

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.

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

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

46 SIDGWICK'S CONDITIONS FOR SELF-EVIDENCE

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

55 The first condition says 'the terms of the proposition must be clear
 56 and precise'; it is commonly called the clarity condition. The second
 57 says 'the self-evidence of the proposition must be ascertained by careful
 58 reflection', while the third says 'the propositions accepted as self-
 59 evident must be mutually consistent' since 'any collision between two
 60 intuitions is a proof that there is error in one or the other, or in both'.
 61 The fourth condition requires general agreement on the proposition,
 62 or a consensus about it: 'since it is implied in the very notion of Truth
 63 that it is essentially the same for all minds, the denial by another of a
 64 proposition that I have affirmed has a tendency to impair my confidence
 65 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.

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

70 I think Sidgwick was a little careless in stating his second condition.
 71 He often contrasted real with merely apparent self-evidence (34 n. 2,
 72 211, 338, 383, 508; also *EEM* 30, 31–2, 33–4, 170), and his condition
 73 is best read as asking us to establish the *apparent* self-evidence of a
 74 proposition, or establish that it really is apparently self-evident. The
 75 picture is this.⁹

76 You start by being in what seems to you to be a mental state of
 77 apprehending a proposition as true just by understanding it. Your
 78 actually apprehending a proposition in this way does not guarantee
 79 that it is true, any more than your seeing a line as bent guarantees
 80 that it is bent. But you need to confirm that you really are in that
 81 mental state, and that is what the second condition requires. Sidgwick’s
 82 justification for it was precisely that we are often mistaken about
 83 our mental states; we tend to confuse genuine intuitions with ‘mere
 84 impressions or impulses which to careful observation do not present
 85 themselves as claiming to be dictates of Reason’ or with ‘mere opinions’
 86 that repetition has given ‘a false appearance of self-evidence’. Since
 87 ‘any strong sentiment, however subjective, is apt to transform itself
 88 into the semblance of an intuition’ (339; also 211–12, 340–1, *EEM*
 89 25, 31), we need to detect any such semblance. This requires ‘careful
 90 reflection’ (339), which for British philosophers of Sidgwick’s day meant
 91 introspection or self-examination more generally.¹⁰ Thus whether you
 92 really are having an intuition ‘can only be decided by each person by
 93 direct introspection or reflection’ (211; also 212, 362, 383, *EEM* 43), or
 94 by that plus a survey of possible distorting causes of your belief, such as
 95 common acceptance in your community or a strong desire that the belief
 96 be true (339). The second condition therefore requires a self-survey.
 97 You are in what seems to you to be a mental state of apprehending
 98 a proposition as true just by understanding it, or of apparent self-
 99 evidence. You then apply introspective and other tests to determine
 100 whether it really is a mental state of that kind. If it is, it involves a
 101 real case of apparent self-evidence, and you can then apply the other
 102 conditions to see if the proposition is simply, and not just apparently,
 103 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.

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

120 Sidgwick's official statement of his third condition requires an
121 apparently self-evident proposition to be consistent only with other
122 apparently self-evident propositions (341; also *EEM* 166), and with this
123 restriction the condition is compelling. If two equally well-grounded
124 beliefs contradict each other, you have no basis for preferring one to
125 the other and should abandon both. But sometimes he spoke as if
126 a self-evident belief must harmonize with all your beliefs about its
127 subject matter, whether apparently self-evident or not (400; also *EEM*
128 25, 31), and this stronger condition is less plausible. If an apparently
129 self-evident belief clashes with one that does not meet that condition,
130 should the latter not yield?

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

140 Unfortunately this reasoning conflicts with his own description of
141 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.

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

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

162 THE CRITIQUE OF DEONTOLOGY

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

171 He sometimes denied that deontological principles satisfy his second
 172 condition: they do not even ‘present themselves as self-evident’, and any
 173 appearance that they do is dissolved by ‘reflection’ (383; *EEM* 25). But
 174 he usually gave a more complex argument. If deontological principles
 175 are left as ‘somewhat vague generalities’, such as ‘you ought to keep
 176 your promises’, they do not satisfy the clarity condition, but they do
 177 seem self-evident and assent to them is ‘approximately universal’. Once
 178 we try to give them ‘the definiteness which science requires’, however,
 179 by adding qualifications that will make their application determinate
 180 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.

