Sidgwick on Consequentialism and Deontology: A Critique

THOMAS HURKA
University of Toronto

In *The Methods of Ethics* Henry Sidgwick argued against deontology and for consequentialism. More specifically, he stated four conditions for self-evident moral truth and argued that, whereas no deontological principles satisfy all four conditions, the principles that generate consequentialism do. This article argues that both his critique of deontology and his defence of consequentialism fail, largely for the same reason: that he did not clearly grasp the concept W. D. Ross later introduced of a prima facie duty or duty other things equal. The moderate deontology Ross's concept allows avoids many of Sidgwick's objections. And Sidgwick's statements of his own axioms equivocate in exactly the same way for which he criticized deontological ones. Only if they are read as other things equal can they seem intuitive and earn widespread agreement; but that form is too weak to ground consequentialism. And in the form that does yield consequentialism they are neither intuitive nor widely accepted. Sidgwick's arguments against a rival view and for his own were, in multiple ways, unfair.

In 1902 Hastings Rashdall said there is 'a general consensus . . . that Ethics must be "teleological" '[1] or consequentialist, and around the turn of the twentieth century consequentialism was indeed the dominant moral theory. Its adherents, however, defended it in different ways. In *Principia Ethica* G. E. Moore took the consequentialist principle to be analytically true, since ' "right" does and can mean nothing but "cause of a good result" ';[2] this seems at times also to have been Rashdall's view.[3] In his later *Ethics* Moore abandoned this reductive analysis of 'right' and took the consequentialist principle to be synthetic, but then declared baldly that, considered on its own, it is self-evident: 'It seems to me self-evident that knowingly to do an action which would make the world, on the whole, really and truly *worse* than if we had acted differently, must always be wrong.'[4] Others too simply asserted the principle's self-evidence,[5] and though that is one way of defending

[1] Hastings Rashdall, 'The Commensurability of All Values', *Mind* 11 (1902), pp. 145–61, at 148; also *The Theory of Good and Evil*, 2 vols. (London, 1907), vol. 1, p. 217; vol. 2, p. 41.

[2] G. E. Moore, *Principia Ethica* (Cambridge, 1903), p. 147; also pp. 18, 21, 24–6, 146, 148, 167–9, 180–1.

[3] Rashdall, *Theory of Good and Evil*, vol. 1, p. 135; vol. 2, p. 42.

[4] Moore, *Ethics* (London, 1965), p. 77; that the principle is synthetic is affirmed on pp. 25–6, 73, 76.

[5] See e.g. John Laird, *A Study in Moral Theory* (London, 1926), pp. 21–2, 25; H. W. B. Joseph, *Some Problems in Ethics* (Oxford, 1931), p. 26; A. C. Ewing, *The Definition of Good* (London, 1947), p. 188; *Ethics* (London, 1953), pp. 66–7, 82.

© Cambridge University Press 2014
Utilitas Vol. 26, No. 2, June 2014
doi:10.1017/S0953820814000089

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

consequentialism it does not offer much in the way of argument to non-believers. A philosopher who did offer more was Henry Sidgwick. In *The Methods of Ethics*[6] he laid down several conditions that a proposition must meet to be genuinely self-evident and argued, first, that no non-consequentialist principles satisfy all four conditions, and then that the principles underlying consequentialism do. Consequentialism emerged after a long discussion as the one moral view that passes all relevant epistemic tests.

This article will critically examine Sidgwick's arguments. Though they have been highly regarded, I will argue that both his critique of non-consequentialism and his defence of his own principles are seriously flawed. In both cases many of their flaws stem from his not having or taking proper account of the concept, later introduced by W. D. Ross, of a prima facie duty, or duty other things equal.[7]

SIDGWICK'S CONDITIONS FOR SELF-EVIDENCE

Since Sidgwick's arguments concern whether certain principles are 'self-evident', we must first determine what he meant by self-evidence. Sometimes he characterized self-evident truths just as ones that 'when their terms are properly understood, the perception of their absolute truth is immediate and irresistible' (229; also 379).[8] But in book III, chapter 11 of *The Methods* he gave a more elaborate account, listing four conditions 'by which self-evident truths are distinguished from mere opinions' (338).

The first condition says 'the terms of the proposition must be clear and precise'; it is commonly called the clarity condition. The second says 'the self-evidence of the proposition must be ascertained by careful reflection', while the third says 'the propositions accepted as self-evident must be mutually consistent' since 'any collision between two intuitions is a proof that there is error in one or the other, or in both'. The fourth condition requires general agreement on the proposition, or a consensus about it: 'since it is implied in the very notion of Truth that it is essentially the same for all minds, the denial by another of a proposition that I have affirmed has a tendency to impair my confidence in its validity' (338–42).

[6] Henry Sidgwick, *The Methods of Ethics*, 7th edn. (London, 1907). Bracketed page references are to this work; earlier editions are cited as *ME1*, *ME2*, etc. References to other ethical writings of Sidgwick's are to the collection of his *Essays on Ethics and Method*, ed. Marcus G. Singer (Oxford, 2000), abbreviated *EEM*.

[7] W. D. Ross, *The Right and the Good* (Oxford, 1930), ch. 2.

[8] For similar present-day accounts see Robert Audi, *The Good in the Right: A Theory of Intuition and Intrinsic Value* (Princeton, 2004), pp. 48–9; Derek Parfit, *On What Matters*, 2 vols. (Oxford, 2011), vol. 2, pp. 490, 508–9.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

There is an initial puzzle about this account. Its four conditions are supposed to be jointly necessary for self-evidence, but the second tells us to ascertain 'the self-evidence of the proposition'. If this condition by itself establishes self-evidence, what work is left for the others to do?

I think Sidgwick was a little careless in stating his second condition. He often contrasted real with merely apparent self-evidence (34 n. 2, 211, 338, 383, 508; also *EEM* 30, 31–2, 33–4, 170), and his condition is best read as asking us to establish the *apparent* self-evidence of a proposition, or establish that it really is apparently self-evident. The picture is this.[9]

You start by being in what seems to you to be a mental state of apprehending a proposition as true just by understanding it. Your actually apprehending a proposition in this way does not guarantee that it is true, any more than your seeing a line as bent guarantees that it is bent. But you need to confirm that you really are in that mental state, and that is what the second condition requires. Sidgwick's justification for it was precisely that we are often mistaken about our mental states; we tend to confuse genuine intuitions with 'mere impressions or impulses which to careful observation do not present themselves as claiming to be dictates of Reason' or with 'mere opinions' that repetition has given 'a false appearance of self-evidence'. Since 'any strong sentiment, however subjective, is apt to transform itself into the semblance of an intuition' (339; also 211–12, 340–1, *EEM* 25, 31), we need to detect any such semblance. This requires 'careful reflection' (339), which for British philosophers of Sidgwick's day meant introspection or self-examination more generally.[10] Thus whether you really are having an intuition 'can only be decided by each person by direct introspection or reflection' (211; also 212, 362, 383, *EEM* 43), or by that plus a survey of possible distorting causes of your belief, such as common acceptance in your community or a strong desire that the belief be true (339). The second condition therefore requires a self-survey. You are in what seems to you to be a mental state of apprehending a proposition as true just by understanding it, or of apparent self-evidence. You then apply introspective and other tests to determine whether it really is a mental state of that kind. If it is, it involves a real case of apparent self-evidence, and you can then apply the other conditions to see if the proposition is simply, and not just apparently, self-evident.

