Mark Schroeder's Hypotheticalism:
Agent-neutrality, Moral Epistemology, and Methodology

Forthcoming in a *Philosophical Studies* symposium on
Mark Schroeder's *Slaves of the Passions*

Tristram McPherson,
University of Minnesota Duluth

tristram@d.umn.edu

Mark Schroeder's *Slaves of the Passions* is the most systematic defense yet constructed of a broadly 'Humean' conception of normative reasons for action, according to which all of one's reasons are ultimately explained by facts about one's psychology. This, however, radically undersells the book, which is breathtaking in its scope, originality, and density of powerful ideas. So: if you care about deep questions in ethics, you should read it. It is also rhetorically Humean, written so engagingly that even those without prior interest in its subject may enjoy thinking through such a superb example of philosophical argument.

Despite admiring much in this book, I will play the usual role of symposium commentator here, and focus on sketching three brief objections to Schroeder's account, in the hopes that they will spur him to develop and clarify his views. §1 argues that Schroeder's crucial account of agent-neutral reasons cannot be made to work. §2 argues that a core element of his distinctive proposal in moral epistemology has awkwardly limited scope. Finally, §3 identifies a tension between Schroeder's view and its central methodological motivation.

1. Agent-neutral reasons.

Schroeder claims that his theory is capable of explaining the existence of *agent-neutral* reasons. On Schroeder's gloss, for a consideration to be an agent-neutral reason to do A is,

roughly, simply for it to be a reason for everyone to do A.[1] This claim is crucial to his project for two reasons. The first reason is dialectical: it avoids the traditional charge that Humeans are unable to accommodate very strong intuitions that certain ordinary moral considerations provide agent-neutral reasons. For example, Schroeder himself suggests that, no matter what your interests, the fact that Katie needs help is a reason to help her (103). Second, agent-neutral reasons play a central role in Schroeder's own account of the weight of reasons, which he takes to be strongly independently motivated. A metaphysically elegant Humean theory of reasons capable of vindicating agent-neutral moral reasons would be impressive. I will argue however, that Schroeder cannot make good on his claim to offer such a theory.

Schroeder calls the version of Humeanism that he defends *Hypotheticalism*. The heart of Hypotheticalism is the following reductive account of reasons:

> **Reason** For *R* to be a reason for *X* to do *A* is for there to be some *p* such that *X* has a desire whose object is *p*, and the truth of *R* is part of what explains why *X*'s doing *A* promotes *p*. (59)

Intuitively, Reason looks like trouble for the existence of agent-neutral reasons. This is because Reason tells us that all of our reasons are grounded in our desires. Thus, all it would take for a proposition *p* to fail to be an agent-neutral reason to *A* on this account is for some possible agent to lack any desire whose promotion by doing *A* would be explained by that proposition.

As Schroeder notes, one might attempt to rebut this worry in at least two ways. First, one might seek to vindicate agent-neutrality by arguing that there is (at least) one desire necessarily shared by all agents. Schroeder rejects this quasi-Kantian approach in favor of a second strategy. Rather than singling out any particular desire as universal, and hence as the ground for agent-neutral reasons, he suggests that if we relax the promotion relation sufficiently, every possible agent will turn out to have some desires or other (not

[1] Schroeder's own formulation (*Slaves of the Passions*. Oxford: Oxford UP, 2007. 18) is more careful, and includes a contextual parameter that permits different strengths of agent-neutrality. I return to this point below. All further references are to *Slaves of the Passions*, unless otherwise noted.

necessarily the same ones) such that the same consideration will turn out to explain why performing a certain action would promote those desires (109).[2]

This strategy requires an *extremely* relaxed interpretation of the promotion relation, and this is just what Schroeder provides. Consider two of his examples. First, he suggests that if you want to get home on time, there is a reason for you to perform any action, provided that "there is *some* chance that [the action] will get you home on time" (112, emphasis his). Call this the *some chance criterion*. Second, he suggests that Mary's *believing that p just in case p* would promote the satisfaction of Mary's desire to have new shoes. His defense of this latter suggestion rests on the idea that being in error about any particular proposition *could*, by a sort of holism about belief formation, entail that Mary fails to get new shoes (114).

