

Can Sidgwick and Ross Converge in Moral Theory?

Sidgwick and Ross are, I think, the most important figures in what Tom Hurka has recently characterized as the Sidgwick to Ewing school. The philosophers in this school largely agreed about metaethics, endorsing non-naturalism and intuitionism. I am sympathetic to their metaethical views, but they won't be my focus here. My focus will instead be on the normative: on the question whether Sidgwick and Ross can converge in moral theory. I will argue that to a surprising extent they can. We can develop an interesting normative synthesis of Sidgwick and Ross.

Now, I should admit up front, convergence can come cheap. It will at least very often be possible to get philosophers' views to converge if you simply split differences between them, or simply allow that one of them is right in one respect and another right in another. Such cheaply acquired convergence may be a route to a plausible philosophical view. But it doesn't reveal anything much about the philosophers whose views you get to converge beyond the fact that, as you see it, each was on to part of the truth. Cheap convergence is not surprising.

I'm going to sketch a route to a synthesis of Sidgwick and Ross that, I will claim, is more interesting and surprising than that. It does involve some of the less interesting things: some difference-splitting and some mixed verdicts to the effect that Sidgwick is partly right and Ross is partly right. But it also derives from some significant commonalities between Sidgwick and Ross which seem to me well worth highlighting and exploring.

I'll focus in turn on three topics: the right, the good, and the appropriate conceptual framework for moral theory. And I'll argue that, with respect to the theory of the good and especially with respect to the theory of the right, Sidgwick and Ross have a lot more in common than they may initially seem to have.

1: The Right

Sidgwick and Ross can seem to disagree fundamentally in the theory of the right. After articulating his self-evident axioms, Sidgwick writes

I have tried to show how in the principles of Justice, Prudence, and Rational Benevolence as commonly recognised there is at least a self-evident element, immediately cognisable by abstract intuition; depending in each case on the relation which individuals and their particular ends bear as parts to their wholes, and to other parts of these wholes. I regard the apprehension, with more or less distinctness, of these abstract truths, as the permanent basis of the common conviction that the fundamental precepts of morality are essentially reasonable. No doubt these principles are often placed side by side with other precepts to which custom and general consent have given a merely illusory air of self-evidence: but the distinction between the two kinds of maxims appears to me to become manifest by merely reflecting on them. I know by direct reflection that the propositions, "I ought to speak the truth," "I ought to keep my promises," – however true they may be – are not self-evident to me; they present themselves as propositions requiring rational justification of some kind. On the other hand, the propositions, "I ought not to prefer a present lesser good to a future greater good," and "I ought not to prefer my own lesser good to the greater good of another," do present themselves as self-evident; as much (*e.g.*) as the mathematical axiom that 'if equals be added to equals the wholes are equal.' (382-383)

In a passage whose target (as usual in Ross's work) is Moore, not Sidgwick, Ross writes

If we are told, for instance, that we should give up our view that there is a special obligatoriness attaching to the keeping of promises because it is self-evident that the only duty is to produce as much good as possible, we have to ask ourselves whether we really, when we reflect, *are* convinced that this is self-evident, and whether we really *can* get rid of our view that promise-keeping has a bindingness independent of productiveness of maximum good. In my own experience I find that I cannot, in spite of a very genuine attempt to do so; and I venture to think that most people will find the same, and that just because they cannot lose the sense of special obligation, they cannot accept as self-evident, or even as true, the theory that would require them to do so. In fact, it seems, on reflection, self-evident that a promise, simply as such, is something that *prima facie* ought to be kept, and it does *not*, on reflection, seem self-evident that production of maximum good is the only thing that makes an act obligatory. (RG 39-40)

At first sight Sidgwick and Ross indeed seem to be in fundamental disagreement here. Sidgwick claims consequentialist principles are self-evident and deontological principles are not self-evident. Ross claims consequentialist principles are not self-evident and deontological principles are self-evident.

