Notes on *The Methods of Ethics*

Posted by Jacob Williams on brokensandals.net. Last updated 2025-01-11.

Book 1

1. “Introduction”

Sidgwick gives two main ultimate ends that people commonly adopt: Happiness and Perfection/Excellence. The end of Happiness comes in two forms: personal and universal. The most important Excellence is doing what’s right, so he treats this as a form of Intuitionism. The book will explore the three methods implied by these different ends

(Egoism, Utilitarianism, Intuitionism). He's more focused on analyzing each than arguing which is best.

A question he poses: why do we typically include psychology in discussing ethics, but not in e.g. math? He thinks it's mainly because we have the question "Why should I do what I see to be right?" (Sidgwick 5) which is *not* just a request for a motive; rather, people would accept answers based on various ultimate principles but we tend to be unsystematic in their application:

> For if there are different views of the ultimate reasonableness of conduct, implicit in the thought of ordinary men, though not brought into clear relation to each other,—it is easy to see that any single answer to the question 'why' will not be completely satisfactory, as it will be given only from one of these points of view, and will always leave room to ask the question from some other. (Sidgwick 6)

> …it seems easier to judge of the mutual relations and conflicting claims of different modes of thought, after an impartial and rigorous investigation of the conclusions to which they logically lead. (Sidgwick 13)

2. "The Relation of Ethics to Politics"

By Sidgwick's definitions, Ethics is about "what ought to be done by individuals" (Sidgwick 15), while Politics is about "what the government … ought to do and how it ought to be constituted" (Sidgwick 15). There is some connection between them, but Politics is generally concerned with Ideal Law, whereas Positive (actual) Law is more often relevant to our moral obligations.

Sidgwick rejects the idea that Ethics should focus on how we ought to behave in an *ideal* society, both because there may be many radically different ideal societies (see quote below) and because there's no guarantee that right behavior in an ideal society would be a good guide to right behavior in actual society.

> …it seems that when we abandon the firm ground of actual society we have an illimitable cloudland surrounding us on all sides, in which we may construct any variety of pattern states; but no definite ideal to which the actual undeniably approximates, as the straight lines and circles of the actual physical world approximate to those of scientific geometry. (Sidgwick 22)

3. "Ethical Judgments"

Sidgwick considers whether morality is based in reason or only supplies ends for reason; and what "ought" means. He rejects the idea that it is (only) an expression of emotion/preference, or a claim about punishments/consequences.

Interesting argument against emotivism: that we we can have “quasi-moral” feelings that diverge from our moral ones, e.g. a repulsion against lying even when we think we ought to lie.

> …all or most men in whom the moral consciousness is strongly developed find themselves from time to time in conflict with the commonly received morality of the society to which they belong… (Sidgwick 30)

Sidgwick thinks “‘ought’, ‘right’, and other terms expressing the same fundamental notion … is too elementary to admit of any formal definition” (Sidgwick 32).

There’s a narrower sense of ‘ought’ in which it implies ‘can’, and a broader one in which it doesn’t (but maybe still implies we can take some steps toward ‘an approximation’?)

Cognition of an ‘ought’ gives a motive to act, but other motives may overpower it.

Even Egoism requires an irreducible ‘ought’ (though he raises an objection on p. 36-37 that he doesn’t refute here).

Even conditional/hypothetical imperatives involve an irreducible ‘ought’:

> …it also implies the unreasonableness of adopting an end and refusing to adopt the means indispensable to its attainment. (Sidgwick 37)

4. “Pleasure and Desire”

Though Sidgwick grants that pleasure/pain stimulate desire/aversion, he argues that not all desire/aversion is directed toward pleasure/pain.

Some interesting claims/comments:

- Different methods of ethics may be associated with different emotions
- Even if “Psychological Hedonism” (that we actually only ever pursue pleasure / avoid pain) were true, our judgments of right/wrong would still play a role in determining what we find pleasurable
- Desire is not necessarily painful. And there is a difference “between the volitional stimulus of desire itself and the volitional stimulus of aversion to desire as painful…” (Sidgwick 54)

5. “Free Will”

Sidgwick downplays the importance of the determinism vs libertarianism debate for ethics. In either case we need to (and do in practice) rely on predictions of future behavior, and in either case we know such predictions will be imperfect.

