The methodological irrelevance of reflective equilibrium[*]

Tristram McPherson
Virginia Tech
dr.tristram@gmail.com

Penultimate Ms.: the official version will be published in the
Palgrave Handbook of Philosophical Methodology

Introduction

John Rawls' method of reflective equilibrium is the most influential methodology in contemporary ethics. This paper argues that this influence is undeserved. Rawls' method is highly implausible. Worse, it is also incapable of performing the work that motivates the search for a moral methodology in the first place. These are bold charges, and I dedicate the bulk of the paper to substantiating them (§§3-5). Several of the objections that I offer have been pressed before. However, when such objections are pressed in isolation from each other, it can seem easy to salvage the spirit (if not the letter) of the method, by judicious adjustment. It is much more difficult to do this once the inadequacies of the method are systematically displayed. I illustrate this point by exploring salient attempts to salvage the spirit of reflective equilibrium by abandoning elements of Rawls' approach (§6). I argue that none of these attempts succeed. I conclude that appeal to the method of reflective equilibrium is not a helpful means of addressing pressing methodological questions in ethics. In a slogan, reflective equilibrium is *methodologically irrelevant*.

1.   Desiderata for moral methodologies

This section and the next provide the background for my critical discussion. The core task of this section is to introduce and motivate three desiderata that I will use to evaluate Rawls' methodology. I begin by characterizing a philosophical methodology, and asking how we could motivate methodological enquiry in philosophy.

It will be useful to treat a philosophical methodology as consisting of two elements. The first is a *method*: a set of instructions or prescriptions for accomplishing some philosophical goal. The second is a *methodological theory*: this is a theory that purports to support adopting the method in question. We can think of the *method* of reflective equilibrium, strictly, as a method in the above sense. I will introduce this method in the next section. Rawls' methodological theory is the broader framework within which Rawls characterizes and defends this method in *A Theory of Justice* and *The Independence of Moral Theory*. I will introduce and discuss elements of this theory in §§3-5.

Focus on methodology in philosophy is controversial. For example, one might think of a method as a set of instructions that must have the following property: that someone who understands them can always tell whether she is following them. Timothy Williamson has recently argued that a method, so understood, is impossible (2008). Williamson's conclusion is controversial, but the best way for the methodologist to reply is perhaps to reject the conception of methodology that Williamson targets. A method could be highly informative and useful even if it is not luminous in this sense.

A simpler challenge to focusing on methodology begins by noting that the methodological question "How should we investigate the nature of justice?" is less intrinsically interesting than the substantive ethical question "What is justice?" One might try to meet this challenge by suggesting that focusing on

methodology is a good way of making progress on answering substantive ethical questions. But this in turn might seem to rest on Cartesian fantasies about philosophers' distinctively methodological competence.

Investigation into methodology, however, can be motivated in at least two ways that do not require such fantasies. First, suppose there are serious challenges to the very possibility of successful enquiry in a domain. In this case, methodological theory might be important as a way of assessing or answering such existential threats. Second, suppose that central substantive disagreements in a domain can be traced back to the differing methods being deployed by the disagreeing parties. This motivates pursuing the substantive disagreement back to its source, and seeking to adjudicate between those competing methods. These motivations gather strength in proportion to the plausibility of the two suppositions. There is a good case to be made for both suppositions in ethics. Consider a brief sketch of each case, in turn.

Perhaps the most familiar approach to normative ethical theorizing centrally involves the appeal to intuitions about possible cases. However, there is vigorous disagreement regarding the role that such intuitions should play in our ethical theorizing. Is it unavoidable?[1] To be avoided at all costs?[2] Should we appeal only to real or realistic cases? Or can we appeal with abandon to elegant but outlandish scenarios?[3] Should intuitive judgments about cases be given priority over other intuitive judgments, such as judgments concerning the plausibility of principles, or concerning the moral significance of certain kinds

---

[1] It is often presupposed that in ethics and other areas of philosophy, appeal to intuition plays a central and possibly ineliminable justificatory role (see for example Kagan 2001, p. 44). This presupposition has recently been challenged by (Williamson 2007) and (Cappelen 2012). The nature of intuitions is itself a controversial issue. For a variety of views concerning the nature and significance of intuitions, see (DePaul and Ramsey, 1998).

[2] Representative worries about ethical intuitions can be found in (Hare 1981, p. 130ff) and (Singer 2005).

[3] Worries about the philosophical use of imaginary cases are suggested by (Dancy 1985) and (Wilkes 1988). Suspicion about appeal to *unrealistic* cases is widespread but rarely defended in detail.

of properties, or concerning the significance of certain theoretical desiderata?[4] Finally, can we augment or even supplant the appeal to intuition by producing empirical results that bear on normative ethical theses?[5]

These disputes (and many other like them) have substantive bite. To take just one example, a methodology that privileged plausible general principles over case-specific judgments (especially those about outlandish cases) would make it much easier to defend standard forms of consequentialism, and much harder to defend the sorts of complex deontological principles proposed by Frances Kamm (2006).

Next consider four representative (but far from exhaustive) anxieties about ethical enquiry, which might appear to cast doubt on our ability to successfully engage in it. First, ethical enquiry appears to contrast strikingly – both in terms of methods and cumulative success – with our scientific paradigms of successful enquiry.[6] This casts doubt on whether substantial theoretical success in ethics is possible. Second, disagreement in ethics is deep, pervasive, and seemingly intractable. This naturally prompts the question of whether we are simply incompetent to answer controversial ethical questions (compare McGrath 2008). Third, philosophers have worried loudly about the distinctive vulnerability of our moral beliefs to emotional influence,[7] to

---

[4] See (Huemer 2008) for principled defense of an unusual assessment of the relative significance of different types of intuitions. Appeals to the plausibility of moral principles are perhaps as old as ethics. Appeals to the moral (ir)relevance of certain properties and to theoretical desiderata are perhaps less familiar. For an example, consider (Unger 1996), which follows (Singer 1972) in appealing to the intuitive idea that mere physical proximity cannot be a morally relevant property. The appeal of theoretical desiderata like simplicity and explanatory power helps to explain why we have seen so many efforts to work out normative theories that explain everything about morality in consequentialist or deontological or aretaic terms. Rossians and particularists, of course, argue that these desiderata are not decisive. See for example (McNaughton 2003).

[5] A good (if now dated) point of entry into this burgeoning empirical literature is (Doris and Stich 2006).

[6] Many challenges to ethics, such as the familiar thought that ethical commitment is in tension with a 'disenchanted' modern worldview, presuppose this contrast. The idea that there is a methodological constrast between ethics and science has been challenged, for example by (Boyd 1997 §4.4).