[9] See also David Phillips, *Sidgwickian Ethics* (Oxford, 2011) p. 60.

[10] An early and influential use of 'reflection' to mean introspection is in John Locke, *An Essay Concerning Human Understanding*, ed. J. W. Yolton (London, 1977), bk. II, ch. 1, para. 4.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

Of these other conditions, the first or clarity condition is closely tied to the second and perhaps inseparable from it. An unclear proposition can lead you to believe falsely that you are having an apparent intuition; more strongly, if a proposition is unclear, it is hard to see how you can have the understanding of it that intuition requires. But the third and fourth conditions involve additional tests. You can apprehend each of two propositions as true and find that they contradict each other, or apprehend one as true and find that other people reject it. The situation when a belief that passes the second test fails the third or fourth can be described in two ways. We can say your initial mental state was not in fact one of intuiting the proposition, so the third and fourth conditions address the same question as the second but by non-introspective means, or we can say you did intuit the proposition but it turned out to be false. The second option, which allows an intuition to be false, seems closer to Sidgwick's texts; it was also preferred in later discussions by A. C. Ewing.[11]

Sidgwick's official statement of his third condition requires an apparently self-evident proposition to be consistent only with other apparently self-evident propositions (341; also *EEM* 166), and with this restriction the condition is compelling. If two equally well-grounded beliefs contradict each other, you have no basis for preferring one to the other and should abandon both. But sometimes he spoke as if a self-evident belief must harmonize with all your beliefs about its subject matter, whether apparently self-evident or not (400; also *EEM* 25, 31), and this stronger condition is less plausible. If an apparently self-evident belief clashes with one that does not meet that condition, should the latter not yield?

His fourth condition put him on one side of a present-day debate in epistemology about 'peer disagreement', and in fact he gave a classic statement of the view that others' disagreement should reduce your confidence in a belief.[12] We can understand his reasoning for the condition as follows. You have a capacity for intuiting self-evident truths that is generally reliable but sometimes leads you astray. If someone with an equally reliable capacity does not share your intuition, one of the two capacities must be malfunctioning, and you have no reason to believe it is not yours (*EEM* 32).

Unfortunately this reasoning conflicts with his own description of intuition as involving 'immediate' knowledge (229), for it implies that

[11] Ewing, *Ethics*, p. 139; Ewing, *Value and Reality* (London, 1973), pp. 43–4.

[12] For a more recent statement see Adam Elga, 'Reflection and Disagreement', *Noûs* 41 (2007), pp. 478–502; the contrary view is defended in Thomas Kelly, 'The Epistemic Significance of Disagreement', *Oxford Studies in Epistemology*, vol. 1, ed. J. Hawthorne and T. Gendler-Szabo (Oxford, 2005), pp. 167–96.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

even in favourable cases you make an inference from the general reliability of a type of mental process to its reliability now. And there is already a conflict with that description in the list of four conditions, since together they make the process of determining self-evidence anything but 'immediate'. Given an initial seemingly intuitive belief, you engage in self-examination, including a survey of possible distorting causes of your belief, test its consistency with other beliefs and check what other people believe. While the conditions as a whole try to secure against error, they also make the determination of self-evidence far from simply a matter of 'immediate and irresistible' perception (229).

At the same time, the conditions make the intuition of moral truth in several ways fallible; certainly your initial sense that a moral judgement seems true is no guarantee that it is, since your belief can fail any of the second, third and fourth conditions. Some commentators have contrasted Sidgwick's fallibilism with the allegedly more dogmatic views of later writers such as Moore, H. A. Prichard, and Ross,[13] but the contrast is overdrawn, since many of those writers likewise insisted that intuition is fallible. Sidgwick's account of self-evidence may have been more systematic than theirs, but its basic content was the same.

THE CRITIQUE OF DEONTOLOGY

The first part of Sidgwick's case for consequentialism was his negative argument that no non-consequentialist principles are self-evident. His target here was what he called 'Dogmatic Intuitionism', a theory that tries to systematize common-sense morality while retaining its deontological character, so it contains both principles about promoting the good and principles constraining how we may do so. His model for it was William Whewell's *Elements of Morality*, from which he said his initial antipathy to deontology derived (xv–xvi).

He sometimes denied that deontological principles satisfy his second condition: they do not even 'present themselves as self-evident', and any appearance that they do is dissolved by 'reflection' (383; *EEM* 25). But he usually gave a more complex argument. If deontological principles are left as 'somewhat vague generalities', such as 'you ought to keep your promises', they do not satisfy the clarity condition, but they do seem self-evident and assent to them is 'approximately universal'. Once we try to give them 'the definiteness which science requires', however, by adding qualifications that will make their application determinate and resolve conflicts between them, the consensus disappears and their

[13] See e.g. Bart Schultz, 'Introduction', *Essays on Henry Sidgwick*, ed. B. Schultz (Cambridge, 1992), pp. 1–61, at 28–9, 59 n. 84.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

self-evidence 'becomes dubious or vanishes altogether' (342–3). In one formulation, in other words, the principles seem to pass the second test of apparent self-evidence and the fourth about consensus but are neither clear nor mutually consistent. In another they pass the first and third tests but are too complicated to pass the second or fourth: the gain in clarity and consistency is at the expense of intuitiveness and universal agreement.

He applied this argument to a multitude of deontological principles: about benevolence, gratitude, distributive and retributive justice, law-abidingness and promise-keeping, truth-telling and more, in a survey running through ten chapters of *Methods* III. Surprisingly, he only briefly mentioned (253, 348) what many philosophers today consider the core of a deontological view: what Ross called the principle of non-maleficence, which makes it more seriously wrong directly to harm another than to fail to give her benefits, for example, more seriously wrong to kill her than to fail to save her life. This principle and the issues it raises barely figured in his discussion.

