

Sidgwick on Virtue

Robert Shaver

Department of Philosophy

University of Manitoba

bshaver@cc.umanitoba.ca

ABSTRACT

Sidgwick's arguments for hedonism imply that virtue is not a good. Those arguments seemed to many wholly unpersuasive. The paper analyzes them, focusing also (especially in the final Appendix) on many changes Sidgwick made on chapter XIV of Book III through the various editions of the *Methods*. From an analysis of the first sections of this chapter, it emerges that Sidgwick employed two different argumentative schemes, one against the view that virtue is the sole good and the other against the much more diffused claim that virtue is one of the goods. These arguments can be fully understood in the context of Sidgwick's general claim that only "desiderable conscious life" is good. Sidgwick's general point is that virtue, insofar as it is valuable as an end, is so because of the feelings or consciousness associated with it.

Sidgwick's arguments for consequentialism seem, for a time, to have been wholly persuasive. Until Prichard's "Does Moral Philosophy Rest on a Mistake?" there are at best few deontologists, and even after Prichard, deontology did not revive until Carritt and Ross.

Sidgwick's arguments for hedonism seem to have been almost wholly unpersuasive.¹ Both ideal utilitarians and their deontological opponents agree that there are intrinsic goods other than pleasure. Virtue is seen not only as good, but as the most important good. This is the view of Hayward, Rashdall, Prichard, Ross, Carritt and Ewing.² Moore, though less enthusiastic, agrees that virtue is

¹ Hayward writes that "Sidgwick has done for [hedonism] what Plato did for his idealistic metaphysics, he has shown that the opposing arguments are almost — if not quite — as strong as the arguments in its favour" (F. H. Hayward, *The Ethical Philosophy of Sidgwick* (London: Swan Sonnenschein, 1901) p. 226).

² See, for example, Hayward, Philosophy ch. 8; Hastings Rashdall, *The Theory of Good and Evil* (London: Oxford University Press, 1924) v. i pp. 64-5, 71-3, 75-6, 94, 100-1, 267, *Ethics* (London: T. C. and E. C. Jack, 1913) pp. 27, 51, 64-6, 70, 72, "Professor Sidgwick's Utilitarianism," *Mind* o.s. 10, 1885; H. A. Prichard, *Moral Writings* (Oxford: Clarendon, 2002) pp. 11-12, 55-6, 61-2, 99-100 (later he claims that virtue is the only good (p. 173; Prichard to Ross,

at least one good.³ It is, then, worth examining Sidgwick's arguments against virtue as a good, to see where, if anywhere, he went wrong.

There is another reason to look at these arguments. Their chapter (III.XIV) of the *Methods* — “the most important chapter” — went through many changes through different editions.⁴ The result is a bit of a mess — hardly the “pure white light” for which Sidgwick is famous.⁵ In the *Appendix*, I document these changes.

One preliminary: Different proponents of virtue mean slightly different things by “virtue.” In *The Right and the Good* Ross thinks of virtue as the possession of certain desires, especially “the desire to do one's duty, the desire to bring into being something that is good, and the desire to give pleasure or save pain to others.”⁶ Prichard and Carritt separate “virtue” — desires such as the desire to help others out of sympathy, or to act courageously “from a sense of shame at being terrified,” without thought of duty — and “moral goodness,” the desire to do one's duty (where one thinks of the action *as* one's duty).⁷ Thus Carritt defines virtuous dispositions as “those which lead people to do impulsively and effectively what reflection would generally or often show to be obligatory.” Sympathy is his main example.⁸ (I shall follow Ross in grouping both sorts of desire as “virtue.”⁹) In *Foundations of Ethics* Ross includes, in addition to desires, acts of will, emotions such as satisfaction at the pleasure of another or sorrow at her pain, and “character,” the state underlying these desires, willings, and emotions which exists even when they do not.¹⁰ Rashdall includes one's will, desires (to do one's duty or to help others), feelings, emotions, and moral beliefs. Moore includes “a love of some intrinsically good consequence which [one] expects to

Bodleian Library, Oxford, Ms. Eng. Lett. d. 116, December 20, 1928)); W. D. Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930) pp. 134, 150-4, *Foundations of Ethics* (Oxford: Oxford University Press, 1939) pp. 275, 283-4, 290-2; E. F. Carritt, *Ethical and Political Thinking* (Oxford: Oxford University Press, 1947) pp. 66, 83, 85-6, 90; A. C. Ewing, *Ethics* (New York: Free Press, 1953) pp. 46-7, 61.

³ G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903) pp. 177-9, 217-19, *Ethics* (New York: Oxford University Press, 1965) p. 102.

⁴ Hayward, *Philosophy* p. 220.

⁵ Brand Blanshard, “Sidgwick, the Man” *Monist* 58, 1974, 349.

⁶ Ross, *Right* p. 134.

⁷ Prichard, p. 16. See also pp. 55-6, 61-2, 154, 160, 216, 218.

⁸ Carritt, p. 85.

⁹ Ross, *Right* p. 161.

¹⁰ Ross, *Foundations* pp. 290-3.

produce by his action or a hatred of some intrinsically evil consequence which [one] hopes to prevent by it” and “the emotion excited by [the thought of] rightness.”¹¹ Ewing includes willings, emotions such as love, and attitudes to others (such as displayed in fairness).¹²

In his initial discussion of virtue, Sidgwick has much the same view. Virtue is “a quality of the soul or mind.”¹³ It is manifested in volitions and (for some virtues, such as gratitude, benevolence, and purity) in emotions or feelings (222-3, 226). One’s motive can be love of virtue or duty, or certain natural affections, such as humility or spontaneous sympathy (223, 225, 226). For some virtues, such as justice and veracity, we do not require either a thought of duty or an emotion, but rather just a “settled resolve to will” (224).

