Phillips, David. *Sidgwickian Ethics*

New York: Oxford University Press, 2012. Pp. xii+163. $65 (cloth).

There is, alas, no truly scholarly, variorum edition of Henry Sidgwick’s masterpiece *The Methods of Ethics* (London: Macmillan, 1907, first edition 1874)*,* the work that Derek Parfit, echoing an early statement by C. D. Broad, has called “the best book on ethics ever written” (*On What Matters*, vol. 1 [Oxford: Oxford University Press, 2011], xxxiii).

 But the absence of a scholarly edition of the *Methods* has not damned the stream of scholarly work on the *Methods*, work of both a historical and an analytical nature. Parfit’s brilliant works, *Reasons and Persons* (Oxford: Oxford University Press, 1984) and *On What Matters*, are both admittedly in significant measure extended analytical engagements with Sidgwick’s ethical and metaethical philosophy. And David Phillips’s new book, *Sidgwickian Ethics*, is an excellent contribution in this Parfitian vein. The aim is straightforward: to provide a clear and crisp assessment of the *Methods*, celebrating what is living in it and burying what is dead. As Phillips announces, in his opening line: “My aim in this book is to interpret and evaluate the central argument of Sidgwick’s *The Methods of Ethics*, in a way that brings out the important conceptual and historical connections between his views and contemporary moral philosophy” (3).

 Phillips succeeds in this task on a great many counts, though it must be allowed that a good deal of what he declares dead in Sidgwick would have been regarded by Sidgwick himself as healthy limb, not up for amputation. Phillips acknowledges this:

My view of the success of Sidgwick’s treatments of the two key conflicts contrasts with his. He thinks he succeeds in reconciling utilitarianism with intuitionism, but is left with some kind of unresolvable conflict between utilitarianism and egoism. By contrast, I shall argue that his treatment of the conflict between utilitarianism and intuitionism is less successful than he supposes, and his treatment of the conflict between utilitarianism and egoism is more successful. His treatment of the conflict between intuitionism and utilitarianism is less successful than he supposes because, while he succeeds in undermining the idea that there is a basic *epistemic* contrast between utilitarianism and intuitionism, he does not succeed in his attempt to support utilitarianism as against the *moral-theoretic* component of dogmatic intuitionism. His treatment of the conflict between utilitarianism and egoism is more successful than he supposes because, while he (on the whole) takes himself to be left with a fundamental contradiction, his arguments instead support a coherent and plausible positive view of practical reason.” (p. 6).

To unpack: Sidgwick did hold that his extended grappling with three “methods of ethics”—egoism, utilitarianism, and a cleaned up version of commonsense morality that he labeled “dogmatic intuitionism” and attributed to William Whewell, all of which represented going rational procedures for individuals to determine what it is right to do—had yielded some philosophical progress, though also a great deal of philosophical (and religious) anxiety. As he put it, in his autobiographical sketch:

I tried to say what I had found: that the opposition between Utilitarianism and Intuitionism was due to a misunderstanding. There was indeed a fundamental opposition between the individual’s interest and either morality, which I could not solve by any method I had yet found trustworthy, without the assumption of the moral government of the world: so far I agreed with both Butler and Kant./ But I could find no real opposition between Intuitionism and Utilitarianism…. The Utilitarianism of Mill and Bentham seemed to me to want a basis: that basis could only be supplied by a fundamental intuition; on the other hand the best examination I could make of the Morality of Common Sense showed me no clear and self-evident principles except such as were perfectly consistent with Utilitarianism. (Sidgwick, *Methods,* xxii-xxiii).

Thus, for Sidgwick, as one sorted out the various methods more or less unconsciously at work in ordinary people’s jumbled practical reasoning, one found both a deep coherence, in the cognitivist utilitarian grounding of commonsense morality, and a deep incoherence, in the irreconciliability of utilitarianism and rational egoism, the “dualism of the practical reason.” Phillips in effect gleans from Sidgwick’s work neither the deep coherence nor the deep incoherence of practical reason.

Given the evident emphasis in Sidgwick’s reflections on his failures, that is, his belief, as the last line of the first edition of the *Methods* put it, that he had come close to reducing practical reason to “a chaos” and a “failure” because of the dualism of the practical reason, the tribute that Phillips pays to the *Methods* might seem a little backhanded, suggesting that Sidgwick himself radically misread his own achievements. And it is possible to remain persuaded that, in the end, Sidgwick had the richer philosophical and historical vision, and that Phillips’s salvage job succeeds chiefly on a narrower academic front, without much sense of how little or Pyrrhic the argumentative victories might be. From a larger Sidgwickian perspective, if not the “Point of View of the Universe,” Phillips seems to be tidying the deck chairs on the Titanic.

