

David Phillips, *Sidgwickian Ethics*, (Oxford: Oxford University Press, 2011), 163 pages. ISBN: 9780199778911 (hbk.). Hardback: £40.00.

David Phillips's tightly argued *Sidgwickian Ethics* deserves high praise. Its ambition is to interpret and evaluate the main meta-ethical and normative doctrines of Henry Sidgwick's *The Methods of Ethics* (ME). It deals with Sidgwick's non-naturalism, his intuitionist moral epistemology, his rejection of the (deontological) rules of common-sense morality, and his "Dualism of Practical Reason".

There is an established body of scholarly and philosophical literature on the topics in Sidgwick to which *Sidgwickian Ethics* is devoted. Its novelty does not therefore lie in the issues of which it treats. Instead, its novelty lies in the fact that it helps us see clearly the relations between its topics.

Sidgwick defends a kind of non-naturalism: "the fundamental notion represented by the word 'ought' or 'right,' which such [moral or prudential] judgments contain expressly or by implication...[is] essentially different from all notions representing facts of physical or psychical experience" (ME 25). In chapter two, Phillips contends that this view is defensible (pp. 20-21), that it is central to Sidgwick's ethics (pp. 13, 21-28), and that Sidgwick's defense of non-naturalism is superior to Moore's (pp. 35-38).

Some balk at the suggestion that non-naturalism is central to Sidgwick's ethics. His subscription to this doctrine has certainly done him no favours, especially when understood to have ontological implications, which to Phillips it does (pp. 27-30). The idea, then, that this view is central to Sidgwick's ethics requires explanation.

One explanation, overlooked by Phillips, is that the commitment is taken as fundamental because it is a premise in an argument for another view that is central to

Sidgwick's ethics, that moral requirements are *à la* Kant categorical and necessary. The analyses of "ought" that Sidgwick does consider seem either to challenge the idea that moral claims are categorical or that they are necessary. This understanding of Sidgwick fits with the fact that he ends his discussion of various attempts to analyze "ought" in terms of natural facts by claiming that his aim has been to "exhibit" that there are "unconditional or categorical" imperatives (ME 35), a claim which he defends until the end of the chapter of ME that includes it, and with the fact that he is keen to show in his skirmish with Stephen (much-discussed by Phillips) that denying non-naturalism leaves one unable to account for the "necessary nature of an ethical first principle" (pp. 26, 29).

In ME, Sidgwick provides an argument for utilitarianism; it relies in part on an appeal to two philosophical intuitions:

(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).

(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).

He thinks these get him utilitarianism's "fundamental" or "first" principle (ME 387, 421): roughly, that we ought to promote aggregate good even at a net cost to ourselves. Sidgwick thinks (U) and (R) are self-evident truths. They satisfy the "conditions" of self-evidence: that intuitions be clear and precise, self-evident on

reflection, consistent with other self-evident intuitions one accepts, and not denied by someone who one believes is no more likely to be in error than oneself (ME 338-342).

He goes on to claim that unlike (U) and (R) the intuitions of common sense (e.g., that we ought to keep our promises) do not satisfy the conditions of self-evidence; in particular, they are not clear and precise in practice (ME 342, 360-361). They represent “mere opinions” about what we ought to do (ME 338). Phillips calls this the “criterial argument” for utilitarianism (pp. 64, 67).

Sidgwick also provides a “proof” of utilitarianism, aimed at the proponent of common-sense morality. In it, he argues that utilitarianism’s ability to support the main rules of common-sense morality and to inject greater clarity, completeness, and system into ethical thinking provides the proponent of common-sense morality with reason to accept it. Sidgwick describes this proof as involving 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). Phillips calls this the “bi-partite ad hominem” argument (p. 63).

Many think that these arguments reveal Sidgwick’s moral epistemology. Phillips agrees. In chapter three, he argues that in the criterial argument Sidgwick relies on a moderate foundationalism, according to which intuitions possess intrinsic epistemic credibility that is amplified with the satisfaction of the third and fourth conditions of self-evidence (pp. 59, 60-62, 76-85). (He thinks, further, that Sidgwick supplies an argument to the effect that intuitions are indispensable to ethics (pp. 53-59)). It is common for commentators to argue that in the bi-partite argument Sidgwick attributes probative status

to common-sense morality and that he thinks that utilitarianism is justified in part by the fact that it coheres with it. Phillips denies this (pp. 74-76).