[7] See, for example (Sinnott-Armstrong 2006, §9.4.3).

rationalization of our existing in-group lifestyles,[8] or to evolutionary pressures.[9] These complement more traditional anxieties about ideological influence.[10] Together, these constitute a laundry list of apparently ethically arbitrary factors that are arguably implicated in the formation of our beliefs about ethics. Finally, many of us take ethical enquiry to aim to discover ethical facts, not to create them. Yet, when answering the most fundamental ethical questions, it can at least appear to be more appropriate to *think hard* than to *go look*. This raises the familiar but nonetheless deep question of how we can hope to discover objective facts about the world from the armchair.

These anxieties complement each other. For example, our apparent sensitivity to untrustworthy inputs might be used to explain persistent ethical disagreement. This explanation would in turn make the armchair distinctively less comfortable in ethics than it is in mathematics (for example), where ideology and emotion have less incentive for purchase, and disagreement is less endemic.[11] Likewise, in the absence of a positive story about armchair reliability, such sensitivity becomes a potentially damning explanation for the striking contrast between ethical enquiry and paradigms of successful scientific enquiry.

The points just sketched motivate the project of moral methodology. The scope and substantive significance of controversy in moral methodology provides rich opportunities for 'dialectical ascent' from ethical to methodological disagreement. And the force of the representative anxieties should prompt interest in methodology, in order to explore whether ethical enquiry can actually attain its goals.

These points also suggest natural desiderata by which we can assess candidate methodological theories in ethics:

---

[8] The locus classicus here is (Mackie 1977, p. 36).
[9] See, among many recent discussions, (Joyce 2006, ch. 6).
[10] See (Railton 2012, essay 12) for helpful discussion of the challenge from ideology.
[11] But see (Clarke-Doane *forthcoming*) for a defense of the comparative epistemological plausibility of moral as opposed to mathematical realism.

Vindication	does the candidate methodological theory address the central challenges to the possibility of successful ethical enquiry?
Adjudication	does the candidate methodological theory purport to resolve the central controversies within moral methodology?
Plausibility	are the distinctive features of the methodology plausible in the face of objection?

Plausibility is a central desideratum on any philosophical theory: a theory that cannot be adequately defended does not deserve our allegiance. **Vindication** and **Adjudication** are desiderata that reflect the two motives for engaging in moral methodology just sketched. If what I have suggested above is correct, a moral methodology that fails to meet these desiderata to any significant degree thereby fails to accomplish the tasks that warrant attention to moral methodology in the first place. In light of this, I will adopt these desiderata as the criteria against which we should assess Rawls' methodology. I now turn to introducing that methodology.

2. Rawlsian hopes

In this section, I sketch some of the distinctive elements of Rawls' method of reflective equilibrium (hereafter, for brevity: *Rawls' method,* or *the method*). I then illustrate the sketch by explaining how Rawls understands the method's role in one of his own central arguments. With this preliminary characterization in hand, I suggest reasons for initial optimism concerning Rawls' ability to meet the desiderata introduced in the preceding section.

We can isolate two structural components of the method of reflective equilibrium, suggested by Rawls' canonical account in *A Theory of Justice.*[12] The first is an account of the central *inputs* to be drawn on in moral theorizing, and

---

[12] Rawls offers a distinct (and inferior) precursor to the method in his (1951). For a brief criticism of one feature of that precursor, see my (2012, p. 532).

the second is an account of the *operations* to be performed with those inputs. The central inputs to the method are what Rawls dubs *considered judgments* about moral matters: those judgments one sincerely and stably affirms in conditions which minimize the influence of strong emotion or self-interest (1999b, p. 42). The operations begin by formulating what one takes to be plausible moral principles, which purport to explain the ethical theses expressed by one's considered judgments. Because it is typically very difficult to find plausible principles that adequately explain the full range of one's considered judgments, the method then requires that one attempt to fit these judgments and principles into a coherent scheme.[13] One does this by modifying, adding, or abandoning principles or particular judgments as one deems appropriate on reflection (1999b, p. 18). Finally, during this process, one is also supposed to take into consideration the leading theories on the topic, and the arguments that can be made in support of each (cf. 1999b, p. 43).

The method can be helpfully illustrated by a canonical use that Rawls puts it to: his explanation of the justification of his conception of the original position in *A Theory of Justice* (1999b, p. 17-18). As a contractualist, Rawls distinguishes two sorts of broadly moral principles: substantive principles that characterize ideally just institutions, and procedural principles that characterize the contractual procedure that generates the substantive principles. Rawls introduces the method after having explained the idea of procedural justification of principles via an idealized contractual situation that he calls the 'original position'.

Rawls explains that in using the method in this case, we seek equilibrium between our considered judgments about justice on the one hand, and the contractual procedure (and the principles this procedure generates) on

---

[13] I have formulated this in individualistic terms. However, Rawls himself often prefers to talk of what 'we' do. The characterization of the method in his (1951) suggests taking this seriously: there he emphasizes the role of the *shared* judgments of the set of 'competent judges'. However, Rawls' characterization of moral theory (discussed in the next section) strongly suggests that in the mature version of the method, each individual is a potentially distinct object of investigation.

the other, making adjustments to each as seems intuitively best (1999b, p. 18). The contractual machinery is in turn supposed to reflect our considered judgments about just procedures. Thus, principles have some claim to be adequate principles of justice because they would be agreed to in the original position (1999b, p. 17). However, if an otherwise plausible account of the original position delivered principles that produced repugnant results when applied, it would be substantively unacceptable. So the description of the contractual situation is apt for adjustment not only in light of our procedural judgments, but also in light of its implications for our judgments about substantive justice.

This sketched application makes the use of Rawls' method appear familiar. It also appears to provide reason for optimism concerning the three desiderata introduced in the previous section.

First, Rawls' methodological theory appears to be provide **Adjudication**, at least assuming that it supports the method just sketched. Consider some examples. The role of considered moral judgments in the method seems to reject priority for either case-specific judgments, or moral principles.[14] This is because, as Rawls emphasizes, considered judgments are possible at all levels of generality (1999a, p. 289). Further, there is no automatic privileging of judgments about realistic cases (or, conversely, of judgments about outlandish cases). Finally, the method identifies considered moral judgments as the privileged inputs to moral enquiry. This contrasts with views that require that moral enquiry begins with our *intuitions*, at least if intuitions are construed as psychological states distinct from beliefs. All of this adds up to a considerable appearance of methodological substance.

We can also make a seemingly strong indirect case for the **Plausibility** of Rawls' methodology. Moral philosophers are credible judges of the plausibility of a moral methodology. And it is remarkably common for such

---

[14] This is somewhat controversial; for example (Singer 1974, p. 516-7) charges that Rawls' method *de facto* advantages particular moral judgments over general principles.

philosophers to endorse the method. Two especially ambitious examples illustrate the strength of that endorsement. First, Michael Smith suggests that Rawls' account of the method successfully systematizes our methodological platitudes (1994, p. 40). Second, Shelly Kagan suggests that *all* practicing normative theorists are at least implicitly committed to something very similar to the method (1998, p. 16).