Determinism undermines “remorse” in a certain sense, but:

> …it appears to me that men in general take at least as much pains to cure defects in their circumstances, organic defects, and defects of intellect—which cause them no remorse—as they do to cure moral defects; so far as they consider the former to be no less mischievous and no less removable than the latter. (Sidgwick 71)

Determinism also means punishment can’t be about retribution, but Sidgwick thinks it should only be about deterrence and reformation anyway. Interesting footnote on that:

> Thus we find it necessary to punish negligence, when its effects were very grave, even when we cannot trace it to wilful disregard of duty; and to punish rebellion and assassination none the less although we know that they were prompted by a sincere desire to serve God or to benefit mankind. (Sidgwick 72)

6. “Ethical Principles and Methods”

On two other ultimate ends:

- Divine will is to some extent out of scope (the subject of theology) and otherwise our way of determining it collapses to some other method/end.
- Conformance to nature can’t be made precise in any way that makes it a good theory.

Sidgwick says that the same *methods* might be used for multiple *ends*, and he’s choosing to categorize by method, which is why he puts Perfectionism and Intuitionism/Duty together.

Sidgwick ultimately sees Utilitarianism as closer to e.g. Intuitionism than to Egoism, though I like this phrasing of one reason to see affinity with the latter:

> …a Universalistic Hedonist may reasonably hold that his own happiness is that portion of the universal happiness which it is most in his power to promote, and which therefore is most especially entrusted to his charge. (Sidgwick 84)

7. “Egoism and Self-Love”

Sidgwick notes there are multiple ways of conceiving of self-interest; self-preservation is not the same as seeking one’s own happiness. He thinks the latter conception of Egoism is more relevant. He also thinks Hedonism—presumably of both the Egoistic and Utilitarian forms—is best conceived of as seeking *only* to maximize quantity of pleasure. If we also evaluated e.g. whether some pleasures “are nobler and more elevated” we would have “a perplexing mixture of Intuitionism and Hedonism” (Sidgwick 95).

8. “Intuitionism”

What is the ‘Intuitional’ method referring to? Sidgwick notes that even Hedonism relies on intuition in a wider sense (the intuition that you should seek pleasure), but he’s concerned with a narrower sense: intuiting rightness without consideration of consequence.

Intuitionism has three forms or phases:

1. direct judgments of conscience based simply on considering the case at hand (Sidgwick 100)
2. a system of rules (Sidgwick 101)
3. “deeper explanation” of the rules (Sidgwick 102) — note, footnote says this may involve intuition in the “wider” rather than “narrower sense”

9. “Good”

Sidgwick distinguishes the modern view of ethics involving a duty/law to be adhered to, from the Greek view where virtue / virtuous conduct is just one “good” that must be compared with other goods. “Good” isn’t definitionally equivalent to “desired” or to “pleasurable”. Sidgwick does seem to think it can be defined in a sort of value-neutral way:

> …a man’s future good on the whole is what he would now desire and seek on the whole if all the consequences of all the different lines of conduct open to him were accurately foreseen and adequately realised in imagination at the present point of time. (Sidgwick 112)

…but IIUC he prefers a more normative definition, saying “the calm desire for my ‘good on the whole’ is *authoritative*; and therefore carries with it implicitly a rational dictate to aim at this end…” (Sidgwick 112)

Sidgwick thinks only things that contribute to happiness or to human perfection/excellence can be good. (There’s an interesting discussion here about whether something like beauty or knowledge, if we acknowledge its value depends on humans being able to experience it, could still be worth doing for its own sake rather than for the potential impact on humans; Sidgwick says no.)

Book 2: “Egoistic Hedonism”

1. “The Principle and Method of Egoism”

Sidgwick notes that belief in Egoism is widespread, often even viewed as the motive for following Christianity (though of course there are opposing strands of Christian thought).

> …it is hardly going too far to say that common sense assumes that 'interested' actions, tending to promote the agent's happiness, are *prima facie* reasonable: and that the *onus probandi* lies with those who maintain that disinterested conduct, as such, is reasonable. (Sidgwick 120)

Once again Sidgwick mentions he's going to focus on the version of Egoism concerned only with quantity of pleasure minus quantity of pain (he presents no further arguments for that here).

Interestingly, he points out that accepting one's own greatest happiness as a goal doesn't necessarily mean you'll use empiricism to pursue it; you might rely on religion or on *a priori* reasoning to conclude that certain actions are in your best interest. But empirical observations are usually at least part of it, and he's going to focus on the "Empirical-reflective" method first.

2. "Empirical Hedonism"

Hedonism assumes that pleasure/pain "have determinate quantitative relations to each other" (Sidgwick 123). (He mentions the possibility of what today we might call *value lexicality*; he doesn't believe in it, but also thinks it would not be a problem. He also briefly mentions the issue of *uncertainty* but thinks it's easily handled—I'm not sure if he's assuming we should use expected value calculations, or what.)