One aspect of Sidgwick's claim that everyday versions of the principles are unclear concerned cases where they conflict. A pluralist deontology view has to weigh duties such as benevolence and promise-keeping against each other but has no precise rules for doing so. It therefore cannot always decisively judge particular acts and is in that way partly indeterminate. The unclarity here is not in the individual principles, as Sidgwick's statement of his first condition would suggest, but in the theory containing them as a whole. A second aspect did concern the individual principles. Here he argued that they often have vague boundaries, so it is unclear exactly which acts fall within their scope and which do not; there is 'a sort of margin or dim borderland' (270) where we are unsure whether a given duty applies. Common sense says we should care more about those who stand in special relationships to us, but does someone's belonging to your race give you stronger reason to promote his happiness? Do a childless person's siblings have a special claim on his estate (*ME* 246–7)? Though it is sometimes clear that you ought to keep a promise, what if circumstances have changed radically since you made it and the person to whom you made it has died and cannot release you from it? Is the promise still binding (306–7)? Some instances of this second unclarity involved versions of the first. Sidgwick often argued that a common-sense principle has conflicting internal elements. To assess the strength of a duty of gratitude, for example, should you look to the size of the benefit another gave you or to the degree of effort he expended (261)? Distributive justice likewise has one side telling us to honour any expectations we have created and so to retain any social rules, for example about economic life, that have long been in place, and another proposing an ideal of justice and

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

telling us to reform any practices that hinder its attainment (273). Unless a deontological view can in each case choose one element in total preference to the other, its judgements about gratitude and justice must weigh competing claims and cannot always do so precisely.

In elaborating these charges Sidgwick made many acute observations about common-sense morality and raised telling objections to particular deontological claims. But our question is how successfully his arguments show the unacceptability of deontology in general or a fundamental failing in it.

Rashdall thought the arguments entirely successful:

> The loose statements of Intuitionists as to the clearness, certainty, adequacy, and self-evidence of the ordinarily received rules of conduct have never been subjected to so searching, so exhaustive, and so illuminating an examination. That task has been done once for all, and need not in detail be done over again.[14]

F. H. Hayward agreed:

> Sidgwick's discussion of this 'common sense' doctrine is admitted by all critics to be extremely able, to be, in fact, the most irrefutable part of the book. . . . Most of the third book requires no commentary whatever; by common consent its chapters are so lucid, and the conclusions they embody are, for the most part, so indisputable, that even to point out their merits would be to gild refined gold.[15]

But in fact Sidgwick's critique was both unpersuasive and unfair.

Many of its flaws stemmed from his not clearly grasping the concept of a prima facie duty and therefore not seriously considering the moderate, as against absolute, deontology it makes possible. This is a deontology in which the principles that sometimes make an act with the best consequences wrong, such as 'You ought to keep your promises', are not absolute or inviolable. They can sometimes, though present in a situation, be outweighed by a competing principle, so the act they recommend is all things considered wrong. Thus in an example of Ross's your duty to keep a promise to meet a friend can still be prima facie binding, because not voided by any exception-clause, but outweighed by a stronger duty of beneficence if you could help the seriously injured victims of an accident you have come upon. And its being still binding is shown by the facts that you should feel, if not guilt, then 'compunction' about breaking your promise, and may owe your friend some compensation for doing so.[16]

That Sidgwick lacked Ross's concept is shown, first, in his general characterization of deontology. He defined this as a view that assumes

[14] Rashdall, *Theory of Good and Evil*, vol. 1, p. 83.
[15] F. H. Hayward, *The Ethical Philosophy of Sidgwick* (London, 1901), pp. viii, xviii.
[16] Ross, *Right and the Good*, pp. 18, 28.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

we can see certain acts as 'right and reasonable in themselves, apart from their consequences; – or rather with a merely partial consideration of consequences, from which other consequences admitted to be possibly good or bad are definitely excluded', or without considering consequences 'except in so far as these are included in the common notion of the act' (200, 337). This last qualification was an astute response to the fact that, as he had noted earlier (96–7), the same act can be described in different ways, some referring to consequences that are not mentioned in others; compare calling an act a 'killing' with calling it a 'pulling of the trigger'. But his talk of seeing acts as 'right and reasonable in themselves' was ambiguous. If he meant seeing them as prima facie right or right other things equal, his description of deontology was sound. If all acts of killing are prima facie wrong, then to know an act is prima facie wrong we need only know that it caused a death and need not know its longer-term effects. But if by 'right' he meant right all things considered, his description fits only absolute deontologies; a moderate deontology must always know a killing's further effects, since they may contain enough good to outweigh its wrongness *qua* killing. Sidgwick did not distinguish these readings, which at the least made his characterization less than ideally clear. And several of his statements point to the all-things-considered one, which fits only absolute deontologies. In one passage his examples of deontological principles were 'that duty should be done "advienne que pourra," that truth should be spoken without regard to consequences, that justice should be done "though the sky should fall"' (200; also *EEM* 25), all of which are absolute rather than other things equal. In another he said a deontological view makes 'the practically ultimate end of moral actions their conformity to certain rules or dictates of Duty unconditionally prescribed', where 'unconditionally' again means without considering all their consequences (96; also 3). But our ultimate aim cannot be just to do acts that are right other things equal; it must be to do acts that are right all things considered. Here again he in effect assumed that the only alternative to consequentialism is an absolute deontology that forbids some acts whatever their consequences; the weaker moderate deontology was excluded.

There was a related error in his claim that in their initial formulations the deontological principles contradict each other. His official statement of his third condition said that any self-evident propositions

> must be mutually consistent. . . . any collision between two intuitions is proof that there is error in one or the other or in both. Still, we frequently find ethical writers treating this point very lightly. They appear to regard a conflict of ultimate rules as a difficulty that may be ignored or put aside for future solution . . . Whereas such a collision is absolute proof that at least one of

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

the formulae needs qualification: and suggests a doubt whether the correctly qualified proposition will present itself with the same self-evidence as the simpler but inadequate one. (341)

Here he took a conflict of duties to involve formal inconsistency and to demand the qualifications only such inconsistency requires. As Ross emphasized, however, conflicting prima facie principles do not contradict each other.[17] That in his example you ought other things equal to keep your promise and ought other things equal to help the accident victims are perfectly consistent claims; each correctly describes one aspect of your moral situation. It might be problematic to say you ought all things considered to keep the promise and all things considered to help the victims, but the principles do not say that. In his more detailed discussions of deontological duties Sidgwick repeatedly said their 'conflicts' and 'collisions' require them to be restated to include exception-clauses; as so stated, they are then neither intuitive nor widely accepted. But in so far as his demand for restatement rested on the claim that otherwise they contradict each other, it rested on an error.

This error does not completely vitiate Sidgwick's negative argument. It just means he could not say conflicts of deontological duty involve both inconsistency and unclarity; at most they involve unclarity. But there is another point where he ignored the concept of prima facie duty.

Anticipating his later argument that utilitarianism systematizes common-sense morality, he often argued that when common sense tries to remove its principles' unclarities it looks to the consequences different acts will have, in particular for overall happiness. This suggests, he argued, that the principles are not really freestanding but are derived from the principle of utility, which is their ultimate ground. But his arguments for this claim often moved directly from the premise that common sense takes some account of consequences to the conclusion that it evaluates only by consequences, where that does not at all follow.