So far, we have considered only Rawls' proposed method, and not the broader methodological theory that underlies it. One might hope that this theory would substantially address some of the pressing challenges to the possibility of successful ethical enquiry, and thereby provide **Vindication**.

Such are the hopes that one might entertain for the methodological significance of the method of reflective equilibrium. In the next three sections, I explore several crucial aspects of Rawls' broader methodological theory. I argue that understanding this theory forces us to abandon each of the three hopes entertained here.

3.   Against Vindication: Rawls on moral theory

A methodological theory should provide a conception of the theoretical project that a given method is suited to accomplishing. Rawls characterizes the method of reflective equilibrium primarily as a means of engaging in what he calls *moral theory* (1999b, §9; 1999a, p. 288). In this section, I explain the role of moral theory in Rawls' account, and argue that it fails to enable Rawls' methodology to meet the **Vindication** desideratum.

In *A Theory of Justice,* Rawls describes moral theory as (roughly) the project of describing our 'moral capacity' (1999b, p. 41). Moral theory is part of a broader domain of enquiry that Rawls calls *moral philosophy,* which includes the way that moral enquiry intersects with the theory of meaning, epistemology, metaphysics, and the philosophy of mind (1999a, p. 287). In *The Independence of Moral Theory* Rawls explains his conception of moral theory in more detail. Moral theory is the study of 'substantive moral conceptions' which

aims to systematically compare those conceptions (1999a, pp. 286; 290-1). Strikingly, when engaging in moral theory, we '…put aside the idea of constructing a correct theory of right and wrong…' (1999a, p. 288). Thus, when studying one's *own* moral conception as a moral theorist, one must treat it simply as an aspect of human psychology (ibid.). So understood, *moral theory* in Rawls' special sense is a far cry from *normative ethics* as it is ordinarily understood. The former consists in normatively detached psychological investigation, while the latter is a normatively committed attempt to answer substantive ethical questions.

The fact that reflective equilibrium is proposed as a method for engaging in moral theory suggests that the goal of the method is to enable us to understand and systematically compare those moral conceptions that are psychologically realized in humans. This psychological conception of the goal of the method might appear to provide Rawls with a quick reply to the methodological challenges that apparently face ethical theorizing (such as those mentioned in §1 above). For example, consider the challenges based in pervasive and seemingly intractable disagreement, or those based on the apparent sensitivity of our moral thinking to untrustworthy inputs. Both of these challenges take their force from the idea that they undermine our reasons to believe that our moral judgments are *true*. But the discussion of moral theory suggests that Rawls' method aims at *understanding* our moral conceptions, and not evaluating their correctness. Treating the ethical theorist as engaging in Rawlsian moral theory might thus seem to inoculate her against these challenges, and hence suggest that the Rawlsian methodology provides **Vindication**. Provided that Rawls' method permits us to accurately characterize our moral sensibilities, the reliability of those sensibilities in latching onto alleged objective truth is irrelevant to the success of the method.

This is the only basis on which one might hope to show that Rawlsian methodology provides Vindication. But it faces a difficult dilemma. Is

describing our moral sensibilities the Rawlsian's theoretical goal, or merely a (possibly necessary) step along the way to achieving that goal?

Rawls himself appears to favor the second horn of this dilemma. Thus, in *Independence...* Rawls' case for the significance of moral theory amounts largely to the conjecture that progress in moral theory may be an essential propaedeutic to broader progress in moral philosophy (1999a, p. 291, cf. also p. 287).[15] And this possibility helps to account for the philosophical interest of moral theory in the first place. If we follow Rawls here, however, investigation of moral theory merely postpones the challenges to the possibility of successful ethical enquiry, rather than answering them. The challenges arise again, as soon as the Rawlsian begins to investigate the correctness of the contents of an ethical thesis. And here, the method appears to have nothing distinctive to say. On this horn of the dilemma, then, Rawls' methodological theory simply fails to meet the **Vindication** desideratum: it says nothing one way or the other about whether or how challenges to the possibility of successful ethical enquiry can be met.

This suggests that we should explore the first horn of the dilemma: the idea that understanding our moral sensibilities is the ethical enquirer's central goal. This idea should engender immediate suspicion, since most practicing normative ethicists would be indignant at the claim that they are seeking only to explore their own psychologies.

David Copp has recently defended a deflationary conception of the goal of normative ethical theorizing that might seem to help the Rawlsian address this suspicion. Copp introduces a technical term: the *proximate goal* of work in a discipline. It is definitive of a proximate goal that it is the metric against which candidate methods within a discipline should be assessed (2012, p. 22).

---

[15] This point is crucial to interpreting certain passages in Rawls. For example: 'There is a definite if limited class of facts against which conjectured [moral] principles can be checked, namely our considered judgments in reflective equilibrium.' (1999b, p. 44). This might seem to suggest a simple response-dependent metaphysics of ethics. However, it is best read instead as making a claim about doing moral theory, not about the final stages of moral enquiry.

Copp argues further that practitioners within a discipline must be able to tell whether they are accomplishing the proximate goal, by their efforts *within* that discipline (2012, p. 23). On this basis, Copp argues that characterizing our moral sensibility as it would be in reflective equilibrium is the proximate goal of normative ethical theorizing (2012, pp. 12-15, especially p. 13).

Copp's argument, if successful, would permit the Rawlsian to claim that their methodology achieves **Vindication**. This is because on Copp's account, the adequacy of methods must be assessed against the proximate goal, and as we have seen, taking our goal to be to characterize our moral sensibility appears to undercut the applicability of the central challenges to ethical enquiry.

Copp's argument, however, rests on a highly implausible assumption: that the goal *relative to which methods in a discipline can be assessed* must also be one that *practitioners can tell whether they are accomplishing, by their efforts within that discipline*. This assumption can be shown to be unwarranted by example. Suppose that you attempt to debunk my astrological theories by appealing to evidence from astrophysics and psychology. It would be absurd for me to protest that your criticisms are inert, because the success of my astrological theorizing must be something that I can determine *by doing astrology*![16] Astrology is an example of a *hopeless* discipline: one whose goals are not achievable by using the characteristic methods of that discipline. The most dramatic challenges to ethical enquiry suggest that it is hopeless in this way.

If this is right, the normative ethicist cannot rest content with the goal of characterizing our moral sensibilities. And this means that Rawls' focus on moral theory at best postpones, rather than answers, the most serious

---

[16] The same point arguably holds for successful disciplines. Consider a controversial but plausible view in the philosophy of science. On this view, science has realist aims, but determining whether science achieves those aims – e.g. by assessing 'pessimistic induction' and 'no miracles' arguments – is something that requires philosophical rather than scientific investigation. I am indebted to Chris Daly for a variant of this example.

methodological challenges to the project of ethical theorizing. In light of this, Rawls' methodological theory does nothing to show that ethical enquiry is not hopeless, and hence does not meet the **Vindication** desideratum for moral methodologies.