How can we measure how pleasurable something is? It's not just how strong a desire we feel for it; sometimes "exciting pleasures are liable to exercise, even when actually felt, a volitional stimulus out of proportion to their intensity as pleasures" (Sidgwick 127) and "some feelings which stimulate strongly to their own removal are either not painful at all or only slightly painful:—*e.g.* ordinarily the sensation of being tickled." (Sidgwick 127)

> I propose … to define Pleasure—when we are considering its "strict value" for purposes of quantitative comparison—as a feeling which, when experienced by intelligent beings, is at least implicitly apprehended as **desirable** or—in cases of comparison—preferable. (Sidgwick 127, bold added)

He points out an apparent contradiction between this and the idea (granted in the last chapter) that a pleasure of lesser *quantity* could be preferable on grounds of quality. IIUC, we should understand him here as using 'quantity' to mean something like intensity, and the contradiction is: if intensity is determined by preferability, how could something have less intensity yet be more preferable? His answer:

> …when one kind of pleasure is judged to be qualitatively superior to another, although less pleasant, it is **not really the feeling itself that is preferred**, but something in the mental or physical conditions or relations under which it arises, regarded as cognisable objects of our common thought. (Sidgwick 128, bold added)

I'm a little unclear on what "desirable" means. Is it that there's an objective truth about how 'desirable' a given experience is? Or are we just concerned with what an individual perceives as desirable? I think he's saying the latter:

> …the statement that 'Pleasure is the Ultimate Good' will only mean that nothing is ultimately desirable except desirable feeling, apprehended as desirable by the sentient individual at the time of feeling it. (Sidgwick 129)

3. "Empirical Hedonism—Continued"

Sidgwick considers "*intrinsic* objections to Egoistic Hedonism; arguments… against the possibility of obtaining by it the results at which it aims." (Sidgwick 132) First he dismisses some by someone named Green that I didn't fully follow.

Then he talks about the need for aiming at things other than pleasure itself. This isn't a problem, since we have a tendency to become directly interested in things we're pursuing even if we originally were only pursuing them as part of some other goal; we could even adopt goals arbitrarily if necessary and develop a desire for them. But anyway, we tend to naturally have various desires other than pleasure.

There might be a special pleasure of "absolute self-renunciation and self-sacrifice" (Sidgwick 138) that is unavailable to Egoistic Hedonists, but that pleasure is rarely achieved anyway.

Paying attention to pleasure, as the hedonist must do in order to figure out what's better and worse, may reduce the pleasure itself, but Sidgwick doesn't think this is a big issue. (Sidgwick 140)

Making precise comparisons between different pleasures/pains is difficult. Sidgwick thinks we can trust our immediate evaluation that something we're currently experiencing is pleasurable, but to compare it with another pleasure we must imagine the other. There is reason to doubt that we do this correctly. The same individual will make different judgments over time. We know our state of mind/body affects how attractive any given pleasure seems (or how frightening a pain seems). How much we enjoy things may change over time; also our ability to enjoy—or to endure without suffering—some things may be an acquired skill, and the skills may be differently available to different people. We can't be sure how differently we experience things from others—Sidgwick says this is the problem with Plato's argument for trusting the judgment of a philosopher who experience both intellectual and sensual pleasure. Our memory may mislead us as past experience gets further in the past—Sidgwick points out this can render older people's advice to younger people misguided. Culture can also, without us noticing, give us beliefs about what's pleasurable that don't match our actual feelings.

It's interesting that his description of the Empirical-reflective method emphasizes the for **"making a number of observations and imaginative comparisons, at different times and in different moods"**. (Sidgwick 147)

But still, while “the Empirical-reflective method” isn’t useless, “it would be at least highly desirable… to control and supplement… by the **assistance of some other method**…” (Sidgwick 150, bold added)

Aside: this is a curious statement: “There seem to be no special states of aversion, determined by bodily causes, and related to certain pains as our appetites to their correspondent pleasures…” (Sidgwick 145)

4. “Objective Hedonism and Common Sense”

Here Sidgwick considers how much we can rely on the common opinions of society to tell which things are most pleasurable. This has difficulties such as: most people not even having much time to explore different sources of pleasure; the influence of moral and aesthetic judgments; inconsistency between what people believe is best for them and what they actually do; disagreement among people; and a lack of clear ranking of different pleasures in common opinion. Sidgwick still thinks we can learn from common opinion, but it isn’t a sufficient complement to the Empirical-reflective method.