At one point he called the duty to compensate those we have harmed a 'simple deduction' from the utilitarian principle of benevolence, which we 'approximately obey' by reversing any reduction in happiness we have caused (281). But if common sense really understood the duty to compensate that way, it would say that if you can either give a person you harmed a certain benefit or give an equal-sized benefit to someone else, it does not matter which you do. Common sense does not say that; it thinks the claim for compensation has priority, and does so even when you can give the stranger a somewhat larger benefit. If you can

[17] Sir David Ross, *Foundations of Ethics* (Oxford, 1939), pp. 88–9.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

give him a vastly greater benefit, say by saving his life, you ought to do so, but that means only that its view is moderate rather than absolute deontology; it is not consequentialism. He likewise argued that since common sense says it can be permissible to lie to a child or invalid for her own good, the only way to decide when lying is permissible is by 'considerations of expediency', or by weighing the total good and bad effects a lie will have (316). Again, however, common sense does not think lying even to an invalid is permitted for just a small increase in overall good; it must be a large one. And Sidgwick's line is especially implausible about the duty not to harm. In one passage he said that when we consider the 'negative duty of abstaining from causing pain to others against their will', which he was unsure whether to class under justice or benevolence, we find the only way to define its limits and determine when we may harm some for the sake of others is 'the Utilitarian formula' (348). But common sense does not think it is permissible to kill one innocent person in order to save just a few more; in the well-known example it does not think you may push one person in front of a trolley to stop it from running over five. It thinks the duty not to kill has substantial weight against consequentialist considerations and is therefore not derived from them. It may say that if the only way to save a million people is to kill one, doing so is permitted and even required. But that is again moderate deontology, not consequentialism. That Sidgwick ignored the moderate view was C. D. Broad's main objection to his critique: he ascribed to common sense and then criticized an 'extreme' form of deontology, where there is a more restrained form that both fits everyday thought and avoids his objections.[18]

This ignoring of moderate deontology was shared by other consequentialists of Sidgwick's day. Having said a deontological view pronounces acts right or wrong 'without reference to their consequences',[19] Rashdall illustrated it by stating its rule of veracity as 'Do not lie under any circumstances whatever',[20] and took 'objections to that sort of Intuitionism which declares that certain rules of action are to be followed irrespectively of consequences' to show that we are 'compelled to accept the utilitarian formula in so far as it asserts that conduct is good or bad only in proportion as it tends to promote the Well-being of human society'.[21] Moore likewise contrasted consequentialism only with 'the strictly Intuitionistic view that certain

[18] C. D. Broad, *Five Types of Ethical Theory* (London, 1930), pp. 217–23.
[19] Rashdall, *Theory of Good and Evil*, vol. 1, p. 80; also 91–2.
[20] Rashdall, *Theory of Good and Evil*, vol. 1, p. 92; also 193–4.
[21] Rashdall, *Theory of Good and Evil*, vol. 1, p. 91; also 83–4. Rashdall acknowledged the possibility of moderate deontology in his later *Ethics* (London, 1913), p. 46.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

ways of acting [are] right and others wrong, whatever their results might be',[22] thereby ignoring the moderate deontology intermediate between those two. We might not expect Sidgwick to do the same, given his greater attention to common-sense morality, but, albeit more subtly, he did.

If we set aside his charge that unqualified deontological principles contradict each other, Sidgwick's remaining objection was that they are unclear, first because they cannot be weighed precisely against each other. How serious a defect is it if a deontological theory cannot always make determinate judgements?

We need to distinguish here between metaphysical and epistemic indeterminacy. There is metaphysical indeterminacy if the moral truth itself is partly indeterminate, so there are cases where it is true neither that one duty outweighs another, nor that the second outweighs the first, nor that their weights are exactly equal. There is merely epistemic indeterminacy if the truth is always fully determinate but we sometimes cannot know it.

If a deontological view says there is metaphysical indeterminacy in the weights of some duties, it is hard to see how its own indeterminacy is any objection to it: it is simply reflecting the moral truth as it takes it to be. We might find the view's incompleteness objectionable if we thought of morality as a device to help us make decisions we cannot otherwise make, but deontologists do not see morality that way; they see it as a matter of objective moral truth, as Sidgwick also did. And if the truth is not completely precise, a theory that describes it accurately cannot be so either. It may be argued on general metaphysical grounds that, whatever its subject matter, the truth is never even partly imprecise; it is always fully determinate. But Sidgwick did not make this argument, and in any case it does not apply to a view that says there is only epistemic indeterminacy.

This was in fact Ross's view. He thought there always is a precise truth about the weight of competing duties, though it is often beyond our power to detect. This is why he called a right act a 'fortunate act': 'If we cannot be certain that it is right, it is our good fortune if the act we do is the right act.'[23] It is again hard to see what the objection to this view could be. There might be one if we knew in advance that we can apprehend all the moral truth there is, but Sidgwick did not make this claim, and it is not plausible. We do not know all the mathematical truths there are, and why think our access to the moral truth is any more complete? A deontological view like Ross's is not metaphysically

[22] Moore, *Principia Ethica*, p. 106; also 148, *Ethics*, pp. 74–7.

[23] Ross, *Right and the Good*, p. 31, also pp. 29–30, 34n., 142–4; *Foundations*, pp. 183, 189.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

problematic, because it thinks the moral truth is fully determinate, and it rests the incompleteness of its own claims on the familiar fact of human cognitive limitations. What is remotely objectionable about that?

It is actually puzzling that Sidgwick made his first charge of unclarity, since his own preferred moral theories have the very same feature. They say that what is right is what will result in the most pleasure possible, either for the agent or for all sentient beings, and that, as he often emphasized, is not something we can always know.

His canonical method for applying these theories was 'empirical hedonism' (123), which identifies the consequences different acts will have and then compares the amounts of pleasure and pain they contain. But even apart from the difficulty of knowing exactly what effects an act will have, he did not think we can compare hedonic states precisely. When he considered two pleasures of his own of the same type, he said, 'it is only to a very limited extent that I can obtain clear and definite results from such comparisons' (143), and the difficulties only increased when he considered pleasures of different types or of different people. There had been proposals from Herbert Spencer and others to measure pleasure 'scientifically', by using laws derived from psychology or evolutionary biology, but he thought they all fail (176–95, 470–3)[24] and hedonistic theories must use the empirical method with all its 'perplexity and uncertainty' (460). But then they cannot always say either that one of two acts has better consequences than the other or that their consequences are exactly equally good.

The puzzle is why he did not consider this any objection to these theories. He thought the difficulty of weighing deontological duties requires that they be restated to include exception-clauses, but the indeterminacy in hedonistic theories was just a practical inconvenience. Why this radical difference in response?