On the way to his big conclusions, Phillips has two superb chapters on “Sidgwick’s Metaethics” and “Sidgwick’s Moral Epistemology” that ought to be required reading for all future Sidgwick scholars. In addition to setting out with exceptional clarity the various elements of Sidgwick’s metaethics—his claims that “it is part of our moral concepts that there is such a thing as moral truth and error,” that there is in fact such a thing as moral truth and error, and that there is a fundamental distinction between judgments of “ought” or “right” and representations of physical or psychical experience—Phillips has a very original and intriguing treatment of the filiations between Sidgwick’s position and J. L. Mackie’s “error theory” of ethical judgment, which accepts the first, realist conceptual, claim, and the fundamental distinction claim, but denies the second claim, the substantive realist claim. He argues, intriguingly and suggestively, that “Sidgwick is tempted by error theory and in a way ought to have been an error theorist,” given his unearthing of what he took to be fundamental contradictions in practical reason (39). That is, the dualism of the practical reason, or conflict between utilitarianism and egoism, if taken in Sidgwick’s own terms, ought to have pushed him into the error theory camp, or at least a campground in that vicinity, where the realist concepts conflict with the stubborn realities.

On the conflict between utilitarianism and commonsense or intuitional morality, Phillips argues, rather less originally but nonetheless incisively, that Sidgwick harbored a double-standard, in that the more fundamental and abstract axioms that he claims to discover beneath commonsense morality (e.g., that similar cases ought to be treated similarly), including the ones that would ground utilitarianism, do not meet the very tests of clarity and precision that he applies in his critique of commonsense morality. That is, “Sidgwick requires that the principles of common-sense morality meet a standard of clarity and precision from which he exempts his own favored principles” (110). Sidgwick’s official argument fails against Whewellian intuitionism, and the “unofficial” arguments hinted at in parts of his work would lead straight to W. D. Ross’s version of intuitionism. At any rate, the epistemic point lacks the moral theoretic thrust.

But most of Phillips’s efforts to save what he sees as the living tissue in Sidgwick are directed less at the reconciliation of utilitarianism and commonsense morality and more at the supposed conflict between utilitarianism and egoism. Sidgwick often seems to cast this as a head on conflict: one simply cannot have, at one and the same time, most reason to act for one’s own good and most reason to act for the general good. This, for Phillips, is what the error theorist could turn to advantage. But Sidgwick is rescued from the prospect of sliding into the error theory by Phillips’s radical reconstruction of his claims about the dualism of practical reason, which render it less destructive of the substantive realist thesis, though with the result that Phillips’s claims for an intuitionist metaethics are far more tentative and qualified than Sidgwick’s or even Parfit’s. In exegetical terms, Phillips claims that there are other possibilities, and “that the most textually plausible conflict-mitigating interpretation of the dualism is the permissive interpretation according to which (1) it is rationally permissible to do what best promotes universal good, and (2) it is rationally permissible to do what best promotes one’s own good” (151). And Phillips would like to nudge this interpretation of Sidgwick toward a fuller, albeit heavily qualified position, “the qualified permissive view,” according to which:

whenever one of two options is better for the agent and another is impartially better, and the impartial difference and the difference for the agent are of the same level of seriousness, it is rationally permissible to choose either option. But when one option is better in one of these ways (i.e., for the agent or impartially) and not worse in the other, one is rationally required to choose the option that is better in one way. And when one option is trivially better in one of these ways (i.e., for the agent or impartially) and the other option seriously better in the other, one is rationally required to choose the option which is in one way seriously better (147).

This is no doubt a promising approach, one somewhat reminiscent of Samuel Scheffler’s arguments, but Phillips does not attempt to defend it fully; he only presents enough of it to signal some of his differences with Parfit, Roger Crisp, Thomas Nagel, and other deeply Sidgwickian contemporaries. Unlike Parfit, he wants to make more of a distinction between intuitions about morality and intuitions about rationality. And in some rather pregnant lines that tie together many themes of the book, he explains:

Nagel explicitly says that it appears to him self-evident that pain is bad … I think it is not self-evident that pain is bad, because it is not self-evident that anything is (really) bad. That is, I think error-theory is not self-evidently mistaken. So to defend the existence of agent-neutral hedonic values, we need to fall back on a weaker claim: if anything is in any way bad, then pain is agent-neutrally bad. The idea, that is, that if there are any genuine values or reasons, then pain is genuinely bad and there are genuinely reasons to avoid pain. (145).