5. “Happiness and Duty”

Ignoring any possibility of rewards in an afterlife, does self-interest always coincide with moral duty? Sidgwick reaches the (to me) obvious conclusion of No (though he does think it could in a society with certain ideal characteristics). Some interesting asides:

- It’s possible (but implausible) to hold that virtue maximizes self-interest *even if* the world is unjust and unvirtuous people are often better off than virtuous ones; it could still be true that of the outcomes achievable for any given individual, virtue produces the best. (Sidgwick 163)

- While the benefits of reciprocity can incentivize us to be moral, they can also incentivize immorality. “…on the principle of Reciprocity we should exhibit our useful qualities chiefly towards the rich and powerful, and abstain from injuring those who can retaliate…” (Sidgwick 168)

6. “Deductive Hedonism”

Sidgwick considers some physical and psychological theories about the origins of pleasure and pain, but finds none of them to be fully convincing or very useful for the purpose of helping us plan. (This chapter feels especially dated and uninteresting.) He concludes we’re mostly stuck with the Empirical-reflective method, despite its limitations.

Book 3: “Intuitionism”

> …throughout these chapters I am not trying to prove or disprove Intuitionism, but merely by reflection on the common morality which I and my reader share, and to which appeal is so often made in moral disputes, to obtain as explicit, exact, and coherent a statement as possible of its fundamental rules. (Sidgwick 216)

1. “Intuitionism”

> A dubious guidance to an ignoble end appears to be all that the calculus of Egoistic Hedonism has to offer. (Sidgwick 200)

A strange point: that the lack of precise prescriptions from Egoism (as discussed in the previous book) might make following moral intuition more appealing. Sidgwick suggests moral intuition is clearer, but also at the end of the chapter points out how the general rules we try to articulate for Common Sense morality don’t actually give clear guidance in difficult situations. I don’t really see a difference: imo, both self-interest and morality *sound* simple to the unreflective person, and then get complex when we think in detail.

There’s an intriguing quote from Butler who seems to see both self-interest and morality as obligatory, but resolves the conflict by saying we can’t be sure what is/isn’t in our best interests and therefore should follow morality. (Sidgwick 200)

There’s some discussion about what we should take into account to judge right vs wrong; compare Parfit’s discussion in *On What Matters* volume 1. Sidgwick focuses primarily on “intention”; note that he thinks foreseen side-effects that we don’t specifically desire still count as intended.

But there’s room, he says, for a range of opinions on how our motives affect rightness; some think we must act purely from duty, some think we must act from self-love, some are in between. (Sidgwick 205–06) He makes a distinction between “formal” and “material” rightness; the former is concerned with whether someone “chooses duty for duty’s sake” (Sidgwick 207). He says “there is no reason why the same principles and methods for determing material rightness… should not be adopted by thinkers who differ most widely on the question of formal rightness”. (Sidgwick 207)

He also distinguishes “subjective” and “objective” rightness—“what [someone] believes to be right from what really is so” (Sidgwick 207)—but doesn’t think a whole lot needs to be said about subjective rightness.

Sidgwick mentions Kant’s universal-maxim test, and seems to think it’s a necessary but not sufficient condition for true moral rules. (Sidgwick 209–10)

Great word: “psychogonical”. Seems to mean something like *psychological origin*.

Would knowing where moral intuitions come from—their psychological causes—tell us that they are or aren’t valid? Sidgwick thinks that, at least by default, it wouldn’t.

- Just knowing there's a cause for a belief doesn't undermine it; if it did, all our knowledge would be undermined. So the burden of proof is on someone who wants to say our moral intuitions *are* undermined. He thinks this is impossible to do in a fully general way, because "ethical propositions… cannot be inconsistent with any physical or psychological conclusions." (Sidgwick 213)

- But knowing the cause also isn't likely to show that "our moral cognition… would be especially free from error", given that we already have reasons to think it's fallible. (Sidgwick 214)

Sidgwick wants to figure out what "the morality of Common Sense" (Sidgwick 215) tells us; this refers to

> …a body of moral truth, warranted to be such by the *consensus* of mankind,—or at least of that portion of mankind which combines adequate intellectual enlightenment with a serious concern for morality… (Sidgwick 215)

(If I understand correctly, we're specifically talking about "a body of moral truth" that tries to generalize ordinary moral intuitions, as opposed to ones like utilitarianism that start directly from intuitions about what the correct system should be; although I'm not clear how e.g. Kant's system stands in relation to this.)

2. "Virtue and Duty"

Sidgwick asks what "virtue" and "duty" mean and how they relate to each other. Ultimately he indicates he'll mostly be concerned with right action, and the distinction won't matter much for his purposes, but he discusses various nuances first. One observation is we only really talk of "duty" when the act in question requires some degree of "moral impulse" as opposed to coming totally naturally (Sidgwick 217). With "virtue", though, we have two competing intuitions —that it's more virtuous to overcome one's desires, and that it's more virtuous to have desires aligned with morality; thus there may really be two separate concepts here. (Sidgwick 225) Also, some virtues are thought to be enhanced when backed by particular emotions (Sidgwick 223).