There is, however, a puzzle at the heart of Sidgwick's epistemology (pp. 65-76). Sidgwick's epistemology is puzzling because he holds that both the criterial and the bi-partite arguments are designed to deal with disagreement about first principles. He offers the first argument in ME III.xiii and the second in ME IV.ii. In ME IV.ii Sidgwick does not mention his first and (for him) stronger argument. The puzzle is this: why does Sidgwick not appeal to the criterial argument in ME IV.ii where it is, Phillips says, by Sidgwick's lights stronger and "clearly relevant" (p. 71)?

There is a reply. The criterial argument is designed to demonstrate that there is a consensus of ethical experts on utilitarianism's first principle. There is no need to address the proponent of common-sense morality here; she accepts utilitarianism's first principle. She does not yet accept utilitarianism. This is where the second argument comes in. It aims to secure the agreement of the proponent of common-sense morality, not on utilitarianism's first principle, but on the utilitarian method of ethics itself. The last is more robust, comprising, *inter alia*, a commitment to hedonism and to maximization, which, Phillips seems to admit, the criterial argument cannot secure (pp. 97, 118).

Phillips's own account of Sidgwick's epistemology is puzzling. Sidgwick, he says, employs "a conception of criterial argument which gives no evidential role to common-sense morality" (p. 76). This fits with his claim that in the bi-partite argument Sidgwick does not grant common-sense morality probative status. However, Phillips also believes that Sidgwick holds that in satisfying the fourth condition of self-evidence, (U) and (R), e.g., gain epistemic support from "ordinary moral opinions" (p. 79). These

statements look to be in tension with each other, since Phillips suggests that ordinary moral opinion includes common-sense moral opinion (p. 79). This suggests that the criterial argument does grant common-sense morality an evidential role. If this is true, then it is puzzling that, as Phillips has it, Sidgwick would deny common sense an evidential role in the bi-partite argument.

Phillips's treatment of Sidgwick's meta-ethics and his moral epistemology is largely friendly. His treatment of Sidgwick's normative conclusions is not.

A key component of Sidgwick's case for utilitarianism is the argument against common-sense morality. Phillips forcefully contends, in chapter four, that Sidgwick's (official) argument is less successful than he thought. Indeed, it is unfair (pp. 101, 110). Sidgwick asks that the rules of common-sense morality be "made precise enough to give determinate verdicts in every case" (p. 101). But he makes no such demands of his own principles, which are too abstract "to enable us to ascertain by immediate application of them what we ought to do in any particular case" (ME 379). (Phillips considers, but rejects, possible replies on Sidgwick's behalf (pp. 103-110))

Sidgwick concludes ME thinking that utilitarianism and rational egoism are equally compelling but contradictory requirements of rationality (ME 508). This so-called "dualism of practical reason" relies on there being good arguments for both parts of the dualism. This is the topic of chapter five. Phillips argues that the intuitive argument for the first principle of utilitarianism is unsuccessful (pp. 120-126), and that the same is true of the intuitive argument for egoism (pp. 126-134). However, he thinks that the argument for egoism is more successful than the argument for utilitarianism. It contains at least one non-tautologous principle and it establishes that there are genuine agent-

relative reasons (pp. 126, 129, 131, 133). He does not think that Sidgwick's arguments establish that egoism and utilitarianism conflict; instead, they "provide the model for a range of historically distinctive and plausible positive views of practical reason" (p. 151).

Phillips is cagey about what the first principle of utilitarianism amounts to. He gives a range of options, including that "one must do what is ultimately conducive to universal good" (p. 118). This is a reasonable account of what Sidgwick gets from (U) and (R). Yet, when it comes to the discussion of the failure of the intuitive argument for the first principle of utilitarianism, he suggests that the conclusion that Sidgwick aims for but does not get is much stronger: "doing what promotes universal goodness is, all things considered, either rationally required or rationally permissible" (p. 126).

This is not the conclusion that Sidgwick wants. He is clear in his discussion of the bi-partite argument for utilitarianism that the proponent of common-sense morality can accept the first principle of utilitarianism (ME 421). It would not be possible to say this if his aim was to arrive at the stronger conclusion from (U) and (R).

If Sidgwick did intend the weaker conclusion, it is harder to see that the argument is a complete failure. It might get us the claim that there are genuine agent-neutral reasons. Not even Phillips is clear that it fails in this regard (pp. 140, 150; cf. p. 144).

Anthony Skelton
University of Western Ontario
askelto4@uwo.ca