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

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

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

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

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

234 Rashdall thought the arguments entirely successful:

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

239 F. H. Hayward agreed:

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

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

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

262 That Sidgwick lacked Ross's concept is shown, first, in his general
263 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.

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

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

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

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

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

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

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

340 At one point he called the duty to compensate those we have harmed a
341 'simple deduction' from the utilitarian principle of benevolence, which
342 we 'approximately obey' by reversing any reduction in happiness we
343 have caused (281). But if common sense really understood the duty to
344 compensate that way, it would say that if you can either give a person
345 you harmed a certain benefit or give an equal-sized benefit to someone
346 else, it does not matter which you do. Common sense does not say that;
347 it thinks the claim for compensation has priority, and does so even
348 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.

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

376 This ignoring of moderate deontology was shared by other
 377 consequentialists of Sidgwick’s day. Having said a deontological
 378 view pronounces acts right or wrong ‘without reference to their
 379 consequences’,¹⁹ Rashdall illustrated it by stating its rule of veracity as
 380 ‘Do not lie under any circumstances whatever’,²⁰ and took ‘objections
 381 to that sort of Intuitionism which declares that certain rules of
 382 action are to be followed irrespectively of consequences’ to show that
 383 we are ‘compelled to accept the utilitarian formula in so far as it
 384 asserts that conduct is good or bad only in proportion as it tends to
 385 promote the Well-being of human society’.²¹ Moore likewise contrasted
 386 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.

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

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

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

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

417 This was in fact Ross's view. He thought there always is a precise
418 truth about the weight of competing duties, though it is often beyond
419 our power to detect. This is why he called a right act a 'fortunate act':
420 'If we cannot be certain that it is right, it is our good fortune if the act
421 we do is the right act.'²³ It is again hard to see what the objection to
422 this view could be. There might be one if we knew in advance that we
423 can apprehend all the moral truth there is, but Sidgwick did not make
424 this claim, and it is not plausible. We do not know all the mathematical
425 truths there are, and why think our access to the moral truth is any
426 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.

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

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

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

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

457 He might have justified the difference if deontological theories
458 involve metaphysical indeterminacy and hedonistic theories only an
459 epistemic one, but he did not make this claim, and it does not
460 apply to a deontology like Ross’s that takes the moral truth to be
461 fully determinate. He could also have argued that the hedonistic
462 indeterminacies concern empirical facts while the deontological ones
463 concern moral ones and would remain even if we had full empirical
464 knowledge. But he again did not make this claim, and would need
465 to give an argument why it matters. Why should indeterminacy
466 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.

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

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

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

492 Ross did address an objection of this type about promises. His re-
493 sponse to a series of problem cases raised by W. A. Pickard-Cambridge²⁶
494 was that the boundaries of a promise are set by conditions attached to
495 it when it was made and accepted, if only tacitly, by both parties. Thus,
496 if they both made it a condition of a promise that circumstances not
497 change radically, and circumstances have changed, the duty to keep the
498 promise is void.²⁷ But it is not clear that this clarification fully meets
499 the objection. Some may agree that promises can have conditions but
500 insist that they be explicitly rather than tacitly accepted. Others may
501 allow tacit acceptance but disagree about what is required for it. Ross
502 described a case where if *A* had stated a condition openly *B* would have
503 accepted it, and said the condition therefore 'implicitly determines the
504 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.

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

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

532 A deontologist could try accepting this second indeterminacy as he
533 does the first, though that would involve abandoning Sidgwick's clarity
534 condition and allowing individual principles to have vague contents.
535 The resulting indeterminacy would then be metaphysical if there was
536 no determinate truth about which precise properties are right-making
537 and epistemic if there was a truth but we cannot always know it. On
538 either view there could be core cases where we are certain a right-
539 making property is present, for example when a promise has explicitly
540 been understood in the same way and all its conditions are satisfied,
541 and peripheral cases where we are not. But both possibilities are
542 problematic if the truth in question is supposed to be self-evident.
543 Mathematical truths have precise contents and are known by us, when
544 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.

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

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

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

574 THE POSITIVE CASE FOR CONSEQUENTIALISM

575 The positive half of Sidgwick's argument, presented in *Methods* III.13,
576 was that the principles defining consequentialism *do* satisfy his
577 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, 1933–8), vol. 2, p. 684.