Do we have a duty to do everything virtuous? Sidgwick notes some tension here too, and gives a couple reasons why we sometimes don't judge someone for failing to do what would be virtuous:

- we don't know their situation well enough to be sure they ought to do it

- for practical purposes: "we think that moral progress will on the whole be best promoted by our praising acts that are above the level of ordinary practice, and confining our censure…to acts that fall clearly below this standard" (Sidgwick 221); notably, this means the "standard" will "vary as the average level of morality varies…" (Sidgwick 221)

3. “Wisdom and Self-Control”

> Wisdom… appears to me to imply right judgment in respect of ends as well as means. (Sidgwick 232)

Our ordinary notion of wisdom may make some assumptions about the compatibility of multiple ultimate aims:

> Common Sense seems to mean by a Wise man, a man who attains at once all the different rational ends; who by conduct in perfect conformity with the true moral code attains the greatest happiness possible both for himself and for mankind (or that portion of mankind to which his efforts are necessarily restricted). (Sidgwick 233)

Is wisdom voluntary? “It is often said…that the cognition of Moral truth depends largely upon the ‘heart’” (Sidgwick 233), but there’s disagreement on what you’re supposed to do to get correct moral cognition. There is agreement that it involves some self-control of “certain violent passions and sensual appetites” (Sidgwick 234) and “desire and fear” (Sidgwick 234), though.

4. “Benevolence”

Sidgwick thinks the common sense notion of benevolence includes both giving benefits to others, and to some degree cultivate feelings of affection for them. The benefits might be toward their happiness or toward making the recipient more virtuous—he focuses on the former as the more important component, but thinks common sense morality is open to the latter too. (There’s a short remark about how Kant thought we couldn’t possibly help someone else achieve virtue—Sidgwick basically points out that the same reasoning would mean we can’t help ourselves develop virtue over time. (Sidgwick 240))

Compared with Utilitarianism, Common Sense’s notion of benevolence expects us to be partial to people based on our relationships to them. (Sidgwick 242) (In listing a tentative hierarchy of partiality, he tentatively includes loyalty to race… yikes. (Sidgwick 246)) Sidgwick notes how expectations change over time and (I assume correctly, unless things in England are much different than here) predicts that obligations to leave inheritances to relatives other than your own children are going to continue to weaken substantially.

He takes the fact that Common Sense allows for sometimes rejecting customs, as evidence that its duties can/should(?) be justified on grounds of other principles (e.g. Utilitarianism). (Sidgwick 247)

Four categories of benevolence duties: (Sidgwick 248)

1. “duties arising out of comparatively permanent relationships not voluntarily chosen, such as Kindred and in most cases Citizenship and Neighbourhood”
2. “those of similar relationships voluntarily contracted, such as Friendship”

3. “those that spring from special services received, or Duties of Gratitude”

4. “those that seem due to special need, or Duties of Pity”

In all cases he finds that while there is widespread agreement that some duties exist, their limits and details are nebulous.

An interesting example he mentions: how are our special duties to friends affected by changes in our feelings toward the friend? Are the duties lessened? Do we have any duty to try to prevent such change in feeling? (Sidgwick 258)

5. “Justice”

> The prominent element in Justice as ordinarily conceived is a kind of Equality: that is, Impartiality in the observance or enforcement of certain general rules allotting good or evil to individuals. (Sidgwick 293)

As you’ll expect by now, he argues we can’t get make this sufficiently detailed by relying on Common Sense.

Sidgwick sees a tension in “Conservative Justice” and “Ideal Justice”. The former refers to the expectations we form based on existing laws/customs and our sense that we are entitled to have those expectations met. The latter refers to the fact that we sometimes judge laws/customs themselves to be unjust.

He further argues that Common Sense doesn’t have a clear enough idea of Ideal Justice to give us a detailed picture of an ideal society. For example, many are drawn to the idea that government should guarantee the Freedom of all individuals, but in practice this requires lots of caveats and tradeoffs.

6. “Laws and Promises”

Sidgwick finds Common Sense indeterminate on many issues about what laws we are obligated to obey and how we recognize a legitimate government. Also, our obligation to obey promises is only clear when a long list of conditions are met; lots of things introduce ambiguity.