4. Against Adjudication: Rawls on wide reflective equilibrium

Unlike the initial hopes that Rawlsian methodology could provide **Vindication**, the initial case that it can provide **Adjudication** appeared substantial. As we saw in §2, the heart of the method is an account of the inputs to ethical theorizing (the considered moral judgments), and an account of the operations one is to perform on those inputs. And both of these elements appear to provide substantive guidance. In this section, I argue that further elements of Rawls' methodological theory show this appearance to be misleading. To see this, we need to further understand the relationship between the method and its (provisional) goal.

As we have seen, Rawls' characterization of the goal of the method is *psychological*: it is to characterize our moral sensibilities. This might seem to render puzzling his suggestion that we should use the method to pursue this goal. After all, it is natural to think that the ordinary methods of social science would be much more appropriate to investigating the target psychological facts.

This challenge can be mitigated by further clarifying the goal of moral theory. Rawls suggests that, if we are 'philosophically motivated', our goal will not be to characterize our *actual* sensibilities. Rather, the sensibility that we are interested in as moral theorists is one that has been *idealized*: one in which the agent's sensibility is in *wide reflective equilibrium*. This equilibrium is a state of harmony between the particular judgments and general principles held by an agent *after* that agent has been presented with all candidate ethical

theories, and all reasonable arguments for (and, one assumes, against) each of them (1999b, p. 43; 1999a, p. 289).[17]

Rawls does not explain why this idealization constitutes an especially apt target for philosophical investigation. I offer two conjectures. First, a non-idealized moral psychology might tend to be a theoretically uninteresting mess. By contrast, a moral conception in wide reflective equilibrium will (by hypothesis) be coherent, and might be hoped to display interesting theoretical structure. Second, a non-ideal moral psychology might be highly dialectically unstable: for example, I might be one reasonable argument away from radically altering my ethical views. By contrast, a psychology in wide reflective equilibrium has a stable moral view that has resulted from careful consideration of all reasonable arguments. In light of this, such a psychology is *dialectically invulnerable*: it will be impossible to convince someone in wide reflective equilibrium out of their moral view simply by presenting them with a reasonable moral argument, because they will, by hypothesis already have carefully considered it en route to arriving at their view.

In light of its extraordinarily idealized character, wide reflective equilibrium is not a state that an ordinary human could ever be in. This precludes at least direct investigation of this state by scientific psychological methods. This might seem to mitigate the puzzle that I mentioned above: social scientific investigation of the target facts is not a serious competitor to Rawls' method, because it is not possible.

One might object that scientific investigation into how our psychologies actually work is nonetheless a potentially invaluable step in attempting to theorize how these psychologies would work in ideal circumstances. Consider one highly speculative example. Our moral conceptions might turn out to typically be organized as Roschian *prototype clusters*, or as *connectionist systems*, as opposed to structures that are well described in terms of moral

---

[17] Rawls only dubs this goal 'wide reflective equilibrium' in his (1999a), but it is clear he is talking about the same state in his (1999b).

principles.[18] If so, understanding these psychological facts might be essential information for the moral theorist. On the one hand, a reasonable default presumption is that the idealized form of a moral conception would share the basic structure with its non-idealized analogue. In this case, knowledge of the psychological structure of our actual moral thought could substantially inform moral theory. On the other hand, one might argue that a prototype organization (for example) is for some reason unacceptably non-ideal. In this case, knowledge of our actual psychological structure might provide important reasons to reject certain moral judgments (say: judgments about non-paradigmatic cases) as artifacts arising from the allegedly unacceptable structure.[19]

The Rawlsian has a compelling reply to this objection: if the objector is correct about the significance of such scientific investigation for moral theory, then the method of reflective equilibrium commends us to engage in such investigation. This is because, as we have seen, the method commends us to consider all moral conceptions, and *all reasonable arguments* for each of them. And the objector is exactly proposing that there will likely be reasonable arguments in favor of moral conceptions (as descriptions of what a psychology would accept in wide reflective equilibrium) that appeal in part to empirically discoverable facts about psychological structure.

This is a fair reply. Unfortunately, the logic of this reply can be generalized to undercut the claim that the method of reflective equilibrium **Adjudicates** central methodological controversies in ethics. To begin to see why, consider a single vivid example. Josh Greene (2008) has drawn on a range of empirical work to argue that the psychological provenance of our characteristically deontological ethical judgments debunks deontological ethics as an objectionable bit of post-hoc rationalization. Set aside whether this

---

[18] See for example (Rosch 1978) and (Bechtel 1987) respectively.
[19] For relevant discussion of the commitments of reflective equilibrium with respect to theoretical structure in ethics, see (Schroeter 2004).

argument is *sound* (it is not). It nonetheless surely has some claim to be *reasonable*. And if so, it will be a required input to reflective equilibrium.

This case illustrates a general lesson. Consider any of the controversies within moral methodology that I mentioned in §1. There are surely reasonable (even if incorrect) views on both sides of these controversies. And these methodological views can potentially be used to mount arguments for or against substantive moral theses (this is why accomplishing **Adjudication** is a central motive for engaging in moral methodology). But this in turn means that the method must embrace such arguments as legitimate inputs to the process of reflective equilibrium. Notice further that this applies as well to methodological arguments that challenge the features that initially appeared distinctive of the method: the characterization of legitimate starting points, and of the operations one is to perform with those starting points. For example, Greene's argument, if accepted, would screen off our deontological considered judgments as illegitimate inputs.

In short, in light of suggesting that the ideal moral enquirer consider all reasonable arguments, the method of reflective equilibrium fails to **Adjudicate** the central controversies within moral methodology. Rather, it in effect *absorbs* all such controversies, as inputs. This undercuts any claim that the method provides an informative answer to pressing methodological questions in ethics. Note that this is true even holding as fixed Rawls' project of moral theory; how we should proceed if and when we attempt to move from moral theory to actually defending answers to ethical questions is even less clear.