578 The principles in question were his axioms of justice, prudence and
579 rational benevolence. The first says: 'if a kind of conduct that is right
580 (or wrong) for me is not right (or wrong) for some one else, it must be
581 on the ground of some difference between the two cases, other than
582 the fact that I and he are different persons' (379); it expresses the idea
583 of universalizability or supervenience. The second initially says 'one
584 ought to aim at one's own good', but Sidgwick worried that, if 'good'
585 can be analysed as 'what one ought to aim at', as he himself believed
586 (112–13), this claim is tautological. He therefore supplemented it with
587 a demand for 'impartial concern for all parts of our conscious life',
588 saying 'the mere difference of priority and posteriority in time is not a
589 reasonable ground for having more regard to the consciousness of one
590 moment than to that of another' (381). He derived his third axiom, of
591 benevolence, in stages. First he said 'the good of any one individual
592 is of no more importance, from the point of view (if I may say so) of
593 the Universe, than the good of any other'. Adding the further claim
594 that 'as a rational being I am bound to aim at good generally . . . not
595 merely at a particular part of it' yielded the axiom in its full form:
596 'each one is morally bound to regard the good of any other individual
597 as much as his own, except in so far as he judges it to be less, when
598 impartially viewed, or less certainly knowable or attainable by him'
599 (382). Combining this axiom's demand for impartiality across persons
600 with the temporal impartiality required by the axiom of prudence yields
601 a fully impartial consequentialist principle.

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

609 Some commentators take the second axiom to make no positive claim
610 but only the negative one that you should not care more about goods
611 at one time in your life than at another, which is compatible with
612 not caring about your good at all.³² But though Sidgwick did say the
613 negative claim is '[a]ll that the principle affirms' (381), this is not on
614 balance the most persuasive reading of his view. A merely negative
615 axiom would not deserve the name 'prudence', which he elsewhere
616 associated with positively seeking your good (7, 25–6, 36, 327). The
617 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.

618 needed to remove the appearance of tautology; it would *be* the axiom.
619 And Sidgwick gave the axiom positive content when he restated it in
620 III.14 as a 'precept to seek . . . one's good on the whole' (391; also 119–
621 20) and when he considered a challenge to it in IV.2 (418–19). Most
622 importantly, a negative principle would not have the connection he
623 thought his axiom has with egoism. In earlier editions of *The Methods*
624 he said: 'The axiom of prudence, as I have given it, is the self-evident
625 principle on which, according to me, rational egoism is based' (*ME3*
626 388, *ME4* 387), but that can only be so if the axiom tells us positively
627 to pursue our good. In the final edition he said only that the axiom
628 is 'implied in Rational Egoism' (386), but a merely negative axiom is
629 no more implied in egoism than it is in utilitarianism. And, crucially,
630 only a positive axiom can generate his 'dualism of the practical reason',
631 the conflict between egoism and utilitarianism that he discussed at the
632 end of his book, could not resolve, and thought threatens the coherence
633 of ethical thought. There can only be a 'contradiction in our apparent
634 intuitions of what is Reasonable in conduct' (508) if one apparently self-
635 evident principle tells us to maximize just our own good and another
636 tells us to maximize the good of all, and that is certainly how he
637 described the dualism elsewhere (*EEM* 43). His self-evident principles
638 were stated in III.13, and there the one that can ground egoism is the
639 axiom of prudence; moreover, that axiom is precisely what he said must
640 be harmonized with the axiom of benevolence if the dualism is to be
641 avoided (498).

642 There has also been dispute about the axiom of benevolence. It cannot
643 plausibly be read as just negative, but Robert Shaver has argued that it
644 describes only what is true 'from the point of view of the Universe', so an
645 egoist can accept it as true yet not be committed to act as it commands
646 so long as he refuses to adopt that point of view.³³ This reading has
647 the merit of making Sidgwick's axiom more likely to satisfy his fourth
648 condition about consensus, but it does not fit many crucial aspects of his
649 text. It attaches great weight to a 'point of view of the universe' phrase
650 he used only twice in the body of *The Methods* (382, 420), introduced
651 tentatively ('if I may say so') and elsewhere analysed reductively, so
652 the good from the point of view of the universe is just what everyone
653 ought morally to desire and pursue (*EEM* 27). The phrase occurs only
654 in one of the premises from which he derived his axiom and neither in
655 the axiom itself nor in his later restatements of it as, for example, an
656 unrelativized 'precept to seek . . . others' good no less than one's own'
657 (391–2; also 496, 500, 507, *EEM* 25, 43, 44). And whereas Shaver's
658 reading implies that an egoist can accept the axiom of benevolence,

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

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

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

695 Though he had rigorously tested deontological principles, Sidgwick
 696 did not subject his own axioms to anything like the same scrutiny. On
 697 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.