A nice quote on the difficulty of establishing government legitimacy even in democracies:

> …it is paradoxical to affirm that the freedom and natural rights of a dissentient minority are effectively protected by establishing the condition that the oppressors must exceed the oppressed in number. (Sidgwick 298)

An example I liked about the complexity of the interaction of duty to obey law and other duties:

> It has sometimes been said that the Law cannot override definite duties; but the obligation of fidelity to contract is peculiarly definite, and yet we do not consider it right to fulfil a contract of which a law, passed subsequently to the making of the contract, has forbidden the execution. (Sidgwick 301–02)

There's also a fun thought experiment about promises: suppose you promise to do something which both you and the promisee initially believe will benefit the promisee. Then you learn that it would actually harm them greatly, but this doesn't change their mind about it. Should you still fulfill the promise? (Sidgwick 308)

7. "The Classification of Duties—Veracity"

Sidgwick explains that although there's a temptation to divide duties into "Social and Self-regarding", (Sidgwick 312) this doesn't really match the Intuitionist way of thinking.

IIUC, Veracity is discussed here because it's an example of a duty which, on the Common Sense understanding, has an unclear relation to selfish vs social interests. There is not consensus on when dishonesty is or isn't allowed, or what qualifies as dishonesty.

Sidgwick gives a few replies to the idea (which I associate with Kant) that we should never lie because social acceptance of lying would undermine the possibility of telling lies (and being believed); one well-stated one:

> In the first place it is not necessarily an evil that men's confidence in each other's assertions should, *under certain peculiar circumstances*, be impaired or destroyed: it may even be the very result which we should most desire to produce: *e.g.* it is obviously a most effective protection for legitimate secrets that it should be universally understood and expected that those who ask questions which they have no right to ask will have lies told them… (Sidgwick 318)

8. "Other Social Duties and Virtues"

Sidgwick notes that Common Sense has somewhat unclear and diverging opinions when "malevolent affections" (Sidgwick 320) are allowed or desirable. He also briefly discusses Liberality, Generosity, and Chivalry as elements of Benevolence.

9. "Self-Regarding Virtues"

> I conceive that according to the morality of Common Sense, an ultimate harmony between (1) Self-interest and (2) Virtue is assumed or postulated… (Sidgwick 327)

This talks a little about Prudence, Temperance, Purity, and Chastity; I didn't get an overall point from it.

10. “Courage, Humility, Etc.”

Courage and Humility are virtues which, again, are only approved of up to certain unclear limits.

11. “Review of Common Sense”

Sidgwick gives “four conditions, the complete fulfilment of which would establish a significant proposition, apparently self-evident, in the highest degree of certainty attainable…” (Sidgwick 338):

1. “The terms of the proposition must be clear and precise.” (Sidgwick 338)
2. “The self-evidence of the proposition must be ascertained by careful reflection.” (Sidgwick 339)
3. “The propositions accepted as self-evident must be mutually consistent.” (Sidgwick 341) (He complains of “ethical writers treating this point very lightly.” (Sidgwick 341))
4. Universality or consensus: “…the denial by another of a proposition that I have affirmed has a tendency to impair my confidence in its validity.” (Sidgwick 341)

He reviews his discussion of the duties/virtues from previous chapters, along with additional discussion of Purity/Chastity, to argue that Common Sense does not give us a set of principles that meet these conditions. (This chapter is very repetitive of the preceding ones.)

12. “Moral Judgment of Motives”

What if instead of looking at Common Sense’s judgments about whether people’s intentional actions are right or wrong, we looked at Common Sense’s judgments about what motives are better or worse? Would this give a clear hierarchy that would let us resolve moral disputes better? No, Sidgwick argues.

13. “Philosophical Intuitionism”

> Utilitarianism is thus presented as the final form into which Intuitionism tends to pass, when the demand for really self-evident first principles is rigorously pressed. (Sidgwick 388)

Sidgwick states some “absolute practical principles, the truth of which … is manifest”: (Sidgwick 379)

- a more general version of the Golden Rule: “…whatever action any of us judges to be right for himself, he implicitly judges to be right for all similar persons in similar

circumstances." (Sidgwick 379)

- we should have "impartial concern for all parts of our conscious life" (Sidgwick 380)—future moments are no less intrinsically valuable then the present moment
- "…the good of any one individual is of no more importance, from the point of view (if I may say so) 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 realised in the one case than in the other. And it is evident to me that as a rational being I am bound to aim at good generally…not merely at a particular part of it." (Sidgwick 382)

He thinks these "present themselves as self-evident; as much (*e.g.*) as the mathematical axiom that 'if equals be added to equals the wholes are equal.'" (Sidgwick 383) This is *in contrast to* statements like "I ought to speak the truth" which "present themselves as propositions requiring rational justification of some kind." (Sidgwick 383)

I find this to be an underwhelming culmination. Sidgwick criticizes Mill for having an inadequate argument "that the general happiness is desirable" (Sidgwick 388) (as opposed to merely each individual's happiness being desirable to the individual), and IIUC the 3rd proposition above is meant to plug the "gap" (Sidgwick 388). But how many people would really think Sidgwick's 3rd proposition here, which seems to amount to saying that we each ought to take the point of view of the universe, is really self-evident?