5. The Implausibility of Rawlsian methodology

So far, I have argued that Rawlsian methodology fails to **Vindicate** ethical theorizing in the face of challenge, and that it fails to **Adjudicate** central controversies within moral methodology. These results are disappointing, but they are compatible with a certain attenuated role for the Rawlsian methodology. It might be claimed to provide a very broad methodological

framework that we have good reason to accept, and *within which* more substantive investigation of the **Vindication** and **Adjudication** questions must be carried out. In this section I argue that it would be a mistake to grant Rawlsian methodolgy even this limited role. The previous two sections have emphasized the limits to the methodological substance of Rawls' account. In this section, I argue that the substance that remains – both in Rawls' method, and in his theoretical rationale for adopting it – should be rejected.[20]

I begin with Rawls' theoretical rationale. As we have seen, Rawls motivates the deployment of the *method* of reflective equilibrium by appealing to the philosophical interest of psychologies that are in the *ideal state* of wide reflective equilibrium. An idealized psychology is supposed to reach wide reflective equilibrium by deploying the method. And Rawls appears to assume that this motivates our deploying the method as a way of investigating what that ideal state would be like. However, this assumed connection between the ideal and the wisdom of our use of the method is far from obvious. The literature on non-ideal theory in political philosophy has taught us that our actual circumstances may require quite different norms than those that would apply in ideal circumstances. Consider a non-ethical example: an *ideal* psychology – one with no computational limits – could play perfect chess by considering every branch in the tree of possible variations. But the average chess player would do miserably worse than usual by trying to mimic this algorithm.

A deeper worry about Rawls' appeal to wide reflective equilibrium is that there is likely no such thing as *the* set of principles that I would accept in wide reflective equilibrium. The core worry here is that idealization can be

---

[20] Considerations of space preclude detailed examination of the plausibility of the role that Rawls gives to moral theory. This is dubious on at least two grounds. First, Rawls give no adequate defense of the 'priority' of moral theory over other questions in moral philosophy (for relevant discussion, see my (2012)). Second, Rawls' morally detached conception of the project of moral theory is ill-defended and unappealing. See (Scanlon 2003, p. 142-8) for an 'interpretation' of Rawls which abandons this feature. This interpretation appears to be motivated largely by Scanlon's recognition that it would make Rawlsian methodology more worthy of our attention.

expected to be *path dependent*: such that the order in which novel arguments (for example) are presented to a subject can substantially affect the outcome of idealization. If the idealization process leading to wide reflective equilibrium is substantially path-dependent, there might be no determinate fact about what I would believe in wide reflective equilibrium. Instead, there would only be facts about what I would believe if I reached equilibrium by this route rather than that one.[21]

Rawls' account is dispositional in nature: it asks us to consider how our psychologies *would* evolve, given reflection on the relevant information. This fact makes path-dependence worries even more pressing. For psychologists claim to have found significant *order effects* in how people evaluate certain ethical thought experiments. That is, they claim that we will tend to provide different ethical evaluations of certain pairs of thought experiments, depending on which of the pair they are asked to think about first (Petrinovich and O'Neill 1996; Schwitzgebel and Cushman 2012).

Consider the vastly many different sequences in which I could be provided with the total set of reasonable ethical arguments and asked to make my moral judgments coherent. Path-dependence suggests that it is very plausible that these sequences will not converge on a single ethical theory, or even on a cluster of similar theories. This is a serious problem for Rawls' assumption that the method is a means of ascertaining what we would believe in wide reflective equilibrium. If there is no unique fact of the matter about what we would believe in such a state, then this assumption rests on a false presupposition.

I now turn from Rawls' methodological theory to the method itself, arguing that it allows our dispositions an implausibly central role. Dispositions enter into the method in two places. First, the characterization of the inputs to the method (considered moral judgments) is dispositional. These judgments

---

[21] In formal updating models, such path-dependence worries are quite compelling: updating procedures that are not path-dependent typically exhibit other objectionable features. For relevant discussion, see (Bonevac 2004).

are (inter alia) those made *confidently,* and *without hesitation* in certain circumstances (1999b, p. 42). Second, Rawls' characterization of adjustment and equilibrium are dispositional. As Rawls says (discussing a particular example): 'Reflective equilibrium requires only that the agent make these revisions with conviction and confidence, and continues to affirm these principles when it comes to accepting their consequences in practice' (1999a, p. 289). Both of these roles for dispositions are very hard to defend.

The dispositional criterion of considered moral judgments means that the method can endorse intuitively monstrous judgments as appropriate starting points for normative theorizing, provided these judgments are held with the right sort of dispositions (Kelly and McGrath 2010, p. 346-8). Some philosophers may shrug this worry off; perhaps substantive content is not the right criterion for assessing a theoretical starting point. Such philosophers might be less comfortable countenancing considered moral judgments that insist on the significance of obviously arbitrary considerations, or considerations that violate plausible formal constraints. But we can imagine someone confidently and stably judging that pain is bad – except on a future Tuesday, or denying the transitivity of the *morally better than* relation.[22] Pretheoretical possession of such judgments would arguably constitute evidence of moral idiocy, not the basis for an unusual but credible moral perspective.

A similar point applies to the role of dispositions in the operations that the method prescribes that we perform on our considered judgments. The dispositional character of the adjustment procedure ensures that any reflectively endorsed disposition to adjustment, no matter how bizarre, will potentially count as legitimate. This is again implausible. Suppose that Joe happens to be robustly disposed, on noticing a conflict between his judgments about capital punishment, to form a novel (and substantively unconnected)

---

[22] For future-Tuesday indifference, see (Parfit 1984, pp. 123-4); for the claim that 'formal' ethical judgments (such as the transitivity of value) are especially trustworthy, see (Huemer 2008, p. 386).

belief about the moral value of art. And suppose that Joe is inclined to reflectively endorse that transition as reasonable. Joe's exercising this disposition is an appropriate object of pity, not methodological endorsement.

The possibility of psychologically robust irrational dispositions is no mere philosopher's fantasy. For example, some ordinary reasoners are robustly inclined to endorse irrational inference patterns such as the gambler's fallacy (Stich and Nisbett 1980, pp. 192-3). Our confidence in the irrationality of such inferences stems from our possession of rationally compelling arguments against them. However, the fact that these arguments are rationally compelling does not guarantee that they would be *psychologically* effective when addressed to those tempted to the gambler's fallacy.[23] Further, the fallaciousness of arguments in favor of the gambler's fallacy does not guarantee those arguments to be psychologically *in*effective (Stich and Nisbett 1980 pp. 196-7).   Thus, it may be empirically plausible that straightforwardly objectionable principles of reasoning such as the gambler's fallacy could survive into a wide reflective equilibrium, on the dispositional account.

The task of a moral methodology is in part to *theorize* substantive and procedural irrationality in ethical enquiry. As this discussion shows, the method of reflective equilibrium instead embraces many clear instances of such irrationality as legitimate, provided those instances are psychologically robust. Beyond this substantive implausibility, however, the method also defers to our dispositions in an objectionably *arbitrary* way.