1. In III.XIV, Sidgwick begins (§ 1.) by rejecting the view that “‘General Good’ consists solely in general Virtue,” or that “Virtue...constitute[s] Ultimate Good,” or is “the sole Ultimate Good” (392, 394, 395). This “ — if we mean by Virtue conformity to such prescriptions and prohibitions as make up the main part of the morality of Common Sense — would involve us in a logical circle; since we have seen that the exact determination of these prescriptions and prohibitions must depend on the definition of this...Good” (392). Sidgwick takes himself to have established that the relevant prescriptions and prohibitions are rules concerning the production and just distribution of goods. We do not know the content of the rules without knowing what is good; being told that what is good is just conformity to the rules is useless.

¹¹ Moore, *Principia* pp. 177, 179. Moore thinks the latter of small value when it lacks the hatred found in the former (218-19).

¹² Ewing, *Ethics* pp. 47, 61, 67, *Second Thoughts in Moral Philosophy* (New York: Macmillan, 1959) pp. 106, 132, 134-5, 140.

¹³ Henry Sidgwick, *The Methods of Ethics* (Indianapolis: Hackett, 1981) p. 222. Subsequent parenthetical references are to the seventh edition of *The Methods*. Parenthetical references to earlier editions give the edition followed by the page. The first edition came out in 1874, the second in 1877, the third in 1884, the fourth in 1890, and the fifth in 1893 (all London: Macmillan). Other works by Sidgwick cited are *Outlines of the History of Ethics* (OHE) (London: Macmillan, 1896), “Some Fundamental Ethical Controversies” (FC), *Mind* o.s. 14, 1889, and *Lectures on the Ethics of T. H. Green, Mr. H. Spencer, and J. Martineau* (GSM) (London: Macmillan, 1902).

Against this argument — call it the circle argument — one could deny that virtue is conformity to rules. One might claim, for example, that to be virtuous is to have certain emotions, or a certain will, or certain knowledge, or a certain disposition. As long as these other things could be specified without introducing other goods, the circle argument is evaded.

Sidgwick considers most of the proposals just suggested.

(i) He admits (§ 2.) that if the virtuous person is simply one with a will to do what she takes to be right, the circle argument fails (394). But he objects, plausibly, that (a) we think some course of action *is* right, and other goods are needed to explain what makes it right, and (b) we think the will to do what one takes to be right is not always the will one ought to have, implying that there are other goods that limit it (394-5). (b) is not conclusive: one could hold that this will is not always the will one ought to have, without thinking that the reason is that other things are good. For one might think, with the deontologists, that an ought-claim such as “one ought not to will to do what one takes to be right” can be justified without relying on any claim about the good (though Prichard, Carritt and Ross themselves would not make this particular ought-claim). I consider deontology below.

(ii) Sidgwick argues that we value certain types of dispositions only because of the feelings or actions that realise them (393-4). He might be wrong here: we might value a disposition, say, to feel sorrow at the pain of others even if one never encounters anyone in pain and so never feels the sorrow. We might think a person who would feel sorrow is better than a person who would not, even if neither encounters a case that makes their difference manifest. But many will see no value in a useless disposition. And it would be odd to value *only* the disposition, and that is what would be needed to avoid the circle argument.¹⁴

(iii) Sidgwick does not consider explicitly the suggestion that virtue consists in having certain emotions. At times, he writes as if virtue is strictly a matter of “conduct” or “action” (395, 396). But he does mention feelings when discussing the character suggestion, and, as noted, feelings are prominent in the earlier discussion of some of the virtues. Presumably, however, he could again argue that

¹⁴ Sidgwick might also note the current scepticism about the existence of these traits or dispositions inspired by situationist psychologists. For discussion of the literature and its upshot for virtue, see Gilbert Harman, “Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error,” *Proceedings of the Aristotelian Society* 99, 315-31 and John M. Doris, *Lack of Character* (Cambridge: Cambridge University Press, 2002).

it would be odd to value *only* feelings. A virtuous person does not only feel sorrow at the pain of others, but acts to alleviate the pain in at least some cases. Moreover, the obvious explanation for why sorrow at pain is good is that pain is bad.

(iv) The proposal that virtue is knowledge is historically important, since it is made by the Stoics, the most prominent defenders of the view that virtue is the only good. Sidgwick objects with another application of the circle argument: if, as he suggests Plato, Aristotle, and the Stoics thought, the knowledge in question is knowledge of what is good, no guidance is forthcoming (376-7, OHE 76).¹⁵ He notes, however, that the Stoics could reply that the knowledge is knowledge of what is to be preferred or rejected, and that what is to be preferred or rejected can be learned by observing how Nature has designed us (378n1, OHE 79-80). For example, Nature has designed us not to mutilate ourselves, hence self-mutilation is to be rejected. Against this, Sidgwick does not make the standard objection that speaking of what is to be preferred or rejected is just another way of speaking of what is good or bad. Nor does he object (though he surely would) to the theological beliefs he argues are needed to give the appeal to Nature normative force (81, OHE 77-9). Instead, he objects that the appeal to Nature's design does not provide consistent guidance: on one interpretation, it recommends rejecting what is "artificial and conventional;" on another interpretation, it recommends accepting what is established; and the Stoics did not show the superiority of one interpretation to the other (378n1, OHE 81-2). Similarly, he earlier objects that we need guidance when we have conflicting impulses, so one must be able to identify the impulses whose selection counts as conforming to Nature. But there is no way to do this: neither the impulses that are most common nor first nor independent of human action are plausible candidates for impulses that ought to be followed (81-2).¹⁶

Sidgwick has a further argument against taking knowledge as the good. Later in III.XIV, he will argue that knowledge is not even one good. I consider this argument below. If it works, it would also rule out the specific sort of knowledge the Stoics valued.