14. "Ultimate Good"

Part of Sidgwick's argument that Intuitionism leads toward Utilitarianism, which also came up earlier but is stated relatively clearly at the beginning of this chapter, is "that most of the commonly received maxims of Duty…are found when closely examined to contain an implicit subordination to the more general principles of Prudence and Benevolence…" (Sidgwick 391)

He thinks he's established that we should aim at promoting the Good of all; but what is the Good? He rejects identifying it with Virtue, or anything like that, on the grounds of circularity. He argues that only Happiness can be a primary Ultimate Good. Other things like Truth and Beauty are also rational to pursue directly, but only because of their consequences and/or the necessity of focusing on something other than happiness in order to achieve happiness.

Book 4: "Utilitarianism"

1. "The Meaning of Utilitarianism"

> …the conduct which… is objectively right, is that which will produce the greatest amount of happiness on the whole… (Sidgwick 411)

Sidgwick thinks the happiness of animals counts, too (I'm a little surprised how briefly he makes this point) (Sidgwick 414). So do future humans; and Sidgwick says something along the lines of the Repugnant Conclusion (and seems to think he's the first to write about the point) although he doesn't note that this could lead to most lives being barely worth living, which is maybe why he doesn't seem bothered by it.

Sidgwick points out "that there may be many different ways of distributing the same quantum of happiness among the same number of persons" (Sidgwick 416)—I assume this means, e.g., you could give 100 hedons to 1 person or 1 hedon to each of 100 people. He endorses a principle "of pure equality" (Sidgwick 417), which IIUC means you should just divide it evenly.

2. "The Proof of Utilitarianism"

Sidgwick suggests Egoists could be challenged to explain why, on the one hand, they think they should consider both their own present *and* future well-being; but, on the other, they don't think they should consider both their own and *other people*'s well-being. Humean ideas about the self being an illusion make the challenge particularly sharp. (I personally think this is the most important theoretical argument against egoism, but Sidgwick decides not to dwell on it.)

He describes another way of trying to argue someone from one moral framework to another:

> Perhaps … what is needed is 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. (Sidgwick 420)

He says book 3 chapter 13 gave "[s]uch a line of argument, addressed to Egoism" (Sidgwick 420); I'm confused by this because I thought that was addressed to people who accept Intuitionism but not pure Egoism. (Anyway, Sidgwick says this only applies when Egoist's claim is "that his happiness or pleasure is Good, not only *for him* but from the point of view of the Universe,—as (*e.g.*) by saying that 'nature designed him to seek his own happiness,'…" (Sidgwick 420–21). I suspect most people today would not frame Egoism that way.)

Next he needs to complete that argument by showing that Utilitarianism fills that role of "more comprehensive principle" for Common Sense morality well. (Sidgwick, see p. 422)

3. "Relation of Utilitarianism to the Morality of Common Sense"

Sidgwick claims that Common Sense morality reflects a basic valuing of pleasure and that Utilitarianism explain/justifies much about Common Sense morality—e.g. Common Sense is ambivalent on some things such as lying to criminals precisely because there are utilitarian considerations pointing in different directions.

He identifies “three distinct lines of argument” for Utilitarianism to endorse deviations from pure impartiality to take account of special relationships between individuals, related to: the value of the “affections” themselves; the value of not disappointing humans’ natural expectations; and various “utilitarian reasons why, generally speaking, services should be rendered to the persons commonly recognised as having such claims rather than to others.” (Sidgwick 439)

A lot of the reasoning Sidgwick uses here makes sense to me but I also feel like you could probably come up with utilitarian stories to justify the exact opposite: this raises both the worry the we won’t be able to get concrete guidance out of utilitarianism, and the worry that we can use it to rationalize any conclusion we want. This is perhaps illustrated in how he justifies the sexual mores of his time (Sidgwick, see p. 451).

Sidgwick talks about how different moral codes of different societies/times often reflect differences in what’s beneficial because of the society’s circumstances. Also, he thinks “the capacity for sympathy in an average member of the community” (Sidgwick 455) varies across history.

4. “The Method of Utilitarianism”

Sidgwick does *not* think we should assume Common Sense morality tracks Utilitarianism perfectly. So how can we create an optimal code?