He might have justified the difference if deontological theories involve metaphysical indeterminacy and hedonistic theories only an epistemic one, but he did not make this claim, and it does not apply to a deontology like Ross's that takes the moral truth to be fully determinate. He could also have argued that the hedonistic indeterminacies concern empirical facts while the deontological ones concern moral ones and would remain even if we had full empirical knowledge. But he again did not make this claim, and would need to give an argument why it matters. Why should indeterminacy resulting from one cognitive limitation be a serious objection while

[24] See also his *Lectures on the Ethics of T. H. Green, Mr. H. Spencer, and J. Martineau*, ed. E. E. Constance Jones (London, 1902), pp. 146–77.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

that resulting from another is not? Not only was his first charge of unclarity unpersuasive in itself, it criticized deontological theories for a feature he did not find at all problematic in his own.[25]

Sidgwick's second objection was that deontological principles are unclear individually, or have vague boundaries. When this is so we again cannot judge acts decisively, but the reason is now internal to one or more principles rather than due to any conflict between them.

Though Ross's concept of prima facie duty effectively answered the charge about conflicts, Ross did much less to address this second one. On the contrary, he tended to state his prima facie duties in just the vague way Sidgwick found unsatisfactory, saying only that there is a duty to keep promises (but exactly when?), show gratitude (for exactly what?), or not harm others (which differs exactly how from not benefiting them?). This second vagueness is more problematic. It is one thing for a moral theory to be indeterminate where it says we cannot have self-evident knowledge, as Ross thought we cannot about conflicts of duty. It is another for it to be indeterminate where it says we can have such knowledge. If we have a self-evident intuition that a property is right-making, should the intuition not tell us exactly what the property is and when it is present? Sidgwick argued that if deontological duties 'can be referred to independent and self-evident principles, the limits of each must be implicitly given in the intuition that reveals the principle' (247). Where there is genuine intuition, in other words, there should at least after reflection be full determinacy, and with the proposed deontological intuitions there is not.

Ross did address an objection of this type about promises. His response to a series of problem cases raised by W. A. Pickard-Cambridge[26] was that the boundaries of a promise are set by conditions attached to it when it was made and accepted, if only tacitly, by both parties. Thus, if they both made it a condition of a promise that circumstances not change radically, and circumstances have changed, the duty to keep the promise is void.[27] But it is not clear that this clarification fully meets the objection. Some may agree that promises can have conditions but insist that they be explicitly rather than tacitly accepted. Others may allow tacit acceptance but disagree about what is required for it. Ross described a case where if *A* had stated a condition openly *B* would have accepted it, and said the condition therefore 'implicitly determines the nature of the understanding' between them.[28] But some may deny that

[25] See also Phillips, *Sidgwickian Ethics*, pp. 101–3.

[26] W. A. Pickard-Cambridge, 'Two Problems About Duty (I.), (II.), and (III.)', *Mind* 41 (1932), pp. 72–96, 145–72, 311–40.

[27] Ross, *Foundations*, pp. 94–8.

[28] Ross, *Foundations*, p. 94.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

this merely hypothetical fact suffices for *B*'s acceptance; he must do something actually to accept the condition. And further issues arise if the parties to a promise do not understand it in the same way. Is its content then determined only by those conditions both accepted, or is it enough if the promissor sincerely believes the promissee accepted a condition?[29] And of any fully determinate account of the duty to keep a promise we can ask Sidgwick's questions: does a duty specified in this precise way really seem self-evident, and is it really universally accepted? Similar issues arise about other deontological duties, such as the duty not to harm. As the voluminous discussion of trolley cases has shown, it is far from easy to say just what makes for the morally more objectionable case of harming as against the less objectionable one of not benefiting, and thus far from easy to say exactly when the duty applies. But should a self-evident intuition that harming is wrong not make it self-evident what harming is?

There might not be a difficulty here if we could always judge particular acts decisively. Then a deontologist could say the situation is like that with a concept and its analysis. In judging particular acts we rely implicitly on certain criteria but cannot always state them explicitly; doing so is the task of ethical theory. The problem is that we cannot always judge particular acts. When we consider difficult cases about promising we are often uncertain whether a past pledge really is still binding, and though we may be confident that it is permissible to throw the switch in the original trolley case, most of us are less certain about one involving a looping track and can disagree about it. Sidgwick's second problem was therefore not just about making implicit criteria explicit.

A deontologist could try accepting this second indeterminacy as he does the first, though that would involve abandoning Sidgwick's clarity condition and allowing individual principles to have vague contents. The resulting indeterminacy would then be metaphysical if there was no determinate truth about which precise properties are right-making and epistemic if there was a truth but we cannot always know it. On either view there could be core cases where we are certain a right-making property is present, for example when a promise has explicitly been understood in the same way and all its conditions are satisfied, and peripheral cases where we are not. But both possibilities are problematic if the truth in question is supposed to be self-evident. Mathematical truths have precise contents and are known by us, when they are known, as precise. If moral truths have a similar status, as

[29] Alan Donagan reports that Whewell took this latter view; see 'Sidgwick and Whewellian Intuitionism: Some Enigmas', *Canadian Journal of Philosophy* 7 (1977), pp. 447–65, at 457–8.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

Sidgwick and Ross both held, should they not be knowable in the same way?

Sidgwick's second objection is therefore a serious one, but the question is whether it applies only to deontological theories. He thought his own consequentialist principles satisfy the clarity condition, but they need to be supplemented by claims about what is good, and there can be indeterminacies in these, either about the comparative weights of different goods or about the boundaries of a particular one, for example, about exactly what counts as a virtue. These difficulties do not arise if, as Sidgwick thought, the only good is pleasure, but others remain. As he recognized (415–16), a consequentialist principle can tell us to maximize either the total good in a population or the average good per member, and it is not obvious which it should do.[30] May there then be core cases of consequentialism where the same act maximizes both the total and the average good and more contestable ones where the two diverge? There can be similar indeterminacies about individual pleasures. Sidgwick thought a pleasure that is twice as intense or lasts twice as long is always twice as good, but one can instead hold that the value of an increase in a pleasure's intensity or duration gets smaller the more intense or lasting the pleasure is. Moore suggested a view of this diminishing-value type, as did McTaggart and Broad;[31] the choice between it and Sidgwick's view can again generate consequentialist indeterminacies.

Sidgwick raised a serious challenge in his argument that if we can intuit a moral truth as self-evident we should be able to intuit it with a precise content. But it is unclear that the challenge applies only to deontological theories rather than more widely, to all claims to self-evident moral knowledge. If it applies more widely, it cannot support consequentialism over rival views.

THE POSITIVE CASE FOR CONSEQUENTIALISM

The positive half of Sidgwick's argument, presented in *Methods* III.13, was that the principles defining consequentialism *do* satisfy his conditions. It too had flaws.

