
Sidgwick's Argument for Utilitarianism and his Moral Epistemology: A Reply to David Phillips

Anthony Skelton



Electronic version

URL: <https://journals.openedition.org/etudes-benthamiennes/675>
DOI: 10.4000/etudes-benthamiennes.675
ISSN: 1760-7507

Publisher

Centre Bentham

Electronic reference

Anthony Skelton, "Sidgwick's Argument for Utilitarianism and his Moral Epistemology: A Reply to David Phillips", *Revue d'études benthamiennes* [Online], 12 | 2013, Online since 10 December 2013, connection on 20 May 2021. URL: <http://journals.openedition.org/etudes-benthamiennes/675> ; DOI: <https://doi.org/10.4000/etudes-benthamiennes.675>

This text was automatically generated on 20 May 2021.

Droits réservés

Sidgwick's Argument for Utilitarianism and his Moral Epistemology: A Reply to David Phillips

Anthony Skelton

- 1 David Phillips's *Sidgwickian Ethics* is a penetrating contribution to the scholarly and philosophical understanding of Henry Sidgwick's *The Methods of Ethics*.¹
- 2 This note focuses on Phillips's understanding of (aspects of) Sidgwick's argument for utilitarianism and the moral epistemology to which he subscribes. In § I, I briefly outline the basic features of the argument that Sidgwick provides for utilitarianism, noting some disagreements with Phillips along the way. In § II, I raise some objections to Phillips's account of the epistemology underlying the argument. In § III, I reply to the claim that there is a puzzle at the heart of Sidgwick's epistemology. In § IV, I respond to Phillips's claim that Sidgwick is unfair in his argument against the (deontological) morality of common sense.

I Sidgwick's Argument for Utilitarianism

- 3 One aim of *The Methods of Ethics* is to provide an argument for classical utilitarianism. For Sidgwick, this involves, in part, showing that there exist certain non-derivatively warranted propositions, i.e., intuitions, on which the truth of utilitarianism depends. He arrives at these in the penultimate chapter of Book III after a long and exhaustive survey of common-sense morality. The most important of these are:
- 4 (U): "The good of any one individual is of no more importance, from the point of view... 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 realized in the one case than in the other" (ME 382), and

- (R): "It is evident to me 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).
- 5 From (U) and (R), Sidgwick infers the maxim of benevolence: "that each one is morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him" (ME 382). This expresses the idea that one is required to give each sentient being's good weight proportionate to its quantity in one's reasoning about what to do. It is clear that for Sidgwick the maxim or principle of benevolence consists in and is equivalent to the requirement that one promote or pursue (surplus) aggregate good (ME xxi, 382, 385, 392, 400).
 - 6 Phillips is not clear on what he thinks Sidgwick gets from (U) and (R). He suggests at times that Sidgwick infers from them the bulk of utilitarianism (SE 64, 68, 102, 125). At other times, he suggests that Sidgwick infers something more like the principle of rational benevolence (SE 118). This is understandable. Sidgwick is unclear. Soon after arriving at (U) and (R) he says "Utilitarianism is thus presented as the final form into which Intuitionism tends to pass, when the demand for really self-evident principles is rigorously pressed" (ME 388; also 406-407). But Sidgwick more frequently describes what he gets from (U) and (R) as the axiom or maxim of rational benevolence or first principle of utilitarianism (ME 387, 400, 418, 419, 421). This serves as the "basis" of utilitarianism (ME 387). (U) and (R) do not alone, then, get Sidgwick utilitarianism or the utilitarian method of ethics.
 - 7 This appears to be confirmed by the fact that Sidgwick does not believe that (U) and (R) supply a method of ethics, that is, a "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). The problem is that the intuitions are "of too abstract a nature, and too universal in their scope, to enable us to ascertain by immediate application of them what we ought to do in any particular case; particular duties have still to be determined by some other method" (ME 379).
 - 8 It is important to note that Sidgwick thinks that the abstractness point applies to (U), (R) and the principle of rational benevolence. He twice describes the last as abstract (ME 382, 462n1).² It would not, then, be plausible to think that he means his comment to apply to (U) and (R) and not the principle of rational benevolence, which could serve as a method of ethics. This fits with the primary focus of the discussion in which Sidgwick lays out his intuitions, which is principles rather than methods. He does not turn to the utilitarian method of ethics until Book IV of ME.
 - 9 That Sidgwick aims to get utilitarianism's first principle from (U) and (R) and not the utilitarian method of ethics makes it possible to reply to Phillips's claim that there is a puzzle at the heart of Sidgwick's epistemology and that his argument against common-sense morality is unfair. Before articulating these replies (in §§ III and IV) it is important to outline the other features of Sidgwick's argument for utilitarianism.
 - 10 Sidgwick's argument for utilitarianism relies in addition on showing that the fundamental moral requirements of common-sense morality do not, presumably in his view unlike (U) and (R), possess the "characteristics by which self-evident truths are distinguished from mere opinions" (ME 338). To qualify as self-evident a proposition must be clear and precise, self-evident on reflection, consistent with other self-evident