¹⁵ For ancient versions of this objection to the Stoics and consideration of Stoic replies, see Gisela Striker, *Essays on Hellenistic Epistemology and Ethics* (Cambridge: Cambridge University Press, 1996) pp. 238-9, 217-219, 320-324.

¹⁶ Sidgwick does not mention the Stoics here, but the problem of selecting impulses when they conflict was raised, and not solved, by the Stoics and their critics. For discussion, see Striker, pp. 219, 258-261.

A second strategy against the circle argument is to admit, as the suggestions above do not, that virtue is at least in part conformity to rules, but to argue that the rules do not need an account of the good to make them precise. For example, one might hold, with the Stoics, that the rules depend only on an account of what is to be preferred or rejected.¹⁷ Or one might hold, with deontologists, that since at least some rules are not justified by appeals to the good, they can be made precise without such an appeal. I have discussed the Stoic suggestion: Sidgwick seems to admit that this evades the circle argument; he argues that it falters when it tries to give an account of what is to be preferred or rejected. But the deontology strategy requires comment.

Sidgwick does not consider the deontology strategy, presumably because he takes his earlier discussion of common sense morality to have shown the need for an account of the good. But even without this earlier argument, he could note that deontologists such as Ross give a role to the good. Ross subsumes four of his seven prima facie duties under the general duty of promoting the good. And the duties that Ross argues cannot be subsumed, such as the duty to keep promises, are independent of the good only in the sense that it can be right to fulfill them even when an alternative action would produce more good. Considerations of the good still enter into deciding whether one has an obligation to keep a particular promise. Thus Ross holds that whether a promise is binding depends on unspoken qualifications, and these qualifications seem to specify that keeping the promise will still bring about the good foreseen at the time of making it (or at least some good). For example, I am not bound to keep a promise to replace a string on a fiddle of someone about to die (or who no longer wants it replaced), but I am bound to keep a promise to make good the financial loss caused by my breaking a string on the dying man's fiddle, since the dying man's heirs would otherwise lose.¹⁸ Further, how easily the prima facie duty to keep a promise can be overridden by other duties depends in part on the value of the promised service to the promisee.¹⁹ More generally, Ross makes it a neces-

¹⁷ Rashdall suggests this as a reply in defence of "Green's Stoicism," though Rashdall himself rejects it ("Utilitarianism" 206-8). Tom Hurka suggests it, noting Rashdall, for the Stoics (Thomas Hurka, *Virtue, Vice, and Value* (New York: Oxford University Press, 2001) p. 9). Sidgwick himself writes that "the Stoic distinction between *Good and Evil*, and *Preferred and Rejected*, is very much wanted by Green," though it is not clear that he has the circle argument in mind (GSM 99).

¹⁸ Ross, *Foundations* pp. 94, 110; also 95-7.

¹⁹ Ross, *Foundations* p. 100.

sary condition on the performance of duty that some good is produced.²⁰ Ross can say all this without failing to be a deontologist, since it remains true that sometimes one ought to produce less good rather than more; but (as I think he would admit) the good still has a role in making precise even duties that are not duties to produce the most good.

Sidgwick does mention what could be considered an example of the deontology strategy. He notes that “qualities commonly admired, such as Energy, Zeal, Self-control, Thoughtfulness, are obviously regarded as virtues only when they are directed to good ends” (392-3). One might disagree: I think we sometimes do treat these qualities as virtues even when they are directed to bad ends. Some admire these qualities in, for example, criminals. But Sidgwick could reply that our admiration is probably limited by consideration of the good — our admiration for a criminal’s zeal may diminish if, say, he is a mass murderer rather than one of Ocean’s Eleven.²¹ And even if this is not so, it would be implausible to think that *all* virtues are independent of the good.

I conclude that the circle argument, while not as clear-cut as it might appear, can be defended with various follow-up arguments. But the biggest objection concerns its target: it hurts only those who take virtue to be the sole good. Although this might be the position of Green and (perhaps) Bradley, and Sidgwick directs the circle argument against Green, it is not the “very commonplace” position of Rashdall, Ross, etc. (GSM 73-6).²² The view that “the character realised in and developed through Right conduct [...] is the sole Ultimate good [...] is not implied in the Intuitional view of Ethics: nor would it [...] accord with the moral common sense of modern Christian communities” (3). “[I]t is not commonly held that the whole Good of man lies in...obedience to moral rules” (391). Sidgwick realises he needs, and goes on to give, a different argument against the view that virtue is one good.

Despite this, Sidgwick is sometimes misleading. After giving the circle argument and considering replies, such as taking the virtuous person to be one who

²⁰ Ross, *Right* p. 162.

²¹ Anthony Skelton noted that this is not Sidgwick’s actual reply. Sidgwick thinks our admiration is “quasi-moral” and that “we certainly should not call them virtuous” (219).