[30] Ross denied that it is self-evident that we should prefer promoting the greatest total good to promoting the greatest average good (*Foundations*, pp. 69–71), and there are similar difficulties about the distribution of happiness in a population, e.g. is it better if it is equal, and if so, how much weight does equality have?

[31] Moore, 'Mr. McTaggart's Ethics', *International Journal of Ethics* 13 (1903), pp. 341–70, at 358; John McTaggart Ellis McTaggart, *The Nature of Existence*, 2 vols. (Cambridge, 1921, 1927), vol. 2, pp. 437–8; Broad, *Examination of McTaggart's Philosophy*, 2 vols. (Cambridge, 1933, 1938), vol. 2, p. 684.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

The principles in question were his axioms of justice, prudence and rational benevolence. The first says: 'if a kind of conduct that is right (or wrong) for me is not right (or wrong) for some one else, it must be on the ground of some difference between the two cases, other than the fact that I and he are different persons' (379); it expresses the idea of universalizability or supervenience. The second initially says 'one ought to aim at one's own good', but Sidgwick worried that, if 'good' can be analysed as 'what one ought to aim at', as he himself believed (112–13), this claim is tautological. He therefore supplemented it with a demand for 'impartial concern for all parts of our conscious life', saying 'the mere difference of priority and posteriority in time is not a reasonable ground for having more regard to the consciousness of one moment than to that of another' (381). He derived his third axiom, of benevolence, in stages. First he said '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'. Adding the further claim that 'as a rational being I am bound to aim at good generally . . . not merely at a particular part of it' yielded the axiom in its full form: 'each one is morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him' (382). Combining this axiom's demand for impartiality across persons with the temporal impartiality required by the axiom of prudence yields a fully impartial consequentialist principle.

It is unclear whether Sidgwick thought his first, or universalizability, axiom is substantive or follows analytically from the concept 'right'. His calling it the axiom of 'justice' suggests the substantive reading; his saying we can arrive at it 'by merely reflecting on the general notion of rightness' (208; also 34, *ME1* 364) suggests the analytic one. But the axiom is not distinctively consequentialist; it is accepted by most deontologists and cannot support consequentialism against them.

Some commentators take the second axiom to make no positive claim but only the negative one that you should not care more about goods at one time in your life than at another, which is compatible with not caring about your good at all.[32] But though Sidgwick did say the negative claim is '[a]ll that the principle affirms' (381), this is not on balance the most persuasive reading of his view. A merely negative axiom would not deserve the name 'prudence', which he elsewhere associated with positively seeking your good (7, 25–6, 36, 327). The negative claim would also then not be just an 'addition' to the axiom

[32] Phillips, *Sidgwickian Ethics*, p. 96; Robert Shaver, *Rational Egoism* (Cambridge, 1998), pp. 74–7, 'Sidgwick's Axioms and Consequentialism', *Philosophical Review* 123 (2014), pp. 173–204, at 179.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

needed to remove the appearance of tautology; it would *be* the axiom. And Sidgwick gave the axiom positive content when he restated it in III.14 as a 'precept to seek . . . one's good on the whole' (391; also 119–20) and when he considered a challenge to it in IV.2 (418–19). Most importantly, a negative principle would not have the connection he thought his axiom has with egoism. In earlier editions of *The Methods* he said: 'The axiom of prudence, as I have given it, is the self-evident principle on which, according to me, rational egoism is based' (*ME3* 388, *ME4* 387), but that can only be so if the axiom tells us positively to pursue our good. In the final edition he said only that the axiom is 'implied in Rational Egoism' (386), but a merely negative axiom is no more implied in egoism than it is in utilitarianism. And, crucially, only a positive axiom can generate his 'dualism of the practical reason', the conflict between egoism and utilitarianism that he discussed at the end of his book, could not resolve, and thought threatens the coherence of ethical thought. There can only be a 'contradiction in our apparent intuitions of what is Reasonable in conduct' (508) if one apparently self-evident principle tells us to maximize just our own good and another tells us to maximize the good of all, and that is certainly how he described the dualism elsewhere (*EEM* 43). His self-evident principles were stated in III.13, and there the one that can ground egoism is the axiom of prudence; moreover, that axiom is precisely what he said must be harmonized with the axiom of benevolence if the dualism is to be avoided (498).

There has also been dispute about the axiom of benevolence. It cannot plausibly be read as just negative, but Robert Shaver has argued that it describes only what is true 'from the point of view of the Universe', so an egoist can accept it as true yet not be committed to act as it commands so long as he refuses to adopt that point of view.[33] This reading has the merit of making Sidgwick's axiom more likely to satisfy his fourth condition about consensus, but it does not fit many crucial aspects of his text. It attaches great weight to a 'point of view of the universe' phrase he used only twice in the body of *The Methods* (382, 420), introduced tentatively ('if I may say so') and elsewhere analysed reductively, so the good from the point of view of the universe is just what everyone ought morally to desire and pursue (*EEM* 27). The phrase occurs only in one of the premises from which he derived his axiom and neither in the axiom itself nor in his later restatements of it as, for example, an unrelativized 'precept to seek . . . others' good no less than one's own' (391–2; also 496, 500, 507, *EEM* 25, 43, 44). And whereas Shaver's reading implies that an egoist can accept the axiom of benevolence,

[33] Shaver, 'Sidgwick's Axioms', pp. 174–84.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

Sidgwick explicitly denied that, saying a consistent egoist must make no claims about the universal good, since if he makes them he can be argued into impartialism (420–1, 497–8). Shaver's reading also makes the axiom massively fail the clarity condition; whereas Sidgwick criticized deontological principles for not always giving determinate moral guidance, a relativized axiom by itself gives no guidance. For that it must be supplemented by 'You ought to adopt the point of view of the universe', a claim Sidgwick, tellingly, never discussed. A relativized axiom also does not by itself yield consequentialism, as Sidgwick said it does (387–8); that too requires the 'ought' claim he never discussed. And, again crucially, a stronger axiom is needed to generate the dualism of the practical reason. A principle merely describing what is true from an optional point of view does not and cannot contradict egoism; only an unrelativized principle telling each person to maximize the good of all does. For a multitude of reasons, he must have intended a principle of that kind.