propositions one accepts, and not denied by a competent judge who one believes is no more likely to be in error than oneself (ME 338-342).

- 11 Sidgwick thinks that when we reflect on the rules of common-sense morality, including the requirements of fidelity to promises, of justice and of veracity, we see that they fail to possess one or other of the characteristics of self-evidence (ME 360). In sum, his argument is that if you leave the main requirements of common-sense morality vague – e.g., you ought to keep your promises – there is little or no dissent from them. In this case, the requirements possess the fourth but lack the first characteristic. However, if you try to give them “the definiteness which science requires”, it is no longer true that there is little or no dissent from them (ME 342). Competent judges disagree on the possible specifications of the rules of common-sense morality. In this case, the requirements of common-sense morality possess the first but not the fourth characteristic. In either case, the rules of common-sense morality fail to qualify as self-evident (ME 342).
- 12 Phillips calls the foregoing argument the criterial argument for utilitarianism (SE 64, 67, 95, 98). He rightly notes that Sidgwick uses a further argument in his attempt to establish utilitarianism. This argument takes the form of a proof directed to the proponent of common-sense morality (ME 420-422), which he describes 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 principle” (ME 420).
- 13 Relying on this strategy, Sidgwick tries to show that the utilitarian method of ethics supports the main rules of common-sense morality in general and that it injects greater clarity, completeness and system into ethical thinking, thereby remedying the practical “defects” which Sidgwick allegedly finds in common sense (ME 422). He thinks that this furnishes the proponent of common-sense morality with considerations sufficient to determine her mind to accept utilitarianism. Phillips refers to this argument as Sidgwick’s bi-partite argument for utilitarianism (SE 63).

II. Sidgwick’s Moral Epistemology

- 14 To understand and reply to Phillips’s claim that there is a puzzle at the heart of Sidgwick’s epistemology we must get clear on its nature.
- 15 It is relatively clear that Sidgwick’s argument for utilitarianism relies on an appeal to an intuitionist epistemology. He thinks that (U) and (R) are, for example, self-evident, that is, non-inferentially warranted propositions or intuitions (ME 382, 383, 387, 421).
- 16 Phillips agrees that the appeal to self-evidence plays an important role in Sidgwick’s account of the justification of (U) and (R) (SE 60-65). It is not the only thing to which Sidgwick appeals, however. Phillips maintains that in the criterial argument Sidgwick relies in addition on coherence considerations to justify (U) and (R) (SE 61-62, 79). (U) and (R) are both inferentially and non-inferentially justified (SE 81). Sidgwick is therefore a proponent of what Phillips calls “moderate foundationalism” (SE 80).
- 17 Phillips’s case for this interpretation depends in part on a particular understanding of Sidgwick’s characteristics of self-evidence. Sidgwick describes these as “four conditions, the complete fulfillment of which would establish a significant proposition,

apparently self-evident, in the highest degree of certainty attainable" (ME 338). The most important of these for our purposes is the fourth condition, which, on Phillips's reading, states that for a proposition to be self-evident it "must be universally accepted" (SE 60). In passing this test, a proposition gains epistemic credibility. In satisfying the test a proposition acquires "consistency with ordinary moral opinions", which themselves have "probative" value (SE 79, 84).