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

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

730 In addition, he did very little to substantiate his claim that his
 731 axioms satisfy the four conditions on self-evidence he had used against
 732 deontological principles. He said the axioms ‘do present themselves as
 733 self-evident’ as intuitionistic principles do not (383), thereby applying
 734 the second condition, but he gave no evidence for this claim, as, to be
 735 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).

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

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

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

770 But Sidgwick's statements of his axioms equivocate just as much
 771 between other-things-equal and all-things-considered claims. This is
 772 evident, first, in his axiom of prudence, which he needed to play
 773 two roles. On the one hand, it had to supply part of the framework
 774 of impartial consequentialism, its temporal neutrality, and for that
 775 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.

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

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

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

810 It may be said in his defence that he did not need all-things-
811 considered axioms. He could accept only other-things-equal ones and
812 arrive at consequentialism by combining them with the negative claim
813 that there are no other self-evident principles, which claim his earlier
814 critique of deontology supplied. By itself each axiom would then have
815 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.

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

822 First, it makes Sidgwick's axioms fail his clarity condition in the
 823 very way for which he criticized deontological principles, since as
 824 other things equal they do not by themselves yield decisive verdicts.
 825 It also means consequentialism does not have a fully self-evident basis:
 826 its ground now includes a negative claim that is not intuited but
 827 depends on a complex critical argument carried on through eleven
 828 chapters of *Methods* III. Though Sidgwick claimed to be certain of
 829 consequentialism, that is not possible if one of its grounds is a discursive
 830 argument that cannot as such warrant certainty. Nor on this reading
 831 can the dualism of the practical reason involve, as Sidgwick said it does,
 832 a contradiction between apparently self-evident judgements, since as
 833 other things equal the axioms do not contradict each other but are
 834 perfectly consistent. In fact, on this reading there cannot even be a
 835 dualism. That would require both a demonstration that there is no
 836 self-evident axiom other than the axiom of prudence – so it can be
 837 affirmed not only as other things equal but also as all things considered
 838 – and a demonstration that there is no self-evident axiom other than
 839 the axiom of benevolence – so it can be affirmed in the same way.
 840 Since there cannot be both those demonstrations, there cannot be both
 841 an all-things-considered egoistic principle and an all-things-considered
 842 utilitarian one. And whereas this interpretation assumes that Sidgwick
 843 understood the concept of other-things-equal duty, we have seen that
 844 he did not, since he repeatedly failed to use it when it was relevant.
 845 How then could he have it in mind it when stating his axioms? I believe
 846 the only conclusion is that Sidgwick equivocated. Just as he tended
 847 to assume that any deontological principles must be absolute or all
 848 things considered, so he often read his axioms as strong enough to
 849 ground fully determinate views such as egoism and utilitarianism. But
 850 he also wanted them to meet his four conditions, and when he thought
 851 of them that way he slid into imagining them as other things equal. He
 852 understood them in one way when drawing conclusions from them and
 853 in another when saying they are intuitive, as he could not have done
 854 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.

855

CONCLUSION

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

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

881 Sidgwick is also often called a paradigmatically fair-minded
 882 philosopher. McTaggart spoke of his 'scrupulous fairness to opponents';
 883 Broad called *The Methods* a 'conspicuously honest' book;⁴² and even
 884 his critics have agreed, with Alan Donagan calling him 'a man of
 885 heroic disinterestedness' and Bernard Williams noting the 'marked
 886 scrupulousness of his arguments'.⁴³ But while this trait is often evident
 887 in his discussions of smaller topics, it can again be missing in his
 888 treatment of larger ones. He took it to be a serious objection to a
 889 pluralist deontology that it cannot always judge acts determinately
 890 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.

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

899 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.