

Review of Henry Sidgwick's book *The Methods of Ethics*

by *Jacob Williams* on 2025-02-01
feedback is welcome at jacob@brokensandals.net

There's an amusing story about how much my philosophical hero, Derek Parfit, adored this book:

Parfit and [a student named Larry Temkin] would become close, but in 1977 the teacher inadvertently nearly stymied the career of the student. Temkin was preparing for his oral exam, a 'comprehensive' exam, which would partially be about utilitarianism and which he was required to pass to move to the dissertation stage. 'What should I read?' he wanted to know. 'Sidgwick's *The Methods of Ethics*', Parfit replied. Temkin had to travel to New York to get hold of a copy, and when he'd read this huge tome, he went back to Parfit. 'What should I read next?' 'He replied, "Reread Sidgwick.'" Temkin reread it and approached Parfit again; 'Is there anything I should read on utilitarianism besides Sidgwick?'. Parfit advised him to reread Sidgwick a third time, rather than devote energy to inferior writings. Eventually it came to the oral exam. Temkin was questioned about various standard books and papers in the field, but not having read them was unable to comment. One examiner was exasperated: "*Larry*, this is supposed to be a *comprehensive* exam—what *have* you read?!!!" At which point another of the examiners, Derek Parfit, 'raised his finger and emphatically interjected, "He has read *Sidgwick*!!!"'¹

I've read *Methods of Ethics* twice (admittedly the first time was many years ago) and can attest that I do *not* feel prepared to pass any sort of comprehensive exam, nor would I expect a third reading to rectify that.

The three 'methods' covered by the book are Egoistic Hedonism (trying to maximize your own happiness), Intuitionism (trying to do what you intuitively perceive to be right regardless of consequences), and Utilitarianism (trying to maximize everyone's happiness). Sidgwick is somewhat equivocal about the decision between egoism and utilitarianism (at the end of the book he seems to say that in situations where behaving morally would go against one's own self-interest, reason cannot tell you which to prioritize; it will simply depend on what you feel stronger "impulses" toward.²) But his discussion of intuitionism and "the Morality of Common Sense" is focused on showing that you can't build a precise, coherent system out of common-sense morality, and that the difficulties point us toward utilitarianism as the solution.

Utilitarians are ... called upon to show a natural transition from the Morality of Common Sense to Utilitarianism... so that Utilitarianism may be presented as the scientifically complete and systematically reflective form of that regulation of conduct, which through the whole course of human history has always tended substantially in the same direction.³

Sidgwick argues that utilitarianism does a good job of unifying and justifying much of common-sense morality (even though they don't line up perfectly):

It may be shown, I think, that the Utilitarian estimate of consequences not only supports broadly the current moral rules, but also sustains their generally received limitations and qualifications: that, again, it explains anomalies in the Morality of Common Sense, which from any other point of view must seem unsatisfactory to the reflective intellect...⁴

I'm not super interested in this as a way of arguing for utilitarianism, since I think [arguing for utilitarianism directly](#) is the easiest way to establish a reason for caring about morality in the first place.

Some topics for further thought:

Utilitarian just-so stories. I think we should be a bit uncomfortable with how easy it is to come up with utilitarian justifications for predetermined conclusions. Sidgwick, for example, argues that his culture's disapproval of extramarital sex makes perfect sense on utilitarian grounds⁵. To me his argument seems like a just-so story motivated more by a desire to vindicate preexisting biases than to seriously determine the true implications of utilitarianism. But lots of discussion around utilitarianism today takes a similar form. As an extreme case, consider the thought experiment of whether to harvest organs from one healthy person to save multiple dying people: most of us treat this as a puzzle where the goal is to prove that utilitarianism *doesn't* have the radical implication (saying yes to the harvesting) that you might initially think it does. And I think most of us quickly find an explanation that satisfies us. But how do we know we're not making the same mistake Sidgwick did? Is there anything Sidgwick could have done (other than *be born a century later*) that would have allowed him to tell whether his arguments about sexual ethics were objectively strong or just a reflection of his culture's prejudices?