²² Rashdall, “Utilitarianism” 208. It is worth noting, however, that some seem to assume that there is but one good — this characterises the debate between Hayward and E. E. Constance Jones (Hayward, “The True Significance of Sidgwick’s ‘Ethics,’” “A Reply to E. E. Constance Jones,” and Constance Jones, “Mr. Hayward’s Evaluation of Professor Sidgwick’s Ethics,” *International Journal of Ethics* 11, 1900-1, 175-87, 360-5, 354-60).

simply wills to do what she takes to be right, he concludes that “reflection shows that [virtues and talents] are only valuable on account of the good or desirable conscious life in which they are or will be actualised, or which will be somehow promoted by their exercise” (395). If this is intended as a restatement of the point that types of character are valuable only because of the feelings or actions that realise them, it is unobjectionable.²³ But Sidgwick immediately starts the next section by writing that “particular virtues and talents and gifts are largely valued as means to ulterior good,” as if this has been established (396). Again, if this is intended as a restatement of the point about character, it is fine — but in both places one has the impression that Sidgwick thinks he has shown something more, namely that the will, desire, and emotion involved in virtue are valuable only as means to something else. Hence the discussion on 396 concerns whether something can be valuable as both means and end; this seems to assume not just that types of character are valuable only because of the feelings or actions that realise them, but also that it has been shown that the feelings or actions are largely valued as means. The circle argument does not show that.²⁴

One possibility is that Sidgwick takes the circle argument to show the presence of another good, and then that, when one looks at the other good, one sees that one finds the virtue good only when it produces this other good. The circle argument does not by itself entail that virtues are valued only as means, but it is the first step in suggesting this.²⁵ Thus after giving the circle argument, Sidgwick writes that

²³ That it is a restatement of this point is clearer in earlier editions. In the third edition, the puzzling claim from 395 is part of the paragraph rejecting dispositions (3.393). In the fourth edition, the paragraph starts with “what has been before said of Virtue regarded as a quality or element of character,” where what has been before said is that dispositions are not ultimately good (4.396). The criticism of virtues and talents is that “reflection shews that they are really conceived as potentialities not valuable in themselves” (4.397).

²⁴ The problem may result from careless revision. The misleading claim quoted from 395 was introduced as part of a paragraph in the second edition that followed the circle argument (2.365; also 3.393). In the second edition, Sidgwick took the circle argument to show that virtue is not even one good. There it made sense to say that virtues are only instrumentally good.

²⁵ I owe this reading to Joyce Jenkins.

our notions of special virtues [...] contain...the same reference to ‘Good’ [...] as an ultimate standard. This appears clearly when we consider any virtue in relation to the cognate vice [...] into which it tends to pass over when pushed to an extreme [...]. For example, Common Sense may seem to regard Liberality, Frugality, Courage, Placability, as intrinsically desirable: but when we consider their relation respectively to Profusion, Meanness, Foolhardiness, Weakness, we find that Common Sense draws the line [...] by reference [...] to the general notion of ‘Good’ (392).

My own answer to the question [...] Why is the ultimate good [...] held to be pleasure? is, that nothing but pleasure appears to the reflective mind to be good in itself, without reference to an ulterior end; and in particular, reflection on the notion of the most esteemed qualities of character and conduct shows that they contain an implicit reference to some other and further good (GSM 107).

Noting the reference to good shows nothing about virtue’s intrinsic desirability or whether it is good in itself. But if once one sees the good, one also sees that this good determines whether the virtue is valuable, it becomes at least plausible to think that the virtue is merely a means to the good (though Sidgwick does not explicitly give an argument for the latter claim until later in the chapter).

2. In the next section (§ 3.), Sidgwick gives what Tom Hurka reads as his argument against the position that virtue is one good.²⁶

Shall we then say that Ultimate Good is Good or Desirable conscious or sentient Life — of which Virtuous action is one element...? [...] [T]he fact that particular virtues...are largely valued as means to ulterior good does not necessarily prevent us from regarding their exercise as also an element of Ultimate Good: just as the fact that physical action, nutrition, and repose...are means to the maintenance of our animal life, does not prevent us from regarding them as indispensable elements of such life (396).

On Hurka’s reading, Sidgwick goes on to raise an objection to this suggestion, showing a problem for thinking of physical motions in this way and then claiming that the same problem arises for virtue. It “seems difficult to conceive any

²⁶ Hurka, *Virtue* p. 9.

kind of activity or process as both means and end, from precisely the same point of view and in respect of precisely the same quality: and in both the cases above mentioned it is, I think, easy to distinguish the aspect in which the activities or processes in question are to be regarded as means from that in which they are to be regarded as in themselves good or desirable” (396). Physical processes are means to living, but qua physical processes have no value in themselves. What is valuable is “human Life regarded on its psychical side, or, briefly, Consciousness” (396). “In the same way, so far as we judge virtuous activity to be a part of Ultimate Good, it is, I conceive, because the consciousness attending it is judged to be in itself desirable for the virtuous agent” (397). Call this the means/end argument.

This is a puzzling argument if it is read, with Hurka, as an argument against thinking that virtue is one good.

(i) After giving the means/end argument, Sidgwick is aware that further argument is needed against the view that virtue is one good: “the Consciousness of Virtue” might still be “a part” of “Ultimate Good...conceived as Desirable Consciousness,” and a part not to be identified with pleasure (398); we might take “ideal goods” such as virtue, which are not desirable merely *qua* feeling, to be a part of Ultimate Good (400); virtue is not rejected as one good until 400-1 and 402. So Sidgwick does not seem to take the means/end argument to show that virtue is not one good. Perhaps the argument is again directed against the suggestion that character, as a disposition, is intrinsically valuable. The disposition could be seen as the means to the consciousness involved in being virtuous, just as physical processes are the means to consciousness. But Sidgwick does not write of character or disposition here — he writes of “virtuous activity” or the “exercise” of virtue — and, in any case, he has already rejected character or disposition, on 393-4 (397, 396).