- 18 The difficulty with this suggestion is that it is in tension with a number of other things that Phillips says. At one point he claims that Sidgwick offers "a conception of criterial argument which gives no evidential role to common-sense morality" (SE 76). This might just be an oversight on Phillips's part, but actually this seems to be the view that he is forced to accept given his interpretation of Sidgwick's bi-partite argument for utilitarianism. The received view of the bi-partite argument is that Sidgwick grants probative value to common-sense morality and that he thinks that utilitarianism's ability to capture and explain key elements of it provides one with justification for believing it. Phillips rejects this reading of the argument in favor of the view that the bi-partite argument is an *ad hominem* argument directed to the proponent of common-sense morality in which Sidgwick grants no evidential role to common-sense morality (SE 74-76). This suggests that Phillips's view is that Sidgwick grants probative status to common-sense morality in the context of his criterial argument but not in the context of the bi-partite argument. This is puzzling, however. Why grant probative status or evidential value to common-sense morality in the case of the criterial argument but not in the case of the bi-partite argument? The puzzle is deepened by the fact that Phillips grants that common-sense morality might, in the end, possess "imperfect certitude" (SE 84).³
- 19 Phillips has two possible replies.
- 20 (1) He might argue that when Sidgwick appeals to "ordinary moral opinions" he is appealing not to the moral beliefs of the plain man but to the moral beliefs of "(well-informed) [moral] experts" (SE 79; cf. SE 61).
- 21 There are two difficulties with this reply. First, some of the so-called moral "experts" or "competent judges" to which Sidgwick appeals are themselves exponents of common-sense moral opinions (e.g., Whewell and Kant). Second, in the discussion of (U) and (R) Sidgwick is keen to note that the "plain man" does not disagree with the principle or maxim of rational benevolence (ME 382; also LE 331-332).
- 22 (2) He might avoid this charge by exploiting a distinction that he makes between common-sense morality broadly construed and common-sense morality narrowly construed (SE 78). He describes the first as consisting of "ordinary moral opinions", including "those best systematized by dogmatic intuitionism and those best systematized by [egoism and utilitarianism]" (SE 79). The second is described as a "subset of common moral beliefs that Sidgwick systematizes as dogmatic intuitionism" (SE 78). The idea might be that Sidgwick appeals to the former but not the latter in the criterial argument.
- 23 There are two infelicities with this suggestion. First, it is not obvious that Sidgwick cuts the distinction between common-sense morality narrowly construed and common-sense morality broadly construed. Phillips relies on scant evidence to show that this is Sidgwick's view. The passage that he does cite to justify his interpretation seems to list things – concern for the general good, prudence, and so on – which find a firm home in common-sense morality narrowly construed (SE 79; see ME 14, PC 564). The situation is

made worse for Phillips by the fact that he appears to give no indication that Sidgwick thinks common-sense morality broadly construed has probative force. (At least in the case of common-sense morality narrowly construed there is some evidence that Sidgwick thinks it has evidential value (see, e.g., ME 373).)