Epistemic challenges. The discussion of egoism spends a lot of time on the difficulty of knowing what makes you happiest. Sidgwick lists a lot of difficulties with trying to figure it out by empirical observation, but also thinks the alternatives to empiricism about this are pretty weak. (One alternative is to make deductions from a scientific theory about "the causes of pleasure and pain"⁶, but he didn't think a solid theory was available.) I don't think I've taken this problem seriously enough in the past. One might hope modern science puts us in a much better position than Sidgwick, but from an individual's perspective it seems hard to tell what the science really says on this topic, given the prevalence of bad science reporting, studies that don't replicate, etc.

The information content of common-sense morality. Obviously, difficulty in figuring out what makes people happy is a problem for utilitarianism, too. But since Sidgwick thinks common-sense morality is "roughly and generally but not precisely or completely adapted to the production of the greatest possible happiness for sentient beings generally"⁷, he thinks the utilitarian should use it as a starting point to be cautiously revised. I wonder: *how strongly* does the existence of some moral norm count as evidence that the norm promotes the general happiness? It seems intuitive to me that it counts at least *some*, but is it more than weak evidence? And what extra context is most useful in strengthening or weakening the evidential value—data about how broadly accepted (in different times and cultures) the norm is, or an account of its origins, or...?

Reciprocity. Sometimes people try to ground morality entirely in some notion of reciprocity, arguing that it's *in our best interests* to be good to others because doing so will lead others to treat us better. I think it's pretty obvious why that's an inadequate theory: we all have opportunities to cheat and benefit ourselves at others' expense. But Sidgwick also points out some less-frequently-mentioned

problems with reciprocity. It can actually encourage immoral behavior: "...a man is not useful to others by his virtue only, but sometimes rather by his vice; or more often by a certain admixture of unscrupulousness with his good and useful qualities."⁸ (I'm thinking of lawyers who lie on behalf of clients, politicians who accomplish constituents' goals using dirty tactics, CEOs who increase shareholder profits by exploiting workers...⁹) "And further, morality prescribes the performance of duties equally towards all... but on the principle of Reciprocity we should exhibit our useful qualities chiefly towards the rich and powerful... while we may reasonably omit our duties to the poor and feeble... unless they are able to excite the sympathy of persons who can harm us."¹⁰ I would add that keeping the approval of the powerful often *depends* on being willing to overlook injustice against the powerless.

Remorse. Sidgwick thinks the debate about whether we have free will or not doesn't really matter much for ethics. One reason people might think it matters is the worry that determinists won't feel "remorse" and will therefore feel less motivated to avoid wrongdoing. Sidgwick downplays this by saying that people find plenty of motivation to fix various issues in their lives that don't involve remorse, so such motivation should be sufficient for fixing their moral failings too¹¹. I think this raises a philosophical question—*should* belief in determinism prevent people from feeling remorse?—as well as empirical questions: *does* belief in determinism prevent people from feeling remorse? And is remorse important in changing behavior? There [seems to be some relevant research](#)¹².

Notes that I took on each chapter are available [here](#).

1. David Edmonds, *Parfit: A Philosopher and His Mission to Save Morality* (Princeton: Princeton University Press, 2023), 125–26.
2. Henry Sidgwick, *The Methods of Ethics* (Indianapolis: Hackett Pub. Co, 1981), 508.
3. *Ibid.*, 425.
4. *Ibid.*
5. *Ibid.*, 451.
6. *Ibid.*, 176.
7. *Ibid.*, 475.
8. *Ibid.*, 168.
9. *Ibid.*, 169 mentions lying lawyers and politicians taking bribes in a closely-related point.
10. *Ibid.*, 168.
11. *Ibid.*, 71.
12. Tyler F. Stillman and Roy F. Baumeister, "Guilty, Free, and Wise: Determinism and Psychopathy Diminish Learning from Negative Emotions," *Journal of Experimental Social Psychology* 46, no. 6 (November 2010): 951–60, <https://doi.org/10.1016/j.jesp.2010.05.012>.