(ii) The problem for physical motion does not seem to be that it is valuable as a means from one point of view and valuable as an end from another point of view. Physical motion does not seem valuable as an end at all, unless Sidgwick is thinking that consciousness is constituted by (rather than caused, perhaps only in part, by) physical motions. Virtue, on the other hand, does seem to many to be valuable as a means (to pleasure, for example) and as an end, and there seems nothing incoherent about this. It is true that different properties of virtue have these values — virtue qua producer of (say) pleasure is valuable as a means; virtue qua, for example, the occurrence of certain desires, is valuable as an end. But it is unclear why this is a problem. Perhaps Sidgwick is assuming that vir-

tue is by definition a producer of some further goods, where this is intended to rule out the possibility that virtue is anything else, such as the occurrence of certain desires or feelings.²⁷ But (a) he has not established this definition; (b) the definition conflicts with his initial descriptions of virtue in III.II, noted above, and his treatment of many particular virtues in Book III;²⁸ (c) he does not need this definition to give the circle argument, since that argument turns not on a definition, but rather on the specific results of the examination of common sense morality earlier in Book III, namely that virtues such as benevolence and justice make reference to further goods to be maximised or distributed fairly (392-3); and (d) the definition does not rule out the possibility that virtues have other qualities (not true of them by definition), and these qualities could be valuable.

I think this section of III.XIV is better read, not as arguing against virtue as one good, but as arguing in favour of “desirable conscious life” as what is good. On this reading, Sidgwick first eliminates physical motions. The point of the means/end discussion is merely that once the distinction between means and end is made, there is no plausibility in thinking physical motions are valuable. In the next paragraph (396-7), Sidgwick eliminates conscious life that is not desirable. (There is no mention of means and end here.) When Sidgwick then turns to virtue, his claim is that “in the same way” we see there is no value to “virtuous activity” apart from “the consciousness attending it.” In the earlier argument against character as a disposition, Sidgwick is careful to say that what might have value is actions *or* feelings.²⁹ Here he is eliminating actions. I do not read “in the same way” as referring back to the means/end argument, but to the more general point made against physical motion and undesirable conscious life, namely that once we view them on their own, neither is valuable. On my read-

²⁷ This is Hurka’s reading (*Virtue* p. 10). He then argues that the means/end argument fails because this definition is inadequate (11).

²⁸ For the latter, see, for example, pp. 239, 243-5, 249, 250, 253-4, 258-60, 262, 322-4, 326, 346.

²⁹ A disposition “can only be defined as a tendency to act or feel in a certain way...and such a tendency appears to me clearly not valuable in itself but for the acts and feelings in which it takes effect [...]. When, therefore, I say that effects on character are important, it is a summary way of saying that...the present act or feeling is a cause tending to modify importantly our acts and feelings in the indefinite future: the comparatively permanent result supposed to be produced in the mind or soul, being a tendency that will show itself in an indefinite number of particular acts and feelings, may easily be more important in relation to the ultimate end, than a single act or the transient feeling of a single moment [...]”.(393-4).

ing, then, Sidgwick is not arguing from a definition of virtue as instrumental to producing other goods and the view that what is by definition instrumental to producing other goods cannot plausibly be itself intrinsically good. He is instead arguing that virtue, insofar as it is valuable as an end, is so because of the feelings or consciousness associated with it.³⁰

One piece of support for my reading comes from the evolution of III.XIV. Sidgwick introduced the argument that physical motions have no value in the third edition. There the argument proceeds as I suggest, as an argument by elimination for the conclusion that desirable conscious life is what is valuable: non-conscious life has no value; undesirable conscious life has no positive value. There is no mention of means and ends (3.395). When the means/end discussion is added, in the fifth edition, it is inserted into the paragraph arguing that physical motions have no value. It would be odd if this insertion were the key argument. It seems better read as merely correcting one who might think that since physical motions are “indispensable elements” of our life, they are intrinsically valuable.

A puzzle about this section remains, however. After concluding that “the consciousness attending” virtuous activity is good, Sidgwick notes that virtue also has value as a means. He then writes that “[w]e may make the distinction [presumably between virtue as means and virtue as end] clearer by considering whether Virtuous life would remain on the whole good for the virtuous agent, if we suppose it combined with extreme pain.” Sidgwick thinks not. One “would hardly venture to assert that the portion of life spent by a martyr in tortures was in itself desirable” (397).

It is tempting to read Sidgwick as arguing that whatever consciousness comes with virtue is less valuable than pleasure, since we would trade that conscious-

³⁰ J. B. Schneewind seems to have a similar reading of the means/end argument, but he takes it to be directed against “the position that conscious virtuous action might be part of the ultimate good if we take the ultimate good to be desirable conscious life, and add that desirable conscious life has many components, of which virtuous action is one” (Schneewind, *Sidgwick’s Ethics and Victorian Moral Philosophy* (Oxford: Clarendon, 1977) p. 315). This describes the position Sidgwick does not arrive at until 400, after the means/end argument, and which he attacks with arguments on 400-1 and 402, considered below. Arguing merely that it is conscious life that is valuable, as opposed to something entirely outside of consciousness, does not rule out taking virtue, insofar as it involves consciousness, to be valuable. But since the reflection argument on 400-1 is very similar to the means/end argument, Schneewind’s reading is understandable.

ness to avoid pain. But Sidgwick could not expect agreement on this point; and showing it would not show that virtue is not a good (though a lesser one). It is, alternatively, tempting to read Sidgwick as arguing that it is the pleasurable consciousness attached to virtue that is valuable.³¹ But since he goes on to treat the consciousness attending virtue as a live option for being good, and different from pleasure, on 398 and 400, this cannot be right either.