But this way of putting it overlooks important differences between the principles they consider. Focus first on the consequentialist principles. The consequentialist principle whose self-evidence Ross denies -which he gets from Moore, not from Sidgwick – is “the only duty is to produce as much good as possible”.

The key consequentialist principle Sidgwick endorses is instead the “maxim of benevolence”. When in III.XIII Sidgwick finds a version of the maxim in Clarke, one way he formulates it is

A rational agent is bound to aim at universal good. (385)

When the maxim is expressed this way, it is clear, as Tom Hurka and I have both in effect argued, that there are two possible interpretations. On the weaker interpretation, the maxim of benevolence says that there is a *prima facie* duty to aim at universal good. On the stronger interpretation, the maxim of benevolence says that there is an absolute duty to aim at universal good. Sidgwick does not distinguish these two interpretations. Some of what he writes suggests the one, some suggests the other. But, once you have made the distinction, it is natural to ask which interpretation gives Sidgwick the better view.

I think the weaker interpretation gives Sidgwick the better view. On the weaker interpretation, the maxim is a good candidate to be self-evident. By contrast, on the stronger interpretation, the maxim doesn't look to be self-evident or indeed to be true. There is surely strong reason to aim at universal good. But there may be conflicting reasons. One kind of potentially conflicting reason is a partial reason. It may be that what best promotes universal good will produce much less good *for me* than some other outcome. In such cases, one view is that I ought to aim at what is best for me not at what is universally best. Another kind of potentially conflicting reason is a deontological reason. It may be that aiming at universal good means violating a deontological principle (either an absolute deontological principle or a deontological principle of *prima facie* duty). In such cases one view is that I ought not to violate the deontological principle. So it is not at all self-evident that there is an absolute duty to aim at universal good: that I have a duty proper to aim at universal good even where doing so is much less good for me or where doing so means violating deontological principles.

But the maxim of benevolence on the weaker interpretation, the maxim of benevolence interpreted as a principle of *prima facie* duty is not merely not the consequentialist principle whose truth Ross explicitly denies in the passage above. It is a principle Ross himself accepts. As he writes elsewhere in the same chapter

It seems self-evident that if there are things that are intrinsically good, it is a *prima facie* duty to bring them into existence rather than not to do so, and to bring as much of them into existence as possible. (RG 24)

Thus, I suggest, notwithstanding the apparent contrasts in the passages with which I began this section, Sidgwick and Ross agree about the key self-evident consequentialist principle.

One possible response here is: the *prima facie* duty interpretation of the maxim of benevolence is *too* charitable. Sidgwick doesn't have the concept of *prima facie* duty. So the *prima facie* duty interpretation can at most be an account of what Sidgwick ought to have said or thought, not of what he did say or think.

I respect this possible response. (I won't take a stand here on whether it is right). But notice that, if our question is whether there is an interesting convergence between Sidgwick and Ross, it might not matter very much whether we say that the *prima facie* duty reading tells us what Sidgwick did think or what he ought to have thought. Either way, we have interesting convergence.

Now focus on Sidgwick's and Ross's views about deontological principles. The deontological principles whose self-evidence Sidgwick denies in the passage with which we began are not the deontological principles whose self-evidence Ross affirms. Sidgwick does not, of course and unfortunately, have the idea that deontological principles might be understood as principles of *prima facie* duty. What he rejects are deontological principles of duty proper. But Ross too rejects deontological principles of duty proper. The deontological principles Ross affirms are deontological principles of *prima facie* duty. The example he gives in the passage of a deontological claim that is self-evident is "a promise, simply as such, is something that *prima facie* ought to be kept".