Both in its account of inputs and of revision, the method does impose constraints on our dispositions. For example, consistency is a success constraint on the equilibrating operations proposed by the method. As Timothy Williamson (2007, pp. 244-6) points out, in appealing to logical relations like consistency, the method appeals to philosophically contestable evidence. This raises the pressing question: if we are allowed to appeal to such

---

[23] Embarrassing autobiographical analogy: in my teens, I lost a little money to a friend who clearly explained the correct reasoning about the Monty Hall case to me, and then, in the face of my intransigence, goaded me into gambling on it.

evidence here, why can we not also appeal to such evidence to override an agent's dispositions to endorse future Tuesday indifference, or the gambler's fallacy? A similar point applies to the method's account of inputs. Moral judgments are ruled out as inputs if made about cases where one stands to gain. This is an objective constraint; it will apply whether one is disposed to endorse it or not. Besides raising the arbitrariness worry just mentioned, this constraint also appears to be an ossified rule of thumb, which is clearly objectionable in many cases. For example, it implausibly requires us to discount judgments about oppression made by the oppressed (compare Kelly and McGrath 2010, pp. 348-9 for this point).

In this section, I have argued that despite giving very little guidance where it is most needed, Rawlsian methodology manages to be highly implausible. The methodological theory underlying the method is dubious because it proposes the method as a way to investigate a theoretical object – a psychology in wide reflective equilibrium – that likely has no determinate character. The method itself is objectionable, because it appeals repeatedly to the subject's dispositions exactly where normative substance is needed.

6. The method without Rawlsian baggage?

Many philosophers claim allegiance to reflective equilibrium without intending to endorse either Rawls' methodological theory, or the precise details of his method. It might thus seem possible that the basic proposal of assembling our considered judgments and seeking equilibrium between them could be salvaged by detaching them from the details of Rawls' own articulation and defense. I cannot consider every possible effort of this kind here. Notably, I cannot rule out the possibility that someone could construct a compelling methodological theory that met my three desiderata, and in turn underwrote something resembling the Rawlsian method. In this section, I consider three related proposals, which strike me as the most initially inviting alternatives to that theoretically ambitious possibility. These involve replacing Rawls'

problematic dispositional characterizations with normative ones, appealing to the inescapability of the method, and appealing to the idea that the method characterizes a distinctively subjective form of rationality.

To begin, notice that jettisoning Rawls' methodological theory – the appeal to moral theory and wide reflective equilibrium – provides some relief from the challenges that I have posed. However, it does nothing to address the implausibility of the dispositional character of the method. One natural amendment here replaces Rawls' *dispositional* characterization of considered judgment and revision with *normative* characterizations. On this approach, one might say that someone's considered judgments are those judgments held by that person which it is *reasonable* to treat as starting points for theorization (whatever that comes to). And with those starting points in hand, one should deploy the reasonable canons of updating (whatever those are).

This amendment avoids the final problem that I raised for the method. For example, the judgment that there is nothing wrong with future suffering, provided that it occurs on a Tuesday, could be dismissed as an unacceptable starting point. And it blunts the force of Williamson's query about why the method is taking on some philosophically controversial views, while leaving others to be settled by the dispositions of the subject: the amended account potentially leaves nothing 'up to the subject' in this way.

I have two (by now predictable) concerns about the method so revised. The first worry is that it is close to vacuous with respect to the **Adjudication** desideratum. Almost anyone who thinks successful moral enquiry is possible would agree that we ought to start with reasonable inputs, and then apply reasonable updating procedures. The central methodological debates in ethics can be characterized as concerning what counts as reasonable starting points and procedures. The second worry is that this version of the method has nothing helpful to say about the **Vindication** desideratum. Nothing in the proposal tells us whether we can successfully answer challenges to the possibility of successful normative ethical theorizing, or how.

A second suggestion sometimes made in defense of the method is that it is somehow inescapable. For example, T. M. Scanlon suggests that '…apparent alternatives to [the method] are illusory.' (2003, p. 149). This is a popular thought, but it is often hard to pin down what it is supposed to come to. In Scanlon's case, it appears connected to the point I have emphasized in §4: that wide reflective equilibrium requires consideration of all reasonable arguments. In light of this, Scanlon suggests that the method '…allows for what might have been seen as alternative methods of justification to be incorporated within it' (2003, p. 151). Indeed, according to Scanlon, the method ultimately '…becomes simply the truism that we should decide what views about justice to adopt by considering the philosophical arguments for all possible views and assessing them on their merits.' (ibid.).

One might reasonably worry that the method was completely *uninformative*, if it was truly compatible with all competing methodologies. But Scanlon rejects this possibility. For example, he immediately notes that the method is inconsistent with a foundationalist method that singles out some subset of the considered judgments as unique bearers of prima facie justification (2003, p. 151). And earlier in the paper, Scanlon defends the justificatory status of considered moral judgments by appealing to something like his own metaethical quietism (2003, p. 146).

Scanlon's talk of "incorporating" alternatives thus needs to be read modestly. He can incorporate alternatives in the same attenuated sense that a coherentist about epistemic justification can "incorporate" foundationalism, by noting that most of us believe that our perceptual beliefs should be sensitive to our perceptual states (for example). In light of this, coherentism will entail that for most of us, beliefs about what our perceptual states are like will have an important justificatory role. But just as this does nothing to show that there is no alternative to coherentism, Scanlon's point does not show that there is no alternative methodology to reflective equilibrium. In both cases, alternatives will consist in views on which the methodological significance of some feature

is not dependent on the subject's psychology.[24] If this is correct, the idea that the method is inescapable does not survive clarification.

A final attempt to defend (something like) the method of reflective equilibrium claims that it characterizes a hyper-subjective species of rationality. The core idea can be introduced by example. Consider Hapless who believes that suffering does not matter if it occurs on a future Tuesday. We, of course, see the substantive error of her ways. But she does not. She is as confident of this judgment as of any of her others. Arguments that her views rest on objectionably arbitrary distinctions leave her cold. Now consider two things that Hapless might do:

1. Carefully make her set of moral judgments more informed and coherent, by her own lights.
2. Transition – for no rationale that makes any antecedent sense to her – to what we recognize to be a more substantively reasonable set of moral judgments.

The first transition appears characterized by a sort of rationality wholly lacking in the second transition. Hapless, reflecting carefully on these two possible futures can be expected to understand the first transition as an improvement, and the second as an alarming ethical conversion. Some defenders of reflective equilibrium seem to suggest that the method is worthy of allegiance because it describes this especially subjective sort of rationality.[25]

This way of defending the method commits it to the normative significance of such subjective perspectives, which is highly controversial. However, I will grant this controversial assumption for the sake of argument, in order to make two more modest points.

---

[24] Note further that even *if* the method were inescapable in some stronger sense, this would fall short of showing that it was a credible way of forming ethical beliefs. Compare the idea that induction might be epistemically unjustified but psychologically irresistible. See (McPherson and Plunkett *forthcoming*) for discussion of related issues.

[25] I read (DePaul 1998, p. 301ff) as suggesting something like this idea, especially in his emphasis on how real alternatives to the method would require the enquiring subject to submit to an 'alien authority'.