It is better to read Sidgwick as again concerned with “virtuous activity” in particular (397). His point is then that if virtuous activity brought no good consciousness with it (whether that consciousness be pleasure or something else), we would not value it. On this reading, when Sidgwick mentions the pain of the martyr, pain is standing in for “no good consciousness” (perhaps because extreme pain leaves one conscious of nothing else). This reinforces the point made throughout the section, that only consciousness has value, without making idle the arguments to come.

If this reading is correct, Rashdall makes a good point against Sidgwick, at least about the argument up to this point. Rashdall objects that virtue consists of desires, volitions, judgments, attentions, and emotions, and that these

are actual elements of consciousness [...] When we pronounce character to have value, we are just as emphatically as the Hedonist pronouncing that it is in the actual consciousness that value resides, and in nothing else. It is the actual consciousness of a man who loves and wills the truly or essentially good and not mere capacities or potentialities of pleasure-production such as might be supposed to reside in a bottle of old port, which constitutes the “goodness” or “virtue” which is regarded as a “good” [...] by the school which Professor Sidgwick is criticizing [...] But for the difficulty which Sidgwick seems to make of the matter, it would have seemed unnecessary to point out that those who make “virtue” an end mean by virtue “virtuous consciousness.”³²

Through this point in the chapter, although his arguments may be successful, Sidgwick has not engaged with his real opponent.

³¹ For this reading, see Schneewind, p. 316.

³² Rashdall, *Theory* i pp. 64, 65; also, directed at the third edition, “Utilitarianism” 224. Hayward makes the same point: “Sidgwick’s argument is sound but unnecessary” (*Philosophy* pp. 221; 199, 222, 230-1).

3. In the next section (§ 4.), Sidgwick clarifies the position of this opponent. To value knowledge is to value something that is within consciousness (a belief) and something which goes beyond consciousness (that the belief is true, justified, etc.). To value virtue is to value something that is within consciousness (volitions, etc.) and something which goes beyond consciousness (that the volitions, etc. are morally good). The arguments above do not conclude that what is valuable must be confined to consciousness. They conclude only that consciousness must be part of what is valuable. Hence they do not rule out knowledge or virtue as valuable.

In the following section (§ 5.), Sidgwick argues that virtue, so understood, lacks value. The first argument is that he finds it “clear after reflection that these objective relations of the conscious subject, when distinguished from the consciousness accompanying and resulting from them, are not ultimately and intrinsically desirable; any more than material or other objects are, when considered apart from any relation to conscious existence” (400-1). Call this the reflection argument.

The reflection argument asks why adding something (a material object) that has no effect on consciousness adds no value, whereas adding something else (the truth or justification of a belief or the goodness of a volition) that has no effect on consciousness adds value. Defenders of knowledge or virtue can reply that the whole formed by a belief and its truth or justification has more value than the belief has on its own, or that the whole formed by a volition and its goodness has more value than the volition has on its own.³³ The difference between these additions and adding a material object is just that no whole with greater value is brought about by adding a material object.

Sidgwick might avoid this reply by taking a certain lesson from the discussion of material objects. The reason for rejecting material objects is that consciousness is unchanged; that is a reason for rejecting knowledge and virtue as well.³⁴ The worry is that lacking effects on consciousness may be conclusive against material objects, since we cannot see how else they could affect value, but not

³³ I put the point in terms of the “holistic” rather than “variability” view, but some might want to put it in terms of the latter — rather than thinking a whole with more value is formed, one might say that the belief or volition itself acquires more value. For the distinction, see, for example, Thomas Hurka, “Moore in the Middle,” *Ethics* 113, 2003, 606-7.

³⁴ I owe this suggestion to Joyce Jenkins.

against other things, such as having correct and justified belief or virtue, that could affect value in other ways.

The second argument is an appeal to “a comprehensive comparison of the ordinary judgments of mankind” (400). There are cases “in which the concentration of effort on the cultivation of virtue has seemed to have effects adverse to general happiness, through being intensified to the point of moral fanaticism.” In such cases, “we shall...generally admit that...conduciveness to general happiness should be the criterion for deciding how far the cultivation of Virtue should be carried” (402).³⁵ Call this the criterion argument.

One problem with the criterion argument is that it fails to show that virtue is not a good. Even if we prefer happiness to virtue, we might think, with Ross, that of two worlds equal in happiness but unequal in virtue, the world with more virtue is better.³⁶

Another problem is that some may disagree. Consider not a moral fanatic, but rather one with so much sympathy that she sometimes helps others when she is not qualified to do so, and so makes matters worse. Some may think a world with such people, and less happiness, is better than a world with less sympathy and more happiness.³⁷ (Others, again, condemn those who “mean well.”)

³⁵ Similarly, Sidgwick writes that “when Virtue and Happiness are hypothetically presented as alternatives, from a universal point of view, I have no doubt that I morally prefer the latter; I should not think it right to aim at making my fellow-creatures more moral, if I distinctly foresaw that as a consequence of this they would become less happy. I should even make a similar choice as regards my own future virtue, supposing it presented as an alternative to results more conducive to the General Happiness” (FC 487). Rashdall replies by inviting “the reader to say whether he can accept [this] as a correct representation of his own moral consciousness — or of Henry Sidgwick’s” (*Theory* i, p. 70).