Though Sidgwick regrettably does not consider the possibility of framing deontology in Ross's way, he would likely to want to reject deontological principles of *prima facie* duty. But, at least if Broad is right, Sidgwick would need new arguments against such principles. In case you don't know the story: In books published in 1930, Ross and Broad independently articulated versions of moderate deontology featuring versions of the concept of *prima facie* duty. Ross did not have Sidgwick particularly in mind. But Broad did. Having endorsed Sidgwick's critique of absolutist deontology, Broad proposes to

state a form of [deontology] which is not open to Sidgwick's objections and is not flagrantly in conflict with reflective common-sense. (218)

The key to this form of deontology is Broad's version of the concept of *prima facie* duty:

The [deontologist] will have to moderate his claims very greatly. He will be confined to statements about *tendencies* to be right and *tendencies* to be wrong. (222; italics in original)

But it isn't merely that Sidgwick and Ross disagree much less than they initially appear to about which consequentialist and deontological principles are self-evident. They also countenance deviations from consequentialism which have significant commonalities.

Here start with Ross. One characteristic way to understand the core of deontology is as a matter of absolute prohibitions. In introducing and deploying the concept of *prima facie* duty, Ross of course rejects this line. Other ways to understand the core of deontology emphasize something like a distinction between acts and omissions or between intending and foreseeing harm. Such treatments make the difference between right and wrong (or whatever) turn crucially on how an outcome is brought about.

Ross's deontology has something quite different at its core. To see this, consider some of his famous remarks in Chapter II of *The Right and the Good*:

The essential defect of the 'ideal utilitarian' theory is that it ignores, or at least does not do full justice to, the highly personal character of duty. If the only duty is to produce the maximum of good, the question who is to have the good – whether it is myself, or my benefactor, or a person to whom I have made a promise to confer that good on him, or a mere fellow man to whom I

stand in no such special relation – should make no difference to my having a duty to produce that good. But we are all in fact sure that it makes a vast difference. (RG 22)

He goes on to use for the duties of fidelity, reparation, and gratitude the term “special obligations,” which he contrasts with the consequentialist “general obligation” to do as much good as possible. Ross makes the existence of special obligations central to his rejection of ideal utilitarianism.

It is tempting to argue that there is something very similar in Sidgwick: that on the best interpretation, the dualism of practical reason combines the general obligation to do as much good as possible with special, partial obligations to yourself.

This of course is not Sidgwick’s official view. On his official view the dualism involves a contradiction. But, even so, a significant number of interpreters have read the dualism in a different way, as not involving a contradiction. To see why, we can begin with the third of the four conditions Sidgwick employs in III.XI in arguing against common-sense morality

The propositions accepted as self-evident must be mutually consistent. Here, again, it is obvious that any collision between two intuitions is a 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, without any slur being thrown on the scientific character of the conflicting formulae. Whereas such a collision is absolute proof that at least one of the formulae needs qualification. (M 341)

By Sidgwick’s own lights, the existence of apparent utilitarian and egoist intuitions that conflict with one another is “absolute proof that at least one of the formulae needs qualification”.

We have already, above, in effect reviewed one part of the argument that the way to qualify the formulae is to make them principles of *prima facie* duty rather than principles of duty proper. We argued that the best charitable interpretation of the maxim of benevolence is that it is a principle of *prima facie* duty. Do something similar with the principle of prudence - the apparently self-evident egoist half of the dualism - and you end up with a view strikingly similar to Ross’s. You end up with some kind of special reason to favor yourself, and with a reason to promote the general good.

Then, if you are synthesizing, you are apt to accept both kinds of non-utilitarian reasons, the Rossian ones and the Sidgwickian ones: deontological reasons and egoistic partial reasons. In doing so you are not, of course, agreeing wholly with either Sidgwick or Ross. But, you can say, Sidgwick has no good argument against deontological principles of *prima facie* duty. And, as we will see in the final section, Ross’s arguments against egoistic reasons are even worse.

2: The Good

Sidgwick’s and Ross’s commonalities with respect to the theory of the good are, I think, ultimately less interesting than their commonalities with respect to the theory of the right. But they are interesting nonetheless, and well worth noticing and reflecting on.