The first is that this strategy only vindicates something that resembles reflective equilibrium for subjects who are relatively normal. To see this, consider Hopeless, who (*pace* the method) confidently treats his moral judgments as most credible when they are unstable and formed in conditions conducive to bias. The hyper-subjectivist will now ask, *against* the proponent of the method: what is Hopeless to do, but privilege those judgments, on pain of (hyper-subjective) irrationality? Similar points apply to people who are certain that any reflective revision of their beliefs will tend to be unreliable and is hence to be avoided (for a relevant historical case, see Holton 1996). Again, hyper-subjective rationality will tell these people to avoid anything resembling the method. The key point here is that hyper-subjective rationality and reflective equilibrium (on any familiar description) easily come apart: all it takes is for someone to strongly endorse doing something directly inconsistent with the method.

My second point is that it is plausible to take normative ethics to be a public theoretical project. When seeking a moral methodology, we are arguably seeking a methodology for a shared enterprise, where that methodology may not always be perfectly acceptable to every potential enquirer. Rawls (to his credit) tends to characterize moral theory and philosophy as public rather than solipsistic projects throughout his work (see especially his (1951) and (1999a), but see his (1999b, p. 508) for a potentially conflicting claim). But it is bizarre to impose the norms of hyper-subjective rationality on any such public theoretical project. Compare the legitimacy of scientific researchers systematically ignoring the work of consistent cranks.

In this section, I have considered three attempts to salvage the method of reflective equilibrium by amending the method, and/or replacing Rawls' own theoretical rationale for the method with some alternative. I have argued that redescribing considered moral judgments and the process of adjustment in bare normative terms leaves the hollow shell of a method, incapable of meeting the **Vindication** or **Adjudication** desiderata. I have also argued that

salient alternatives to Rawls' underlying rationale for deploying the method fail. The failure of Rawls' methodology thus extends beyond his particular characterization, and to salient attempts to defend closely related ideas.

Conclusions

In these conclusions, I do three things. I review the case against the method of reflective equilibrium. I offer a provocative diagnosis of its continuing influence despite its vices. And I close by sketching a framework for making progress in developing a methodology that fulfills the **Vindication** and **Adjudication** desiderata.

The core of this paper argued for three claims about Rawls' reflective equilibrium methodology. First, this methodology fails to provide **Vindication**: it does not address the central challenges to the possibility of successful normative ethical theorizing. The role of moral theory in Rawls' methodology at best permits his methodology to postpone, rather than answer, these challenges. Second, the method fails to provide **Adjudication**: it does not provide a non-trivial means of settling central debates within moral methodology. Rather, the method forces every enquirer to confront those debates piecemeal. Third, Rawlsian methodology fails to be **Plausible**. This is true of the theoretical framework that is supposed to motivate the method: the appeal to ideal psychologies in this framework makes the account less rather than more plausible. It is also true of the method itself: the role of dispositions in the account of the method's starting points and procedures for revision are each hard to swallow in what is intended to be a normative proposal.

Many of these criticisms have been raised before. What accounts for the continued influence of Rawlsian methodology in the face of such criticism? I have two conjectures. The first is that it can be a moving target: presented with a criticism of some aspect of Rawls' methodology, it is easy to imagine jettisoning that aspect, and retaining something that still appears recognizably Rawlsian. I have attempted to address this temptation in this paper by

systematically probing the many weaknesses of the methodology. I have also sought to highlight the poverty of natural attempts to salvage something distinctive without accepting the whole Rawlsian package.

My second conjecture is less charitable. Philosophers working in ethics have been trained to approach ethical questions in a familiar range of distinctive ways. Suppose that one wants to continue doing so, without becoming bogged down in methodological enquiry. It is very convenient to be able to appeal to a method that (a) appears highly credible within the profession, and (b) appears on its face to license continuing doing what one was already doing.

If the method were credible, such allegiance might be helpful. But because it is not, sanguine acceptance of the method has become a barrier to substantive progress in moral methodology, by providing the illusion that we have already achieved significant methodological success. Uncritical allegiance to the method also prevents ethicists from thinking clearly about their own enquiry, what they might reasonably hope to achieve by it, and what credible alternatives there might be to the approaches that they have been trained to use.

Decisively rejecting the Rawlsian method thus clears the way for systematic enquiry into moral methodology: enquiry that aims to develop plausible unified methodological proposals that can **Adjudicate** central controversies within moral methodology, and **Vindicate** such enquiry in the face of skeptical challenge. I will close the paper on a constructive note, with the conjecture that we can hope to make progress in this project by focusing on two related questions:

Q1: What is the nature of ethical thought and talk?
Q2: What is our goal in engaging in ethical enquiry?

The first question is the central question of *metaethics*. Focusing on this question is important because different metaethical views will provide us with different characterizations of the subject-matter that ethical enquiry is attempting to explore. And (to put it crudely) it is of great methodological

consequence whether that subject-matter is best understood by analogy to the subject-matter of sociology, or of mathematics, or of wine-tasting (see my 2012 for more careful discussion).

Suppose that metaethical enquiry suggested that ethical thought and talk is about a certain class of mind-independent facts. This might seem to tell us that we should deploy methods apt for revealing those facts to us. However, this might be too quick. The second question mentioned above is important because many philosophers have taken the practical significance of ethical enquiry to suggest different sorts of goals for this enquiry. For example, some might want to discover ethical principles that have some sort of transcendental validity; others to find principles which we can justify to each other, others still to find principles whose practical implementation can be expected to have morally good effects.[26]

Clear focus on these two questions is important, because it is very hard to make significant arguments in ethics that do not implicitly commit one to ruling out some answers to these questions. On the one hand, recognition of this point is dialectically useful: some ethical theorists talk past each other, in light of presupposing different ranges of tenable answers to these questions. On the other, these questions are pressing for anyone interested in ethical theorizing, because neither debates about the correct metaethic, nor debates about the best goal for ethical theorizing are trivial. In light of this, practicing ethicists should care deeply about how consilient their approaches to ethical theorizing are with the range of significant answers to these two questions.

---

[26] Interest in these goals for *ethical enquiry* is not always clearly distinguished from the suggestion that related phenomenon might play a role in constituting *ethical facts*. For discussion of the transcendental strategy that is sensitive to this issue, see (Shah 2010). The idea that the normative claims that we can justify to each other might be important even if they are not thereby true is most familiar in the context of political philosophy; for an especially clear statement of this idea, see (Estlund 2008, pp. 4-5). For the idea that moral methodology should provide practices of moral justification that satisfy various moral desiderata, see for example (Tobin and Jaggar 2013, §2).

Works Cited

Bechtel, W. (1987) 'Connectionism and the Philosophy of Mind: an Overview' *The Southern Journal of Philosophy* 26 (Supplement), 17–41.