³⁶ Ross, *Right* p. 134.

³⁷ Rashdall gives the following (now unconvincing) example: “On what other grounds can we either explain or justify [Common Sense’s] emphatic condemnation of suicide in cases where it is clearly conducive to the happiness of the individual and of all connected with him?” (“Utilitarianism” 219). Elsewhere he gives examples of bullfighting, Roman wild-beast and gladiatorial fights, German students’ face-slashing duels, coursing, pigeon-shooting, and drunkenness (“and we should think a man’s conduct in getting drunk worse instead of better if he had carefully taken precautions which would prevent the possibility of his doing mischief...while under the influence of his premeditated debauch”) as worse than their absences even if they maximise pleasure (*Theory* i pp. 97-9). (Rashdall also argues, more ambitiously, that Sidgwick is inconsistent in taking pleasure but not virtue as good. I examine this argument in “*Utilitarianism*,” forthcoming in the *Oxford Handbook of the History of Ethics*, ed. Roger Crisp).

I should note an alternative interpretation. J. B. Schneewind suggests that Sidgwick's point on 402, and earlier when rejecting the good will, is that "there is a limit to the extent to which we think it supremely good to act according to one's moral convictions: and the limit is determined by the utilitarian principle. Here a dependence argument shows that the good will [or virtue more generally] cannot be an ultimate good, for its limits are determined by the claims of another good and its own directives may be overridden in the name of that other good."³⁸

The problem is that Sidgwick seems to understand by an "ultimate" or "intrinsic" good a good which is good as an end rather than just as a means. He does not seem to mean a good that is not limited by other goods. For example, when he considers whether virtue could be both a "means to ulterior good" and "an element of Ultimate Good," he comments that "it seems difficult to conceive of any kind of activity or process as both means and end...and...it is...easy to distinguish the aspect in which the activities...are to be regarded as means from that in which they are to be regarded as in themselves good or desirable" (396). Sidgwick takes his opponents to hold that virtue (and other things) "are ends independently of the pleasure derived from them" (401). He explains that he means by "'Ultimate Good' [...] that which is Good or Desirable *per se*, and not as a means to some further end" (407n). If so, showing that one good is limited by another does not show that the limited good is not ultimate. It might be merely a lesser ultimate good. Thinking that we always trade virtue for happiness, even if true, does not show that the value of virtue is dependent on its production of happiness. Schneewind's interpretation has the advantage of making the criterion argument valid. But it has the disadvantage of not fitting Sidgwick's claims about means and ends and, more importantly, makes Sidgwick talk past those, like Ross, who hold the usual understanding of "intrinsic" or "ultimate."

Sidgwick might have tried a different strategy. In the case of knowledge, the appeal to the ordinary judgments of mankind proceeds by noting that knowledge is valued in proportion to the happiness it brings. The connection to happiness can explain why we value even apparently fruitless knowledge, both because we know that such knowledge can "become unexpectedly fruitful" and

³⁸ Schneewind, p. 314. This interpretation also fits the passage concerning Liberality, etc., quoted earlier.

because its pursuit is itself pleasurable and shows a disposition likely to produce fruitful knowledge (401). Call this the proportion argument.

It does not follow that knowledge is good only as a means. First, each piece of knowledge might have the same value, with the proportion claim made true by combining these values with the differing amounts of happiness produced.³⁹ Second, pieces of knowledge might vary in value, but not so much that less valuable knowledge that produces greater happiness is ranked lower than more valuable knowledge that produces less happiness. Since the proportionality claim is hardly precise, it might be hard to discount this possibility. Third, Ross's example of two worlds equal in happiness and unequal in knowledge convinces some that the proportionality claim is false.⁴⁰

Sidgwick might have tried a parallel argument for virtue. He does not explicitly say that virtues are valued in proportion to the happiness they bring. But he makes the similar claim that utilitarianism explains our ranking of duties (425-6); it is plausible to think that the "minor" virtues — Sidgwick lists caution, decision, good humour, meekness, mildness, gentleness, placability, mercy, liberality, politeness, and courtesy (236, 253, 321, 324-5) — are minor because they are usually less productive of happiness than virtues such as benevolence and justice; he argues that virtues such as purity, courage and humility, which seem to be admired independently of happiness, really do, insofar as they are admired, contribute to happiness (or other virtues) (332, 334, 355, 356n, 429, 450-3, 456); and when Sidgwick later writes that in III.XIV he "tried to show that Common Sense is unconsciously utilitarian in its practical determination of those very elements in the notion of Ultimate Good or Wellbeing which at first sight least admit of a hedonistic interpretation," he suggests that the utilitarian ranking of pieces of knowledge is to be paralleled by a utilitarian ranking of virtues (453-4).

Just as with the proportionality argument against knowledge, this is hardly conclusive. But two of the reasons for resisting the proportionality argument in the knowledge case seem less telling in the case of virtue. Few proponents of virtue would claim that each instance of virtue has the same value. And perhaps, given the importance ascribed to virtue, it would be difficult to hold that differences in the values of virtues are sufficiently small as to exclude the possibility

³⁹ Skelton noted that Ross himself rejects this; see *Right* p. 139.

⁴⁰ Ross, *Right* p. 139.

that a more important virtue that produces less happiness has more value than a less important virtue that produces more happiness.