Such a stance is profoundly deflationary of the ambitions of Parfit’s *On What Matters,* which holds, by contrast, that reasons are more basic than morality, that some beliefs are intrinsically credible (self-evident), and that Mackie’s error theory simply failed to account for the moral dimensions of our thinking and veered towards nihilism. On balance, Parfit may, I believe, be closer to Sidgwick, who would scarcely have regarded the error theory with the same equanimity that Phillips brings to it. Even Sidgwick’s student, Bertrand Russell, who came around to the error theory well before Mackie, found it profoundly unsatisfactory.

Consider in this connection the gloss that, in the second edition of the *Methods,* Sidgwick put on the dualism and its possible resolution by the appropriate Theistic premise:

For, if we find an ultimate and fundamental contradiction in our apparent intuitions of what is Reasonable in conduct, we seem forced to the conclusion that they were not really intuitions after all, and that the apparently intuitive operation of the Practical Reason is essentially illusory. Therefore it is, one may say, a matter of life and death to the Practical Reason that this premise should be somehow obtained…. it seems plain that in proportion as man has lived in the exercise of the Practical Reason—as he believed—and feels as an actual force the desire to do what is right and reasonable as such, his demand for this premise will be intense and imperious. Thus we are not surprised to find Socrates—the type for all ages of the man in whom this desire is predominant—declaring with simple conviction that ‘if the Rulers of the Universe do not prefer the just man to the unjust, it is better to die than to live.’ And we must observe that in the feeling that prompts to such a declaration the desire to rationalize one’s own conduct is not the sole, nor perhaps always the most prominent, element. For however difficult it may practically be to do one’s duty when it comes into conflict with one’s happiness, it often does not seem very difficult, when we are considering the question in the abstract, to decide in favour of duty. When a man passionately refuses to believe that the ‘Wages of Virtue’ can ‘be dust,’ it is often less from any private reckoning about his own wages than from a disinterested aversion to a universe so fundamentally irrational that ‘Good for the Individual’ is *not* ultimately identified with ‘Universal Good.’ (*Methods,* 2nd edition, 468-69).

It is, plausibly, this spectre of an unfriendly universe, of a perverse Cosmos and a fragmented self in which reason is schizophrenic or indeterminate, and the Wages of Virtue all too often dust, that most deeply disturbed Sidgwick, even if he sometimes articulated the issue in drier, more limited terms. His Socratic enterprise of testing commonsense morality ended up threatening to collapse into Aeschylean horror. He could see how a truly tragic world might at least enable pure sacrifice, moral heroism, but he did not think most people would really be able to embrace such a worldview beyond a certain point. He doubted that he could. The allusions here to ‘wages’ are clearly to Tennyson’s poem of that title, the last stanza of which runs:

The wages of sin is death: if the wages of Virtue be dust,

 Would she have heart to endure for the life of the worm and the fly?

She desires no isles of the blest, no quiet seats of the just,

 To rest in a golden grove, or to bask in a summer sky;

Give her the wages of going on, and not to die.

(*The Works of Alfred Lord Tennyson,* [London: Wordsworth Poetry Library, 2008], 656)

In A. C. Bradley’s commentary on Tennyson, this poem is invoked in reference to Tennyson’s “In Memoriam” and the following possible explication of Tennyson’s attitude floated: “And would it have been just to make him merely that he might die? … Or perhaps … the idea is rather: To make him such that he thinks himself immortal when he is really not so, would be unjust.” (*A Commentary on Tennyson’s In Memoriam* [London: Macmillan, 1923], 81-82). For Sidgwick, who adored Tennyson, this question of having the heart to endure when the Cosmos has turned out to be so productive of cruel delusions about the nature of the self and its duties, and so horribly hopeless with respect to immortality and the moral structure needed to underwrite the righteous, was surely another aspect of his own “disinterested aversion” to an “irrational” universe.

This side of Sidgwick is not to be found in Phillips’s otherwise excellent book. Where he finds triumph, Sidgwick found tragedy.

Bart Schultz

University of Chicago