- 24 Second, it does not seem possible to think that while common-sense morality narrowly construed has no probative status, common-sense morality broadly construed does. The latter includes the former (SE 79). In addition, the most obvious arguments (appeals to agreement, for example) provided for thinking that common-sense morality broadly construed has probative status seem to imply that common-sense morality narrowly construed has it, too.
- 25 Perhaps Phillips's main concern is simply to show that common-sense morality narrowly construed has no "special" epistemic role in justification (SE 79). But this is consistent with it having some epistemic role, and this is all that is needed to secure the received view of the bi-partite argument.
- 26 Phillips's problems can be traced back to a misunderstanding of the fourth of Sidgwick's "characteristics" of self-evidence. In his attempt to separate self-evident truths from mere opinions Sidgwick is not *pace* Phillips concerned to show that the justification of self-evident propositions is amplified when it possesses the characteristic of being agreed to by others. It would after all be odd for Sidgwick to appeal to Kant for justification in this regard when he rejects so many of Kant's other views (see, e.g., ME 209-210, 222-223, 389-390, 394-395).
- 27 Rather, Sidgwick's idea seems to be that dissent functions as a potential defeater of an intuition, perhaps explaining why he describes the absence of dissent as an "indispensible negative condition of the certainty of our beliefs" (ME 342). On this view, a proposition qualifies as self-evident only when rational disagreement about its truth is absent or rationally explained away. A proposition lacks self-evidence, not when it lacks universal acceptance, but when there is disagreement about its truth and where "I have no more reason to suspect error in the other mind than in my own" (ME 342). Sidgwick is worried about disagreement that cannot be rationally explained away, because disagreement, when undefeated, impugns a proposition's claim to self-evidence. This need not involve holding the claim that agreement enhances probative status.
- 28 Of course, Phillips argues that his account of the tests is supported by Sidgwick's discussion of similar tests in his writings on epistemology. But this provides inconclusive results. In these writings, Sidgwick says only that a proposition counts as self-evident when it "does not conflict with the beliefs of other persons competent to judge" (LK 465; SE 82).

III. The Puzzle in Sidgwick's Moral Epistemology

- 29 Phillips maintains that there is a puzzle at the core of Sidgwick's moral epistemology. He holds that Sidgwick relies on both the criterial argument and the bipartite ad hominem argument to establish ethical first principles (SE 65-69).⁴ Sidgwick offers an *independent* criterial argument in ME III.xiii, only to abandon it in ME IV.ii, where he offers only the ad hominem bipartite argument. The puzzle is this: why does Sidgwick

abandon the *independent* criterial argument in IV.ii where it is “clearly relevant” (SE 71; also 69)?

- 30 Phillips considers and rejects a solution according to which the independent criterial argument drops out in IV.ii because the context makes only bipartite arguments relevant (SE 73). This seems to me to be the most persuasive solution, however.
- 31 As noted, in the context of the criterial argument Sidgwick aims to establish the first principle of utilitarianism, and that there are no other such competing propositions within common-sense morality. He thinks that the proponent of common-sense morality accepts utilitarianism's first principle (ME 382, 421), and that competent judges accept it, too, e.g., Kant and Clarke (ME 384-385).
- 32 The criterial argument does not establish that the correct method of moral reasoning is utilitarian. Indeed, it does not appear by Sidgwick's own lights to *directly* or *immediately* establish any method of moral reasoning (ME 379). The criterial argument cannot, then, alone establish a method of ethics. This is where the ad hominem bi-partite argument enters the picture. In the context of the bi-partite argument, Sidgwick is trying to establish the utilitarian method of ethics. The common-sense moralist disagrees, of course, with utilitarianism (LE 331-332). It is unsurprising, then, that in IV.ii where the utilitarian method of ethics is the focus he attempts to show that there are considerations that determine the mind of the common-sense moralist to accept the utilitarian method.⁵
- 33 In light of what the arguments are designed to do, it is not at all puzzling that Sidgwick would deploy two arguments and that he would not rely on an independent criterial argument in Book IV. In Book IV the focus is on establishing a method of moral reasoning, and the criterial argument, Sidgwick admits, does not alone establish this. He thus needs another argument to get from the principles that he thinks are self-evident to the utilitarian method of ethics.
- 34 This solution to the puzzle stands even when we agree with Phillips that the criterial argument does not secure the truth of the hedonism or the truth of the maximizing conception of rationality that Sidgwick's utilitarianism includes (SE 68, 97, 102, 111n7, 118). If Sidgwick does establish hedonism and maximization in some way other than by appeal to intuition, this is not sufficient (with (U) and (R)) to establish utilitarianism. This would show only that the maxim of benevolence amounts to the requirement that one maximize surplus pleasure for the aggregate. Sidgwick would presumably still have to establish the “method of utilitarianism”, which is the focus of ME IV.ii (ME 422). In addition, at least in the case of hedonism, a bi-partite argument is unnecessary since an argument is provided for the view before the bi-partite argument is presented (ME 391ff.). Indeed, Sidgwick says that the first principle of utilitarianism that he is working with at the outset of the bi-partite argument is that we ought to aim at “universal happiness” (ME 418).