Difficulties include the need to account for humans as they actually are—including their existing moral beliefs/feelings, which can’t be magically changed instantly—and the difficulty of accounting for how society and culture are constantly changing (so that it’s difficult to predict what the ideal code will be in the future).

> …the utilitarian, in the existing state of our knowledge, cannot possibly construct a morality *de novo* either for man as he is (abstracting his morality), or for man as he ought to be and will be…. if [the utilitarian] keeps within the limits that separate scientific prevision from fanciful Utopian conjecture, the form of society to which his practical conclusions relate will be one varying but little from the actual, with its actually established code of moral rules and customary judgments concerning virtue and vice. (Sidgwick 473–74)

5. “The Method of Utilitarianism—Continued”

Sidgwick thinks utilitarians should be “much less concerned with correcting and improving [society’s moral code] than … with realising and enforcing it.” (Sidgwick 475) But sometimes it is worth trying to improve. He discusses some considerations to take into account when deciding whether to try to change society’s morals in some way:

- how the pushback you’ll get will effect you and your future level of influence (Sidgwick 481)

- how likely your new rule is to gain wide adoption, and whether the (possibly failed) attempt to promote it will "give a certain aid to the forces that are always tending towards moral anarchy" (Sidgwick 482)

- whether deviating from your own habits and the morals encouraged by your society may weaken your moral discipline (Sidgwick 482–83)

Sidgwick notes that merely adopting a *stricter* rule than whatever is accepted in Common Sense morality, is not risky, *as long as* you only do so for yourself rather than pushing the rule on society generally.

What about when an individual utilitarian thinks it would be beneficial to deviate from a general moral rule? Sometimes this really just means we should conceive of the rule in a more nuanced way. (Sidgwick 485) But what about cases where you really would want most people to follow the rule in all circumstances, but there's no harm (and some benefit) in just having a few people break it? Sidgwick thinks this is OK in theory, and uses celibacy as an example that Common Sense even approves of (universal celibacy would be bad, but there's no danger of that happening, so individuals being celibate is fine). (Sidgwick 487) But he thinks that's an unusual case; normally it would be hard to have sufficient confidence that your own deviation from the rule wouldn't influence others to a dangerous level of deviation. (Sidgwick 488)

There's some discussion of the possibility that utilitarians ought to conceal their true opinions on certain moral questions from the public (Sidgwick 489)—this strikes me as pretty elitist and also, in the absence of specific examples, kind of silly.

Although utilitarianism doesn't fundamentally distinguish between the obligatory and the supererogatory, Sidgwick thinks it's still useful to make such a distinction; "…it is natural to us to compare any individual's character or conduct, not with our highest ideal…but with a certain average standard and to admire what rises above the standard…" (Sidgwick 492).

Interesting ending:

> …one who values conduct in proportion to its felicific consequences, will naturally set a higher estimate on effective beneficence in public affairs than on the purest manifestation of virtue in the details of private life: while on the other hand an Intuitionist… still commonly holds that virtue may be as fully and as admirably exhibited on a small as on a large scale. A sincere Utilitarian, therefore, is likely to be an eager politician… (Sidgwick 495)

"Concluding Chapter": "The Mutual Relations of the Three Methods"

Sidgwick revisits an argument from Book 4 chapter 2 (actually, I thought it was two separate arguments, but he seems to treat them as one here) meant to argue Egoists into Utilitarians, and notes its weaknesses. He says that "for morality…to be made completely rational" (Sidgwick 498), we'd need to show it always coincides with self-interest. He doesn't seem optimistic this can be done, unless there is a God to ensure it. Otherwise, when there are genuine "cases of a recognised conflict between self-interest and duty… the conflict would have to be decided by the comparative preponderance of one or other of two groups of non-rational impulses." (Sidgwick 508) The final paragraph hints, IIUC, that certain views on epistemology might provide an alternate justification for us to believe in "the reconciliation of duty and self-interest" (Sidgwick 508).

I'm curious what "Butler's famous argument against the vulgar antithesis between Self-love and Benevolence" (Sidgwick 501) is. From what Sidgwick goes on to say, it sounds like something to do with the benefits of sympathetically enjoying others' welfare.

Appendix: "The Kantian Conception of Free Will"

This is an excerpt from a journal article of Sidgwick's. He distinguishes three notions of freedom—"Good" or "Rational", "Neutral" or "Moral", and "Capricious" (Sidgwick 512)—and argues that Kant slides between the first two in problematic ways.

References

Sidgwick, Henry. *The Methods of Ethics*. Hackett Pub. Co, 1981.