The core problem with Schroeder's strategy can be introduced by considering a parallel to this case. When a person has one false belief, it can sometimes be the case that he is more likely to achieve his aims if he has another relevant false belief. For example, suppose that I am looking for my keys, which in fact are in the kitchen. If I falsely believe that Amy has my keys, also possessing the false belief that Amy is in the kitchen could lead to my finding my keys, while possessing the true belief that she is far away would not. Examples like this show that possessing any arbitrary true belief *could*, by the same holism about belief formation that Schroeder appeals to, ultimately entail that I fail to find my keys. Thus, if the promotion relation were as weak as the some chance criterion and the Mary example suggest, *believing that p just in case not-p* would 'promote' the satisfaction of an agent's desires in the same way that Schroeder claims that believing an arbitrary truth would.

This example suggests that Schroeder's some chance criterion threatens to produce *explosion*: the consequence that every consideration gives every agent a reason to perform every act. Consider again the desire to get home on time. For just about anything that I might do – scratching my ear, running the other way, trying to drive my car off a cliff – there is some possible scenario in which this act leads me to get home on time. The problem remains whether we read chances in a subjectivist or objectivist way. If chances

[2] The plurals in this formulation are required in order to accommodate what Schroeder calls the *strong modal status* of agent-neutral reasons (106), which requires that one would still have the reason even if one lost any particular desire.

are just subjective conditional probabilities, a standard view is that we ought never to assign zero conditional probability to propositions other than logical falsehoods. If chances are objective, there are vanishingly tiny chances that for any *X*, my doing *X* will lead me to get home on time by some quantum miracle.

This is not enough to demonstrate explosion, because the claim that every agent has a reason to perform every act is very different from the claim that *every consideration explains why* performing that act promotes some desire of every agent. Schroeder does not tell us how he is thinking of explanation as it figures in Reason, so it may seem that he could resist explosion by filling out this part of his view. However, his view does constrain the relevant conception of explanation: it must be such that (to return to Schroeder's example) *that Katie needs help* is an adequate explanation of why helping Katie promotes at least two of any agent's desires. The case for (something close to) explosion can now be restated as a conjecture about this explanation relation: any explanation relation on which *that Katie needs help* explains why helping Katie promotes (some of) any agent's desires, will also be a conception on which *that Katie needs help* explains why *harming* Katie promotes any agent's desires, and so on.

For Schroeder's view to be viable, there would need to be conceptions of promotion and explanation that meet three conditions. First, these conceptions would together need to be *selective*: generating *some* agent-neutral reasons from a base of arbitrary agents' desires without generating explosion. The initial intuition that the Humean could not explain agent-neutrality stems from the apparent diversity of possible desires. Once we relax the promotion relation, the same diversity generates the intuition that the Humean cannot explain selectivity. Thus, an adequate theory would need to enable us to see *why* the relations are selective.

Second, an adequate theory would make these promotion and explanation relations intelligible, by showing why they represent the correct normative characterization of the relation between considerations and desires. This constraint is difficult to satisfy, because intuitively plausible versions of the promotion relation are either quite restrictive (as Schroeder notes in his critical discussion of alternatives), or threaten explosion, like his own proposal.

Finally, in order to meet Schroeder's goal of allaying worries that Humeanism has counterintuitive implications, the account would need to provide reason to believe that it can vindicate something close to common-sense morality. This task again seems Herculean, because the promotion and explanation relations are *content-neutral* features of the theory. On its face, the hypothesis that relaxing a content-neutral relation will vindicate the agent-neutrality of a *highly content distinctive* class of reasons, like the moral reasons, looks like wishful thinking.

I cannot believe that the hypotheticalist could satisfy these three conditions. However, one of the virtues of this book is Schroeder's practice of noting alternative ways of developing the core Humean idea, and this issue is no exception. Schroeder suggests that we can make the task of explaining agent-neutral reasons easier, by restricting the domain of agents who need to share a reason in order to for it to count as agent-neutral. He offers two salient options. The first is an 'Aristotelian' precisification, on which an agent-neutral reason is one shared by all possible *human* agents (but which perhaps had no grip on an intelligent but diabolical Martian agent). The second is a sort of crude relativism that restricts the domain to *agents around here* (117-121).