Ross's appeal to worlds of equal happiness and unequal virtue remains. But here Sidgwick might note that Ross's verdict is controversial — in my experience, at best half agree with Ross — and, as with knowledge, Ross's intuition might be explained away, given how difficult it is to imagine something normally so useful as making no difference. Slight variations in the presentation of the case also seem to hurt Ross. For example, say I could increase virtue by writing an inspirational book in moral philosophy, but this would make no difference to the amount or distribution of happiness in the world. (Say the book increases the number of actions done out of duty, but that in all these cases self-interest would have led to the same action.) Many think it does not matter whether I write the book or not.

This, at any rate, seems the sort of argument Sidgwick should have stressed. If he had, the supporters of virtue who came later would at least have had to work harder. It is regrettable that so much of III.XIV is spent on other matters.

Appendix

III.XIV in the final edition incorporates, not always smoothly, changes made over the first four editions. It may be helpful to briefly chart these changes.

In the first edition, Sidgwick notes that “the majority of moral persons would probably declare that Virtue is the *chief* good [but] very few would maintain that the *only* thing in life intrinsically desirable is the habit of obeying moral rules” (1.369). Against the view that virtue is the only good, he gives a quick version of the circle argument (1.369, 376). There is also a version of the argument against dispositions as being of value, but Sidgwick does not take this to count against virtue, but rather to specify that virtue is a matter of “conscious action and feelings” (1.369). Virtue is rejected as a good, along with other objective relations such as knowledge, on the basis of the reflection argument (1.371-2). The criterion argument against virtue and the proportion argument against knowledge do not appear.

In the second edition, the circle argument is expanded to roughly its final form (2.364-5). Virtue is dismissed on the basis of it: it follows from the circle argument that “we cannot, without manifest divergence from Common Sense, introduce [virtue] in a scientific explanation of the nature of Ultimate Good”

(2.365). This is a blunder not made in the first edition, and not wholly corrected until the fifth: the circle argument excludes virtue only as the sole good, not as one good. The prominence of the circle argument, and the focus on virtue as the sole good, may be due to Bradley's *Ethical Studies* and *Professor Sidgwick's Hedonism*, both of which appeared between the first and second editions. Bradley sometimes claims that the sole good is "function," and sometimes treats "function" and "virtue" as interchangeable.⁴¹

In the second edition, knowledge is presented as an alternative to virtue or happiness and is rejected by the reflection argument and the proportion argument (2.366-9).⁴²

In the third edition, Sidgwick expands the argument against dispositions to roughly its final state (3.393). One difference is that he takes it to be a further argument ruling out virtue, and not just dispositions, as one good (3.394). Unlike the circle argument, this argument could show that virtue is not one good, provided one thought of virtue just as a disposition — but as Rashdall notes, defenders of virtue need not think this. This may explain why, by the fourth edition, the argument is taken to discredit only dispositions, and not virtue in general.⁴³

The third edition also adds the arguments against physical processes and mere survival. They are introduced to limit what is valuable to conscious life (3.395).

The fourth edition adds the argument against the will to do what one takes to be right (4.394). Sidgwick recognises that, even if this argument succeeds, the good will could still be one good: it would be a paradox to "affirm [subjective rightness of will] to be the sole Ultimate Good" but not "paradoxical to regard the settled will to realise our duty as an essential part of ultimate good: while at the same time recognising that there are effects of right volition [...] which are also in themselves good" (4.394).⁴⁴ He then objects that if I "suppose that the

⁴¹ See, for example, F H. Bradley, *Professor Sidgwick's Hedonism*, in Bradley, *Collected Essays* (Oxford: Oxford University Press, 1925) pp. 95, 96n, 97, 98, *Ethical Studies* (Oxford: Clarendon, 1927) pp. 136, 137, 138; *Hedonism* pp. 97, 98, *Studies* pp. 140-1.

⁴² The second edition is very similar to "Hedonism and Ultimate Good," *Mind* o.s. 2, 1877, published in the same year.

⁴³ The upshot in the third edition is that "virtues or talents, faculties, habits or dispositions of any kind" are not goods (3.393). In the fourth and later editions, the upshot is that "faculties, habits, or dispositions of any kind" are not goods (4.393, 393).

⁴⁴ This is noted by Schneewind, pp. 313-14.

effects of a man's acting in accordance with his conception of what is right will be on the whole bad — according to an estimate of badness framed without taking into account the subjective rightness of the volition — , I find that this consideration of them appears to me finally decisive of their badness. In my view, therefore, this Subjective rightness of volition is not Good in itself, but only as a means" (4.395). There is no further discussion of virtue as one good. In effect, Sidgwick runs the criterion argument not against virtue in general, but against one account of virtue, as the good will.

The fifth edition (which for III.XIV is the same as the later editions) takes seriously the concession made regarding the good will: not only the good will, but also virtue more generally, has not been excluded as one good by the circle argument. Sidgwick then restores virtue to the place it had in the first edition, as an objective relation like knowledge. (Virtue has this place in the *Lectures* as well (GSM 126).) It is rejected by the reflection argument, as in the first edition, and by the new criterion argument, which generalises the point made against the good will in the fourth edition.⁴⁵

⁴⁵ Thanks to Darcie Fehler, Adam Muller, Emily Muller, Jeff Verman, Sandy Vettese and Andrew Webb for discussion of some of the examples; to Tom Hurka for discussion of many of the moves in the paper; and to Joyce Jenkins and Anthony Skelton for detailed comments on an earlier draft.