as relying solely on appeal to self-evidence for establishing moral theories.<sup>79</sup> Self-evident truths are thought by many to be certain, infallible propositions admitting of no further justification or epistemic amplification. As Sverdlik puts it, “They are [often thought to be] as certain as it is possible for a proposition to be.”<sup>80</sup> Sidgwick, of course, denies that philosophical intuitions are infallible (VB 585, 587; ME 211; EP 109). He thinks that one can be in error about the intuitive basis of a proposition. To Sverdlik's mind, this is a problematic omission, because to admit that intuitions are fallible means that Sidgwick “has no way of eliminating the possibility of delusive self-evidence.”<sup>81</sup> All he can claim is that the propositions he believes to be self-evident “are, at best, very probably self-evident.”<sup>82</sup> But, Sverdlik continues, “a proposition that is very probably absolutely certain [defeasibly so] is surely just one that is very probably true. A proposition that is very probably true is one whose certainty can be augmented.”<sup>83</sup> Therefore, he concludes, Sidgwick's self-evident propositions are propositions whose certainty can be augmented.

<sup>78</sup>For an excellent account of Sidgwick's anxieties about open and violent conflict with common-sense morality, see Schultz, *Eye*, 335–508.

<sup>79</sup>Sverdlik makes a number of specific and dubious remarks about the nature of self-evidence. He claims, for instance, that “if a proposition is self-evident it would be obvious whether it is so” (“Methodology,” 549). No argument is offered for thinking this, and it does not seem to be a necessary element of the appeal to self-evidence or to Sidgwick's view. I deal with his general remarks in what follows.

<sup>80</sup>Sverdlik, “Methodology,” 548.

<sup>81</sup>*Ibid.*, 551.

<sup>82</sup>*Ibid.*

<sup>83</sup>*Ibid.*, 552.

Sidgwick grants common sense an evidentiary role in establishing moral theories.<sup>84</sup> Hence, self-evident propositions are supported by appeal to coherence with common-sense morality.

There are two problems with this argument. First, it does not get one to the Rawlsian position, since this position officially rules out appeal to self-evident intuitions, even fallible ones. It is, even on Sverdlik's view, a version of pure coherentism.<sup>85</sup> Second, it is a *non-sequitur*. It does not follow from the fact that an apprehension of self-evidence is fallible that it is a proposition whose certainty is capable of augmentation. Sidgwick seems to think that intuitions are the highest we can expect to get in ethical inquiry. So the fact that these intuitions are not as certain as it is humanly possible to get entails only that in ethics we are not able to achieve the greatest possible certainty in our fundamental intuitions. It does not follow from the fact that the intuitions Sidgwick arrives at are fallible that they are not of the highest certainty that is possible in moral thought. In other words, it does not follow that they can be augmented by further propositions. At best, what follows is that this is the highest possible certainty we can arrive at.

Of course, Sidgwick cannot rule out the possibility of delusive self-evidence, but apparently no epistemological view can rule out with certainty that we may be deluded, including the view put forth by Rawls. Sidgwick was well aware of the problem of delusive self-evidence and the defeasibility of intuitions. It is this fact that lies behind his endorsement of the four tests which he describes as "*useful*, as a means of guarding against error" (LK 461; italics in original) and as "a valuable protection against the misleading influence of our own irrational impulses on our judgements" (ME 339). The tests are a means of excluding "the more subtle and latent effect of these in fashioning our minds to a facile and unquestioning admission of common but unwarranted assumptions" (ME 339) and for "correcting the oversights of ordinary intuition" (VB 586). The tests work to exclude "all such error as arises from the special weaknesses and biases of individual minds, or of particular sections of the human race" (EP 109).

It is clear that the tests are supposed to help screen out intuitions that are due to bias, or inattention or agent fallibility or other such sources of error. Sverdlik seems to think that so long as a proposition is apprehended as self-evident there is no other way to remove doubt about it except by reference to further justification. But Sverdlik ignores the possibility that some of the tests are "indispensable negative" conditions "of the certainty of our [self-evident intuitions]" (ME 342; also EP 109-10) and are not intended to bring us to absolute certainty but only to test the cogency of intuitions. As I see it, Sidgwick believes that if a putatively self-evident or intuited truth does not fail the tests, then doubt about it is reduced as much as is possible, even if not entirely (EP 109). But the reduction of doubt does not consist in reliance on further factors for justification of the axioms. He emphasizes the importance of having ways of insuring that the truths intuited as true actually are true. He cautions us to "assure yourself of the self-evidence of

---

<sup>84</sup>I hold this to be sufficiently undermined by arguments provided above.

<sup>85</sup>It might, however, show that the mixed view is best. My other arguments should disabuse one of this.

what appears self-evident, by careful examination" (LK 461). The tests are meant to help us achieve the highest form of assurance that we can attain.

I O

In this essay I have argued that Sidgwick's moral epistemology is neither some form of Rawlsian reflective equilibrium nor an epistemology with elements of both foundationalism and coherentism. Instead, it is a form of foundationalism according to which the epistemic justification of moral theories consists in appeal to intuitions or self-evident propositions, aided by *ad hominem* arguments and negative arguments against rivals (among other philosophical tools).<sup>86</sup>

APPENDIX I: SIDGWICK'S PHILOSOPHICAL INTUITIONS

1. "It cannot be right for A to treat B in a manner in which it would be wrong for B to treat A, merely on the ground that they are two different individuals, and without there being any difference between the natures or circumstances of the two which can be stated as a reasonable ground for difference of treatment" (ME 380).
2. "The mere difference of priority and posteriority in time is not a reasonable ground for having more regard to the consciousness of one moment that [*sic*] to that of another" (ME 381).
3. "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).
4. "As 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).
5. "Happiness (when explained to mean a sum of pleasures) ... [is] the sole ultimate end" (ME 402; also LE 107, 126–31).
6. "The greater *quantum* of pleasure is to be preferred to the less, and that *ex vi termini* the larger sum made up of less intense pleasures is the greater quantum of pleasure" (LE 110; italics in original).

---

<sup>86</sup>I wish to thank Wayne Sumner, Thomas Hurka, Arthur Ripstein, Gopal Sreenivasan, Gurpreet Rattan, Danny Goldstick, Donald Ainslie, Matt Leisinger, audiences at the 2003 International Society for Utilitarian Studies in Lisbon and at the 2003 Moral and Legal Philosophy Workshop at the University of Toronto, and three anonymous referees for the *Journal of the History of Philosophy* for helpful and insightful comments on previous versions of this article, and the Social Sciences and Humanities Research Council of Canada for kindly providing me with funding.