Neither of these proposals seems to me to address the explanatory tasks just sketched. On the one hand, these tasks appear only slightly less intractable for the Aristotelian version of the view. On the other, the relativistic restriction looks to me like an under-motivated stopping point, especially given the crucial role that agent-neutral reasons play in Schroeder's account of the weight of reasons. If one is prepared to relativize this much, it is surely more attractive to be serious about one's Humeanism and relativize to each agent. On the latter view, however, the claim to have captured agent-neutrality (within a domain of one!) would be absurd.

I close this section by suggesting a hypothesis. One of the many impressive achievements of this book is the elegant case Schroeder makes for thinking about the weight of reasons in terms of correctness conditions for placing weight, and this in turn in terms of agent-neutral reasons (Ch. 7). This account may provide the beginnings of an attractive unified approach to the weight of reasons with explanatory power in both ethics and epistemology. However, the Humean has few options if she wishes to adopt this account, and if I am right, Schroeder's preferred option cannot work. My hypothesis is that

ultimately one will have to choose: Hypotheticalism *or* agent-neutral reasons, and with the latter an attractive account of the weight of reasons.

2. The epistemology of the normative.

Perhaps surprisingly, Schroeder argues that his reductive theory permits an attractive vindication of a familiar-looking intuitionistic epistemology in ethics. His proposal begins with an account of what it is to have a desire. Roughly, on this account, to have a desire is to have a state which leads one to focus on considerations that (given one's beliefs) will tend to promote the object of one's desire, and to form beliefs that those considerations are reasons (155-8). I will call this the *basic belief mechanism*. The basic belief mechanism being purely causal, the beliefs that result will have an 'intuitive' feel: the mechanism will lead to certain considerations phenomenologically *striking* one as reasons, and one might typically form one's normative beliefs as direct result of such appearances.

Schroeder proposes an account of the epistemology of the normative that takes the perceptual analogy suggested by such 'strikings' very seriously. First, as with perception, one's belief that there being dancing at the party is a reason to go there will standardly be caused by a state (the phenomenological striking) which has the same content as the belief. Second, again as with perception, there is a neat explanation of how there is (in the 'good case') a non-accidental connection between that striking and the state of the world that makes it true. This is because the desire which is partly constituted by the phenomenological striking also plays a crucial role in explaining why the belief is true, on the reductive story Schroeder offers. Finally, Schroeder suggests that, just as ordinary perception immediately generates (defeasible but typically sufficient) reasons to believe things are as they appear to be, so we can hope that something's striking one as a reason can similarly be a basic source of justification (174).

Schroeder finds his moral epistemology appealing because it vindicates the Aristotelian thought that someone who desire the right things to the right degree is thereby able to make correct moral judgments (176).[3] It also has three features that are unusual for

[3] One might think it odd that a Humean could talk about desiring the right things, but on Schroeder's view, the right things to desire would be the things that there is most (right kind of) reason to desire. Provided the right kinds of reason for desiring and action don't come apart (a significant further worry, given how

an epistemology that falls out of a reductive local metaphysics. First, his view is in an important sense epistemically non-revisionary. Second, the epistemic implications of his view fall out of the specific nature of his view, and not some general feature of reductivism. Third, his epistemology nonetheless does not take the form of a method of reliably applying the reductive theory itself in order to generate true specific beliefs.

Very quickly, consider two candidate views that contrast with Schroeder's in these respects. On the one hand, the moral epistemology defended by Richard Boyd is basically just an instance of the epistemology that he takes to be appropriate to investigate broadly social-scientific kinds.[4] It is thus revisionary of standard intuitionism, and falls simply out of the broadly naturalistic character of Boyd's moral metaphysics, not the specific candidate specification of goodness that he offers. On the other, one might develop a Humean 'common point of view' story[5] to provide both a reductive account of the content of moral norms, and a (potentially revisionary) method for discovering that content. Schroeder's account is arguably appealing in part because of the way that it contrasts with such views.

Schroeder's discussion of his account, however, obscures a central explanatory shortcoming. The basic belief mechanism can seemingly only explain one's intuitive beliefs about one's *own* reasons (call these *self-concerning* reasons judgments), and not one's intuitive beliefs about *others'* reasons. The explanation is straightforward: *my* desires are not constituted by the salience of considerations that would explain the satisfaction of *your* desires. The same point applies to Schroeder's 'Aristotelian' thought: even the most virtuous desires will not constitutively generate the 'seeming' that *someone else* has a reason to do something. Further, Schroeder's explanation of the mechanism in the case of self-concerning reasons judgments seems to have no natural extension to the case of other-concerning judgments.