The thing to begin with, which Ross never to my knowledge notices, is just how similar the most important alternatives to hedonism he and Sidgwick consider are. Ross’s alternative is famously, of course, the theory of the good that he accepts. As he frames it (anyway in *The Right and the Good*), four things are intrinsically good: virtue, knowledge, pleasure, and desert: the allocation of pleasure to the virtuous and pain to the vicious rather than *vice versa*.

When Sidgwick argues for hedonism in III.XIV, the pluralistic theory he argues against in sections 4 and 5 has it that in addition to happiness

cognition of Truth, contemplation of Beauty, Free or Virtuous action [are] in some measure preferable alternatives to Pleasure or Happiness – even though we admit that Happiness must be included as a part of Ultimate Good. (ME 400)

Both lists include Ross's three basic intrinsic goods, virtue, knowledge and pleasure. Sidgwick includes also contemplation of Beauty (which Ross too may add in the *Foundations*). Sidgwick definitely does not include Ross's higher-order good, desert. But the lists are still strikingly similar.

Neither Sidgwick nor Ross ultimately endorses the theory in this simple and (to my mind) appealing form. They deviate from it in opposite directions. Sidgwick argues in III.XIV against it and in favor of hedonism. He (implausibly, I think) denies that virtue, knowledge etc. have any independent intrinsic value.

Ross, by contrast, has a strong anti-hedonist impulse that leaves him unsatisfied with the simple pluralist view. In *The Right and the Good* this impulse leads him to endorse the view that virtue is lexically prior to pleasure: that any extra amount of virtue however small compensates for any extra amount of pain however great. But the implausibility of this view can be shown by employing Ross's own favorite argumentative resource in these contexts: what we can call "abstract world arguments". Consider two states of the universe including nearly equal amounts of virtue – differing only in that one person on one occasion was less well motivated and acted less well in universe 1 than in universe 2 – but such that universe 1 includes also widespread and intense pleasure and universe 2 widespread and intense pain. On Ross's view we should choose universe 2. But surely we should instead choose universe 1.

In a way you get to a synthesis here by splitting the difference. But it's not just a compromise between distant and unrelated theories each of which tracks part of the truth. It's a compromise that starts with a simple and attractive pluralist view that Sidgwick and Ross both moot; and you compromise by rejecting their ill-advised moves in opposite directions away from that simple and attractive pluralist view.

3: The Conceptual Framework

To complete the synthesis you need to think about the conceptual framework - about what the fundamental ethical concepts are. Here, I suggest, both Sidgwick and Ross have something important and distinctive to contribute. I don't claim that you get a synthesis here by focusing on interesting commonalities between Sidgwick and Ross. Rather you get the best conceptual framework by combining what is distinctively right in Sidgwick (and missing in Ross) with what is distinctively right in Ross (and missing in Sidgwick).

What is distinctively right in Ross and missing in Sidgwick is the concept of *prima facie* duty, the notion, as Jonathan Dancy puts it, of the contributory. I take it to be pretty uncontroversial that Sidgwick really doesn't have the concept (or, anyway, deploy it at crucial points). As what I've already said will have suggested, I think when you introduce the concept of *prima facie* duty you have reason to reject Sidgwick's verdicts about both deontology and egoism: to think that he is too optimistic in thinking that he can dismiss deontology, and too pessimistic in thinking of the dualism of practical reason as involving an unresolvable contradiction

What is distinctively right in Sidgwick is that – unlike most of the Sidgwick to Ewing school – he makes a distinction between the generically normative and the specifically moral. And, when pressed, it is clear that the generically normative is for him more important. I don't suppose that this claim about

Sidgwick will be terribly controversial. That he makes such a distinction is part of the explanation why he treats egoism as a method of ethics.

Ross, by contrast, makes no such distinction. His unitary concept of duty blends the generically normative with the distinctively moral. This means that Ross does not possess the contemporary concept of a normative reason. By way of evidence for this claim, consider a striking contrast between Ross's views and Parfit's. At the very start of *On What Matters* Parfit introduces the concept of a normative reason. Since it is indefinable, it cannot be explained by giving a definition. Instead, Parfit says

We must explain...concepts [like "a reason"] in a different way, by getting people to think thoughts that use these concepts. One example is the thought that we always have a reason to avoid being in agony. (OWM 1, 31)

Parfit's appeal to this thought at this point suggests that he thinks one of the clearest and most compelling examples of a reason is the reason to avoid pain for oneself.