In the first edition of *The Methods* Sidgwick took a claim about your own good to be primary and derived the axiom of benevolence from it as a kind of corollary (*ME1* 359–60), but later editions treated the two axioms as coordinate and of equal self-evidence. More specifically, they took both to rest on a more abstract thesis 'obtained by considering the similarity of the individuals that make up a . . . Mathematical or Quantitative Whole' and saying it is irrational to care more about one part of a whole, say, the pleasure in one temporal part of your life, than about another (380–3).[34] This thesis, however, cannot give independent support to the axioms. Imagine someone who argues that the total pleasure in your life is only part of a whole comprising all your hedonic states, both pleasures and pains, so it is irrational to promote only your pleasures; you ought to promote your greatest total of both pleasures and pains. Sidgwick would object that a whole comprising both pleasures and pains is not a relevant whole for moral purposes, but then someone who rejects the demand for temporal impartiality can likewise deny that the whole comprising all the times in his life is relevant, and an egoist can do the same about a whole including all people. To apply the thesis about wholes we must first know what a relevant 'similarity' is, which is precisely what the axioms address.

Though he had rigorously tested deontological principles, Sidgwick did not subject his own axioms to anything like the same scrutiny. On the contrary, his treatment of them was extremely lax.

[34] The role of this thesis in Sidgwick's arguments is emphasized in Terence Irwin, *The Development of Ethics: A Historical and Critical Study*, vol. 3 (Oxford, 2009), pp. 496–505.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

For one thing, he did not consider possible counterexamples to them. The axiom of prudence requires full temporal neutrality, but Derek Parfit has objected that it is not unreasonable to care less about pleasures and pains in your past than about ones in your present or future. If while suffering temporary amnesia you are told that you either had a long and very painful operation yesterday or will have a shorter, less painful one today, you will, perfectly reasonably, hope you had the longer operation yesterday.[35] The same point was made earlier by McTaggart. He asked us to consider two people, each with a perfect memory of his last ten years but no memory of earlier ones, and a perfect forecast of his next ten years but none beyond that. If the first person has had ten years of agony and will now have ten of pleasure while the second has had ten years of pleasure and will now have agony, will the first not feel much happier today?[36]

This is a philosophically subtle objection, but there are more straightforward ones to the impartiality of the axiom of benevolence. Sidgwick acknowledged one when he said, right after stating the axiom, 'the duty of Benevolence as recognised by common sense seems to fall somewhat short of this', because it does not require you to sacrifice your good for just slightly greater benefits to others (382; also 252–3, 499), i.e. is less demanding. His response was that if 'a "plain man," in a modern civilised society', were 'fairly brought to consider . . . whether it would be morally right for him to seek his own happiness on any occasion if it involved a certain sacrifice of the greater happiness of some other human being', he 'would answer unhesitatingly in the negative'. But he gave no evidence for this claim about the plain man – he had not done an opinion survey – and it is highly dubious; surely many would reject his axiom as too strict.[37] Nor did he consider the even more telling objection from non-maleficence: that the axiom is wrong to make it permissible to kill one innocent person if that will produce slightly greater benefits for others. Here the claim that a plain man would 'unhesitatingly' take the consequentialist line is simply not credible.

In addition, he did very little to substantiate his claim that his axioms satisfy the four conditions on self-evidence he had used against deontological principles. He said the axioms 'do present themselves as self-evident' as intuitionistic principles do not (383), thereby applying the second condition, but he gave no evidence for this claim, as, to be fair, it would be hard for him to do. He seems simply to have assumed

[35] Derek Parfit, *Reasons and Persons* (Oxford, 1984), pp. 165–7.

[36] McTaggart, *Nature of Existence*, vol. 2, pp. 348–9. Similar claims were made in E. F. Carritt, *The Theory of Morals* (London, 1928), p. 26; Ewing, *Value and Reality*, p. 281.

[37] Parfit calls Sidgwick's claim about the plain man 'simply false' (*On What Matters*, vol. 1, p. 453).

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

that they satisfy the clarity condition, not considering difficulties like the ones about the measurement of goods or total versus average good mentioned above, and he did not discuss the question of their consistency. This is presumably because he thought his axioms of prudence and benevolence contradict each other, and wanted to leave that topic until he addressed the dualism in his final chapter. But if he was saying his axioms satisfy all the conditions whereas deontological principles do not, should he not have acknowledged this conflict earlier? And his argument that there is consensus on his axioms was extremely thin. His claim about the common-sense duty of benevolence was in part an attempt to show that his third axiom is widely accepted, but, as I have said, it is not persuasive. And his more explicit claims about consensus mentioned the opinions only of philosophers, and then only of two, Clarke and Kant, as if their assent could stand in for all moral theorists' and then for all people's generally (384–6). How is an appeal to two people's beliefs a serious test of universal agreement?[38] He did say he was considering the opinions of those moralists who have been 'most in earnest' in seeking 'genuine intuitions', but his initial description of dogmatic intuitionism said its proponents seek just such intuitions (101), and it can hardly suffice for denying that someone like Whewell was 'in earnest' that he did not share Sidgwick's conclusions.

Sidgwick did not, therefore, make a serious case that his axioms satisfy all his conditions. And what he did say reflects a fundamental unfairness in his argument, one that again turns on his lacking the concept of prima facie duty.

Though he did not put it this way, his argument about conflicts of duty rested at bottom on the charge that deontologists equivocate between other-things-equal and all-things-considered claims. The unqualified principle 'you ought to keep your promises' makes acts falling under it right only other things equal, since some other principle can make them on balance wrong; it therefore does not yield decisive verdicts. Clear verdicts do follow from a principle that contains exception clauses and so makes acts falling under it all things considered right, but that principle neither seems self-evident nor is accepted by all.

But Sidgwick's statements of his axioms equivocate just as much between other-things-equal and all-things-considered claims. This is evident, first, in his axiom of prudence, which he needed to play two roles. On the one hand, it had to supply part of the framework of impartial consequentialism, its temporal neutrality, and for that purpose had to have an other-things-equal form, saying you ought to

[38] The first edition of *The Methods* introduced its axioms through a discussion of Kant and Clarke (*ME1* 357–64), and though in later editions Sidgwick gave a 'more direct' statement of his views (*ME* ix), the material on these philosophers remained.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

promote your own greatest good through time only when that will not prevent you from promoting more good for others. Then it could be extended to require temporal impartiality about others' lives when it was combined with the axiom of benevolence. But it also had to be the basis for rational egoism or the self-evident element within it, and there it had to make an all-things-considered claim, saying that, no matter what else will result, you ought to promote your own good. That is certainly how he interpreted its central claim when he attributed it to Butler and Clarke (119–20), restated it in an article of 1889 (*EEM* 43) and discussed it in connection with the dualism (498).