One might take Schroeder's moral epistemology to be attractive even given this incompleteness. However, it seems to me that its incompleteness undermines its plausibility as an explanation of self-concerning reasons judgments. To see this, note that

---

different *desiring* is on his account from *intending*), possessing such desires would thus (by the basic belief mechanism) lead to one making reliable intuitive judgments about what one has most reason to do.

[4] "How to be a Moral Realist." In Stephen Darwall, Allan Gibbard, and Peter Railton, eds. *Moral Discourse and Practice*. New York: Oxford UP, 1997. 105-136.

[5] *A Treatise of Human Nature* Vol. 1. Eds. David Fate Norton and Mary J. Norton. Oxford: Clarendon Press, 2007 (1739), §3.3.1.

self- and other concerning reasons judgments plausibly have identical phenomenology. In the ordinary version of each case, it just strikes one that *R* is a reason for [one, anyone] to *A*. Here it is instructive to contrast cases like perceptual experience or memory that might at first appear analogous to the 'basic belief mechanism' epistemology that Schroeder suggests. My judgments about what I remember are plausibly intuitive in just the sense that Schroeder suggests: I typically move seamlessly from seeming to remember that *P* to the judgment that I remember that *P*. By contrast, I typically need to go through a different and considerably more complicated process in order to come to the conclusion that *you* remember that *P*. Unlike normative judgments, then, self-and other-concerning judgments about memory thus have a strikingly *contrastive* phenomenology. This difference in phenomenological feel seems to me to cast doubt on Schroeder's ability to plausibly appeal to an epistemic parallel to such cases.

One might seek to vindicate this parallel to memory by resisting the claim that other-concerning reasons judgments are intuitive in this way. For example, one might note that if I judge that you have a reason to order the bok choy, I need to at least implicitly appeal to beliefs about your perspective. My beliefs about your tastes in vegetables seemingly play the same role in this case as my beliefs about your psychology that underwrite my judgments concerning whether you remember something.

This point is of limited help, because there are other cases in which we plausibly make intuitive other-concerning normative judgments without appealing to inferences about their psychology. These include paradigmatic moral judgments. For example, suppose again that Katie needs help. I can seemingly engage in just the same intuitive process in making the judgment that you have a reason to help her as I do when judging that I have a reason to help her. It thus appears that, despite their seamless phenomenology, Schroeder's account cannot give a unified explanation of the foundational status of our self- and other-concerning normative judgments.

A normative naturalist has choices in how to spell out her epistemology. On the one hand, she can do so in very general schematic terms (*a la* Boyd), in which case the details of a reduction look to be of dubious relevance to the defensibility of the proposal. On the other, she can identify a more committal epistemic mechanism (*a la* Hume's common point of view). However, such a committal epistemology will only be attractive

given an extensionally committal view of the metaphysics of the normative, which Schroeder judiciously declines to provide. Schroeder attempts to navigate between these extremes: to show that his reductive metaphysics generates a distinctively plausible epistemology, without committing himself to a determinate hypothesis about the extensional content of this metaphysics. Such navigation would be attractive if it succeeded. However, if I am correct to take phenomenology seriously in thinking about this issue, the partial nature of Schroeder's account raises troubling questions as to whether it could be the correct story even about the range of cases where it is strongest.

3. Motivating Hypotheticalism

Slaves of the Passions begins with a motivating idea. Suppose that Ronnie likes to dance and Bradley dislikes dancing. According to Schroeder, intuitively,

(i) that there is dancing at the party is a normative reason for Ronnie to go to the party, but a reason for Bradley to stay away from it, and
(ii) the differing facts about Ronnie's and Bradley's psychologies *explain* their respective reasons in this case.

Call this *Schroeder's datum*. As Schroeder recognizes, only some of our reasons seem intuitively amenable to the type of explanation offered by Schroeder's datum: agent-neutral reasons like the fact that Katie needs help appear difficult to fit into this mold.