In a starkly contrasting passage, Ross writes

We are never conscious of a duty to get pleasure or avoid pain for ourselves, as we are conscious of a duty to give pleasure to or prevent pain for others. (F 277, italics added)

That is, I take it, the difference between Parfit and Ross can be captured like this:

Ross: We never have a *prima facie* duty to avoid pain for ourselves.

Parfit: We always have a reason to avoid pain for ourselves.

There are two possible ways to understand this difference. On one interpretation Parfit and Ross here share the same concept – “*prima facie* duty” and “normative reason” are just different names for this shared concept – and the difference between them is a substantive difference in opinions about what reasons or *prima facie* duties we have. On the other interpretation Parfit and Ross do not here share the same concept – Ross's term “*prima facie* duty” refers to a different concept than Parfit's term “normative reason” - so the difference here between Parfit and Ross is conceptual rather than substantive.

The latter diagnosis seems to me clearly right. To see why, notice first how radical Ross's substantive view would be if he were using the same concept as Parfit. Ross would not merely be denying that we have any *special* reason to avoid pain for ourselves; he would be denying that we have *any reason whatever* to avoid pain for ourselves. Other people's pains and pleasures would give us reasons for action; but our own pains and pleasures would give us no reasons for action. It would surely be exceptionally uncharitable to attribute such a view to Ross. Then, second, notice that there is a way to illuminate the difference between Ross's fundamental concept and Parfit's. Parfit's concept of a normative reason is, as Gibbard puts it, flavorless. It can be alternately expressed by talking about what makes sense. By contrast, Ross's fundamental concept is morally loaded; it is naturally expressed (as Ross himself does) by talking about “duty” and “obligation”. Ross denies that we have a *moral duty* to avoid pain for ourselves; Parfit thinks we always have a *normative reason* to avoid pain for ourselves. These claims do not straightforwardly conflict because they employ different fundamental concepts.

We need the concept of the generically normative to understand the appeal of egoism. Sidgwick has that concept; hence his (reluctant) sympathy for egoism. Ross lacks it; hence his failure to understand the appeal of egoism. In this crucial way, Sidgwick does better than Ross.

Thus neither Sidgwick nor Ross straightforwardly possesses the contemporary concept of a normative reason. To have that concept, you need to make two distinctions: between the generically normative and the specifically moral, and between the contributory and the overall. Sidgwick has the concept of the generically normative as opposed to the specifically moral, but, as we saw, he lacks or

anyway doesn't deploy at key places the concept of the contributory. Ross has the concept of the contributory; that's his key innovation, the idea of prima facie duty. But he lacks the concept of the generically normative as opposed to the specifically moral.

The synthesis of Sidgwick and Ross I'm sketching is, then, both conceptual and substantive. Conceptually, I suggest, we should think both in terms of the contributory and the generically normative; that is, we should think in terms of normative reasons. And when we do that, we are primed to see the appeal of a view that draws substantively both from Sidgwick and from Ross. Sidgwick and Ross are both committed to consequentialist reasons, to the reason to promote the good as one crucial element of normative theory. And they both think that there is another crucial element of normative theory, something agent-relative. They disagree about what this agent-relative element is. Sidgwick thinks there is nothing right about deontology but something right about egoism. Ross thinks there is something right about deontology but nothing right about egoism. Once we get the conceptual scheme right and see the similarities between Sidgwickian egoistic reasons and Rossian deontological reasons, we should admit reasons of both kinds: we should agree with Sidgwick about egoism and with Ross about deontology. And we should embrace the plausible pluralist view about the good that both Sidgwick and Ross articulated but unfortunately deviated from in opposite directions.