The axiom of benevolence did not play two roles, and therefore was needed only in its all-things-considered form. But this cannot be how Sidgwick understood it when he said it is the object of a consensus including Kant. For Kant clearly did not think you ought always to do what will result in the most good for all people. As a deontologist – in fact, the best-known one – he accepted the principle of benevolence (minus its claim about your own pleasure) only in an other-things-equal version, so it can be and often is outweighed by other principles; thus he did not think benevolence permits you to lie or kill. Whereas Sidgwick's defence of consequentialism required the axiom in a strong form, Kant accepted it only in a much weaker one.[39]

In fact Sidgwick's axioms are open to the same general objection he made to deontological principles. Only when they are stated as other things equal is it at all plausible to say they appear self-evident and are generally accepted, though even then there can be objections to them like those mentioned above. But in that form they cannot ground either egoism or utilitarianism, as Sidgwick needed them to do. To do that they must make all-things-considered claims, but then they are less intuitive and will be widely rejected: many will deny that you should always do only what is best for you, and many will also deny that you should always do what has the best consequences for all. As he said of deontological principles, Sidgwick's axioms may satisfy some conditions for self-evidence in one form and others in a different form, but in no form do they satisfy them all.

It may be said in his defence that he did not need all-things-considered axioms. He could accept only other-things-equal ones and arrive at consequentialism by combining them with the negative claim that there are no other self-evident principles, which claim his earlier critique of deontology supplied. By itself each axiom would then have a weaker form, but adding the negative claim would make it *de facto*

[39] This point about Kant is also made in Irwin, *The Development of Ethics*, vol. 3, p. 518. If Kant lacked the concept of prima facie duty, his principle of benevolence would have included exception clauses but would still not have been a consequentialist one.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

stronger, by making it *de facto* all things considered.[40] Like Shaver's interpretation of the axiom of benevolence, this one gives the axioms a better chance of satisfying the consensus condition; it also fits Sidgwick's later claim that he had only shown the consequentialist principle to be one moral axiom but not that it is sole or supreme (421). But it has other interpretive flaws.

First, it makes Sidgwick's axioms fail his clarity condition in the very way for which he criticized deontological principles, since as other things equal they do not by themselves yield decisive verdicts. It also means consequentialism does not have a fully self-evident basis: its ground now includes a negative claim that is not intuited but depends on a complex critical argument carried on through eleven chapters of *Methods* III. Though Sidgwick claimed to be certain of consequentialism, that is not possible if one of its grounds is a discursive argument that cannot as such warrant certainty. Nor on this reading can the dualism of the practical reason involve, as Sidgwick said it does, a contradiction between apparently self-evident judgements, since as other things equal the axioms do not contradict each other but are perfectly consistent. In fact, on this reading there cannot even be a dualism. That would require both a demonstration that there is no self-evident axiom other than the axiom of prudence – so it can be affirmed not only as other things equal but also as all things considered – and a demonstration that there is no self-evident axiom other than the axiom of benevolence – so it can be affirmed in the same way. Since there cannot be both those demonstrations, there cannot be both an all-things-considered egoistic principle and an all-things-considered utilitarian one. And whereas this interpretation assumes that Sidgwick understood the concept of other-things-equal duty, we have seen that he did not, since he repeatedly failed to use it when it was relevant. How then could he have it in mind it when stating his axioms? I believe the only conclusion is that Sidgwick equivocated. Just as he tended to assume that any deontological principles must be absolute or all things considered, so he often read his axioms as strong enough to ground fully determinate views such as egoism and utilitarianism. But he also wanted them to meet his four conditions, and when he thought of them that way he slid into imagining them as other things equal. He understood them in one way when drawing conclusions from them and in another when saying they are intuitive, as he could not have done had he clearly grasped Ross's concept.

[40] Anthony Skelton, 'Sidgwick's Philosophical Intuitions', *Etica & Politica / Ethics and Politics* 10 (2008), pp. 185–209, at 203–4; Shaver, 'Sidgwick's Axioms', pp. 185–94, 200.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

CONCLUSION

The above criticisms are directed at Sidgwick's specific arguments about deontology and consequentialism, but they also tell against some more general claims about him.

He is often praised as, if not a lively writer, then an admirably clear one, his prose free of jargon and obscurities and his meaning easy to detect. This is indeed often true, especially when he was making smaller points about, for example, specific aspects of common-sense morality. But a general truth about *The Methods of Ethics* is that the more important a topic is, the less time Sidgwick spent on it, and what he said about it was then often far from clear. Why, for example, did he think the existence of a God who rewards virtue in an afterlife would resolve the dualism of the practical reason? Moore and Broad found this baffling: even if egoism and impartial consequentialism would then agree about which acts are right, they would give competing explanations of why they are right.[41] Sidgwick did not address this issue in the few pages he gave his book's main conclusion at its close. The same is true of his ethical axioms, his statements of which were, as we have seen, multiply ambiguous. Does the axiom of prudence make only a negative claim forbidding time-preference or also a positive one telling us to promote our good? Is the axiom of benevolence relativized to a universal point of view or unrelativized? And how strong a claim does each axiom make? Is it just other things equal or all things considered? Some aspects of his overall argument require the one and some the other, but he never decided clearly between them. Despite the axioms' centrality to his view, he gave them brief equivocal statements and left them at that.

Sidgwick is also often called a paradigmatically fair-minded philosopher. McTaggart spoke of his 'scrupulous fairness to opponents'; Broad called *The Methods* a 'conspicuously honest' book;[42] and even his critics have agreed, with Alan Donagan calling him 'a man of heroic disinterestedness' and Bernard Williams noting the 'marked scrupulousness of his arguments'.[43] But while this trait is often evident in his discussions of smaller topics, it can again be missing in his treatment of larger ones. He took it to be a serious objection to a pluralist deontology that it cannot always judge acts determinately because it cannot weigh duties precisely, but no objection to hedonistic

[41] Moore, *Principia Ethica*, p. 103; Broad, *Five Types*, pp. 159–60, 253.

[42] McTaggart, 'The Ethics of Henry Sidgwick', *Quarterly Review* 205 (1906), pp. 398–409, at 412; Broad, 'John McTaggart Ellis McTaggart', *Proceedings of the British Academy* 13 (1927), pp. 307–34, at 309.

[43] Donagan, 'Sidgwick and Whewellian Intuitionism', p. 456; Bernard Williams, 'The Point of View of the Universe: Sidgwick and the Ambitions of Ethics', *Making Sense of Humanity* (Cambridge, 1995), pp. 153–71, 154.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press

theories that they cannot compare pleasures precisely. He tested deontological principles rigorously with his four conditions but barely applied those conditions to his own principles. And though one of his main objections to deontological principles turned on the difference between their other-things-equal and all-things-considered forms, his defence of his axioms equivocated on the same point and involved the same ambiguity. In arguing against deontology and for consequentialism he applied, and not just once, a double standard.[44]

tom.hurka@utoronto.ca

[44] Material in this article is extracted from chapters 5 and 7 of my book *British Moral Philosophers from Sidgwick to Ewing* (Oxford, forthcoming). For helpful discussion I am indebted to Roger Crisp, Brad Hooker, Robert Johnson, Derek Parfit, David Phillips, Rob Shaver, Wayne Sumner and Peter Vallentyne.

https://doi.org/10.1017/S0953820814000089 Published online by Cambridge University Press