Schroeder argues that the best way to make progress in thinking about reasons for action is nonetheless to begin by focusing on cases like Schroeder's datum, and not on cases like Katie's needing help. This is because, according to Schroeder, if you want to know what a reason is, you should begin by focusing on cases of reasons that are reasons for some people and not for others (207). This point is in turn motivated by a general methodological principle that states that, in trying to discover the nature of some kind, you need to compare instances of that kind with foils. For example, if you want to know what an ancestor is, focus on someone who is an ancestor to some but not to others, and try to identify the difference. In what follows, I accept this principle for the sake of argument, and examine whether it really supports the use of Schroeder's datum.

The first problem is that Schroeder's datum appears to be a less specific application of the methodological principle than it could be. As Russ Shafer-Landau argues in this

symposium, there are certain types of desires that we typically think *do not* provide reasons: for example, desires based upon false beliefs. So, a fully general application of the methodological principle would seemingly ask us to include in our initial contrast class the contrast between the sorts of desires that do seem to provide reasons, and those that do not.[6] Thus, beginning with Schroeder's datum appears *ad hoc* relative to the general methodological principle.

Schroeder's use of the datum also appears to conflict with the substantive theory of reasons that he takes it to motivate. This is because, as we saw in §1, Schroeder endorses a *very* relaxed promotion relation. Because of this, if Bradley is anything like a normal human being, he surely has some aim that would on Schroeder's theory be promoted by going to the party, in virtue of the presence of dancing there. This would entail that the fact that there is dancing at the party is a reason for *both* Ronnie and Bradley to go. However, if this is a consequence of Schroeder's theory, then Schroeder's theory entails the falsity of the intuitive presupposition of an existential contrast in reasons posited in Schroeder's datum. This in turn would make Schroeder's case for his theory violate another highly plausible methodological dictum: that a theory cannot be legitimately motivated by a thesis incompatible with the truth of the theory.

Conclusions

In the concluding chapter of *Slaves*, Schroeder sets out what he takes to be the virtues of his Hypotheticalism. Simplifying greatly, the most central of these are:

1. The virtues of the account of weighing reasons
2. That Hypotheticalism offers satisfying accounts of moral motivation and moral epistemology

[6] Unlike many contemporary Humeans, Schroeder declines to add idealizing conditions to his account of reasons in response to such intuitions. His account is thus arguably less responsive to the methodological principle than such accounts. One might think that Schroeder could reply to Shafer-Landau's point by appealing to his attack on *negative existential intuitions*: intuitive judgments that some consideration is *not* a reason for a certain person to perform a certain act. Schroeder offers a fascinating argument that such intuitions are 'not to be trusted' (92-5). However, this is a dubious response for two reasons. First, cases of desires based on false belief can be framed in completely specific terms, and part of Schroeder's attack on negative existential intuitions involves the prediction that these intuitions will weaken when our descriptions of the case become more specific. Second, Schroeder cannot debunk all such intuitions because doing so would undercut his ability to apply his methodological principle. His use of this principle, after all, depends upon our intuitive ability to accurately assess that agents like Bradley *lack* certain reasons.

3. That the argument for it respects the general methodological principle.
4. Virtues that follow from its being a reductive theory

In these comments, I have briefly sketched my doubts about all but the last of these alleged virtues. Against the first, I have argued that there are reasons to think that it is impossible for Hypotheticalism to deliver the agent-neutral reasons required for the account of weighing reasons (§1). Against (part of) the second, I have suggested that Hypotheticalism fails to underwrite a plausibly general account of moral epistemology (§2). Finally, against the third I have argued that Schroeder's attempt to motivate Hypotheticalism by appeal to the general methodological principle both appears intuitively *ad hoc*, and appears inconsistent with central predictions of the theory itself (§3). If anyone is capable of showing that these objections to Hypotheticalism are illusory, it is Schroeder. In the meantime, however, I cannot help but think that he has marshaled one of the most impressive collections of philosophical insights and arguments in recent memory in defense of a view that turns out to be hostile to the very fortifications that he has sought to provide.[7]

[7] I am indebted to Derek Baker, David Plunkett, Alex Silk, and Ralph Wedgwood for conversation related to these comments, and to Janet Levin for inviting me to present at the symposium on *Slaves* at the 2010 Pacific APA. Special thanks are due to Mark himself for helpful discussion, and for being so nice about my seriously misreading one of his views in the APA symposium version of these comments.