IV. Sidgwick's Argument Against Common-sense Morality

- 35 Phillips holds rightly that Sidgwick's criterial argument involves showing that (U) and (R) do, but that the main rules of common-sense morality do not, possess the “characteristics by which self-evident truths are distinguished from mere opinions”

(ME 338; see SE 95). Sidgwick is especially keen to show that the rules of common-sense morality lack the first characteristic of self-evidence, namely, clarity and precision in practical direction (ME 342-343; ME 360-361). This is a serious defect in an ethical method: “we study Ethics...for the sake of practice: and in practice we are concerned with particulars” (ME 215; cf. viii).

- 36 Phillips charges that in this argument Sidgwick is ultimately unfair to his opponent. He demands that the main rules of common-sense morality be “made precise enough to give determinate verdicts in every case” (SE 101). He makes no such demand of his own principles, however. The demand appears not to apply to his principles because these are “of too abstract a nature, and too universal in their scope, to enable us to ascertain by immediate application of them what we ought to do in any particular case; particular duties have still to be determined by some other method” (ME 379).⁶ Phillips argues that it is unfair of Sidgwick to ask of common-sense morality that it furnish us with practically precise principles but not to ask the same of, for instance, (U) and (R) (SE 101-103).
- 37 This is a strong argument. One possible reply for Sidgwick is to admit that his philosophical intuitions are not themselves precise enough to provide determinate ethical verdicts in every case, but deny that this impugns them. His intuitions are of principles, not methods, and there is no requirement that principles deliver determinate verdicts in every case *immediately*. Phillips himself grants that a proposition can be self-evident without being immediately practical. He claims that Sidgwick's case for egoism depends on the truth of the following claim: “if the distinction between any one individual and any other is real and fundamental, then “I” ought to be concerned with the quality of my existence as an individual in a sense, fundamentally important, in which I ought not to be concerned with the quality of the existence of other individuals” (SE 127; also ME 498, FC 484). This principle, Phillips maintains, is “self-evident (intuitively compelling)” (SE 129), but, like (U) and (R), it is not precise enough to provide determinate ethical injunctions in every case. Perhaps, then, Sidgwick thinks (U) and (R) are self-evident in this sense.
- 38 The problem with this reply is that it does not demonstrate the superiority of Sidgwick's principles to the rules of common-sense morality. For the proponent of common-sense morality might attempt to argue that (at least some of) its rules are self-evident in the same way as the principles (U) and (R) are. In this case, the proponent of common-sense morality might suggest that, e.g., the claim that “I ought to keep my promises”, is a self-evident in the same way as (U) and (R) and that we should accept it as such even though it cannot “be made precise enough to give determinate verdicts in every case” (SE 101). It does, after all, like (U) and (R), have some practical value.
- 39 Sidgwick might, then, try a different reply to Phillips. He might grant again that his principles are not *immediately* clear and precise. They are in this sense no better than the rules of common-sense morality. However, unlike the rules of common-sense morality, Sidgwick might argue, his principles are *mediately* clear and precise. His principles satisfy the second condition of self-evidence *mediately* rather than *immediately*. This is because (U) and (R), together with other arguments, point to a *method of ethics* that is on the face of it precise enough to give determinate verdicts in every case, namely, utilitarianism. The principles (U) and (R) together with arguments about the good and with ad hominem arguments point to the method of utilitarianism which is (at least in principle) capable of delivering determinate ethical verdicts in

every case. This makes Sidgwick's principles superior to the rules of common-sense morality; they are mediately applicable.