Bonevac, D. (2004) 'Reflection Without Equilibrium' *Journal of Philosophy* 101.7, 363-388.

Boyd, R. (1997) 'How to be a Moral Realist' *Moral Discourse and Practice*. Darwall, S., A. Gibbard, and P. Railton (eds). New York: Oxford University Press. 105-136.

Cappelen, H. (2012) *Philosophy Without Intuitions*. Oxford: Oxford University Press.

Clarke-Doane, J. (*forthcoming*) 'Moral Epistemology: The Mathematics Analogy' *Noûs.*

Copp, D. (2012) 'Experiments, Intuitions, and Methodology in Moral and Political Theory' *Oxford Studies in Metaethics* Vol. 7. R. Shafer-Landau (ed.). Oxford: Oxford University Press, 1-36.

Dancy, J. (1985) 'The Role of Imaginary Cases in Ethics' *Pacific Philosophical Quarterly* 66, 141-153.

DePaul, M. R. (1998) 'Why Bother with Reflective Equilibrium?' *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. M. R. DePaul and W. Ramsay (eds). Lanham, MD: Rowman and Littlefield.293-309.

--- and W. Ramsey, eds. (1998) *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*.

Doris, J. and S. Stich. (2006) 'As a Matter of Fact: Empirical Perspectives on Ethics' *The Oxford Handbook of Contemporary Philosophy*, F. Jackson and M. Smith (eds), Oxford: Oxford University Press, 114-152.

Estlund, D. M. (2009) *Democratic Authority*. Princeton: Princeton University Press.

Greene, J. (2008) 'The Secret Joke of Kant's Soul' *Moral Psychology Vol. 3: The Neuroscience of Morality*. W. Sinnott-Armstrong, (ed.). Cambridge, MA: MIT Press, 35-80.

Hare, R. M. (1981) *Moral Thinking: its Levels, Method, and Point*. Oxford: Clarendon.

Holton, R. (1996) 'Reason, value and the muggletonians' *Australasian Journal of Philosophy* 74.3: 484-487.

Huemer, M. (2008) 'Revisionary Intuitionism' *Social Philosophy and Policy* 25, 368-92.

Joyce, R. (2006) *The Evolution of Morality*. Cambridge, MA: MIT Press.

Kagan, S. (1998) *Normative Ethics*. Boulder, CO: Westview.

---. (2001) 'Thinking About Cases' in *Moral Knowledge*. E. F. Paul, F. D. Miller, and J. Paul, (eds). Cambridge: Cambridge UP, pp. 44-63.

Kamm, F. (2006) *Intricate Ethics*. Oxford: Oxford University Press.

Kelly, T and S. McGrath. (2010) 'Is Reflective Equilibrium Enough?' *Philosophical Perspectives* 24: *Epistemology*, 325-359.

McGrath, S. 'Moral Disagreement and Moral Expertise.' *Oxford Studies in Metaethics* Vol. 3. R. Shafer-Landau (ed.). Oxford: Oxford University Press, 2008. 87-107.

Mackie, J. L. (1977) *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.

McNaughton, D. (2003) 'An Unconnected Heap of Duties?' *Ethical Intuitionism: Re-Evaluations*. P. Stratton-Lake (ed.). Oxford: Oxford University Press.

McPherson, T. (2009) 'Moorean Arguments and Moral Revisionism' *Journal of Ethics and Social Philosophy* 3.2, 1-24.

---. (2012) 'Unifying Moral Methodology.' *Pacific Philosophical Quarterly* 93, 523-549.

---, and Plunkett, D. (*forthcoming*) 'Deliberative Indispensability and Epistemic
    Justification' *Oxford Studies in Metaethics* Vol. 10. R. Shafer-Landau
    (ed.). Oxford: Oxford University Press.

Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Petrinovich, L. and O'Neill, P. (1996) 'Influence of wording and framing effects
    on moral intuitions.' *Ethology and Sociobiology* 17, 145–71.

Railton, P. (2003) *Facts, Values, and Norms*. Cambridge: Cambridge University
    Press.

Rawls, J. (1951) 'Outline of a Decision Procedure for Ethics.' *Philosophical
    Review* 60, 177-97.

---. (1999a) 'The Independence of Moral Theory' *Collected Papers*. S. Freeman
    (ed.). Cambridge, MA: Harvard University Press, 286-302.

---. (1999b) *Theory of Justice* Rev. ed. Cambridge, MA: Belknap.

Rosch, E. (1978) 'Principles of Categorization' *Cognition and Categorization*. E.
    Rosch & B. Lloyd (eds). Hillsdale, NJ: Lawrence Erlbaum Associates, 27–
    48.

Scanlon, T. M. (2003) 'Rawls on Justification.' *The Cambridge Companion to
    Rawls*. Ed. S. Freeman. Cambridge: Cambridge University Press, 139-67.

Schroeter, F. (2004) 'Reflective Equilibrium and Anti-theory' *Noûs* 38.1, 110-134.

Schwitzgebel, E. and Cushman, F. (2012) 'Expertise in Moral Reasoning? Order
    Effects on Moral Judgment in Professional Philosophers and Non-
    Philosophers' *Mind and Language* 27.2, 135-153.

Shafer-Landau, R. (2003) *Moral Realism: A Defence*. Oxford: Clarendon.

Shah, N. (2010) 'The Limits of Normative Detachment' *Proceedings of the
    Aristotelian Society* 110.3, 347-371.

Singer, P. (1972) 'Famine, Affluence, and Morality.' *Philosophy and Public
    Affairs* 1.1, 229-243.

---. (1974) 'Sidgwick and Reflective Equilibrium' *Monist* 58.3, 490-517.

---. (2005) 'Ethics and Intuitions' *Journal of Ethics* 9.3-9.4, 331-52.

Sinnott-Armstrong, W. (2006) *Moral Skepticisms*. New York: Oxford University
        Press.

Smith, M. (1994) *Moral Problem*. Oxford: Blackwell.

Stich, S. and R. Nisbett. (1980) 'Justification and the Psychology of Human
        Reasoning' *Philosophy of Science* 47, 188-202.

Tobin, T. W. and Jaggar, A. M. (2013) 'Naturalizing Moral Justification'
        Metaphilosophy 44.4, 409-39.

Unger, P. (1996) *Living High and Letting Die: Our Illusion of Innocence*. New
        York: Oxford University Press.

Wilkes, K. (1988) *Real People: Personal Identity without Thought Experiments*.
        Oxford: Clarendon.

Williamson, T. (2007) *The Philosophy of Philosophy*. Oxford: Blackwell.

---. (2008) 'Why epistemology can't be operationalized' *Epistemology: New
        Philosophical Essays*. Q. Smith (ed.). Oxford: Oxford University Press.