- 40 The problem with the rules of common-sense morality is that unlike (U) and (R) there is no route from them to a method of ethics that is capable of delivering determinate verdicts in every case. Unlike (U) and (R) the main rules of common-sense morality do not seem to point to a method of ethics beyond themselves that is capable of delivering determinate verdicts in every case while at the same time relying on the resources found in common-sense morality itself. The difficulty for common-sense morality is that the main rules that it offers are put forward as though they are to function as a method of ethics. It offers them as immediately self-evident. Sidgwick might be right that these fail to be immediately self-evident. He cannot use this against them. His own principles fail to be immediately self-evident. He can, however, use the fact that these rules also fail to be mediately self-evident. His own principles do not, it seems, have this failing, and so they are a better approximation of clarity and precision.
- 41 There are two worries about Sidgwick's reply.
- 42 (1) It is not clear that the utilitarian method of ethics is capable of delivering determinate verdicts in every case. Sidgwick himself notes that there are problems with utilitarian moral reasoning (ME xxiii, 87, 140-150, 414, 461-462). The main problem is that we lack the epistemic power and resources to figure out the short and long term outcomes of the institutions, rules and actions that we seek to evaluate.
- 43 The only reply Sidgwick has to this (self-inflicted) criticism is to argue that the problems that affect the utilitarian method of reasoning are more tractable than those that affect the common-sense method of moral reasoning. He might claim that we know almost nothing of how to solve the problems that Sidgwick points out for the requirement of (say) justice without appeal to utilitarianism or other (non-common-sense-based) principle, whereas we seem to be relatively clear at least on how to solve the problems with the acquisition of empirical information and for overcoming biases that plague utilitarian forms of moral reasoning. These problems more closely mirror those that are encountered in other domains of inquiry. In this case, Sidgwick might argue that his principles are closer approximations to clarity and precision than the rules of common-sense morality.
- 44 (2) Sidgwick must establish more clearly that common-sense morality does not possess the resources within itself for the purpose of dealing with its own practical defects. He will have to say more about what counts as common-sense morality, such that there is no element of common-sense morality that might help with its own practical defects, and about how he understands the alleged defects in common-sense morality.

NOTES

1. Abbreviations: SE = David Phillips, *Sidgwickian Ethics* (Oxford: Oxford University Press, 2011); ME = Henry Sidgwick, *The Methods of Ethics*, seventh edition (Indianapolis: Hackett, 1981); PC = "Professor Calderwood on Intuitionism in Morals," *Mind* 1 (1876), 563-566; EP = "Establishment of

Ethical First Principles," *Mind* 4 (1879), 106-111; LK = *Lectures on the Philosophy of Kant and Other Philosophical Lectures and Essays*, ed. James Ward (London: Macmillan, 1905); LE = *Lectures on the Ethics of T. H. Green, Mr. Herbert Spencer, and J. Martineau*, ed., E. E. Constance Jones (London: Macmillan, 1902); FC = Henry Sidgwick, "Some Fundamental Ethical Controversies," *Mind* 14 (1889), 473-487.

2. He also frequently says that the principle of rational benevolence is an intuition (ME 388, 400, 421, and 462n1). This is misleading. Officially it is represented as an inference from (U) and (R) (ME 382).
 3. It does this by possessing one of the characteristics of self-evidence (SE 84).
 4. Phillips finds support for this at EP 106-107.
 5. He concedes that he cannot convince the egoist (ME 420-421).
 6. He must mean some method rather than some other method.
-

AUTHOR

ANTHONY SKELTON

University of Western Ontario