On doing better, experimental-style

Jonathan M. Weinberg

Abstract Timothy Williamson devotes significant effort in his *The Philosophy of Philosophy* to arguing against skepticism about judgment. One might think that the recent "experimental philosophy" challenge to the philosophical practice of appealing to intuitions as evidence is a possible target of those arguments. However, this is not so. The structure of that challenge is radically dissimilar from that of traditional skeptical arguments, and the aims of the challenge are entirely congruent with the spirit of methodological improvement that Williamson himself exemplifies in the Afterword of his book.

Keywords Skepticism · Judgment · Intuitions · Experimental philosophy · Philosophical methodology

1 Introduction

Timothy Williamson's *Philosophy of Philosophy* manifests terrific amounts of intellectual energy, and a substantial portion of it is generated by a creative tension between two of his chief declared metaphilosophical tasks. One is his "opposition to philosophical exceptionalism" (p. 6),[1] by articulating a conception of philosophical judgment as continuous with ordinary and scientific cognition. Another is his defense of philosophy as a legitimately armchair discipline, consisting "of thinking, without any special interaction with the world beyond the chair, such as measurement, observation, or experiment would typically involve"

---

[1] All otherwise unreferenced page citations are to Williamson (2007).

J. M. Weinberg (✉)
Department of Philosophy, Indiana University, 1033 E. Third St., Sycamore Hall 026, Bloomington, IN 47405-7005, USA
e-mail: jmweinbe@indiana.edu

(p. 1). For many other philosophers, the restriction to the armchair is a matter of deep commitment to philosophy as an matter of aprioristic, or foundational, or purely linguistic and/or conceptual analysis. But Williamson offers no such in-principle reason for this restriction, and it is hard to see what such reason could be offered, consistent with that first goal. Arguments here thus must turn more on the matter of what happens to work in practice. This immediately suggests the questions: just how well do armchair judgments really work in practice? And how could we tell?

Although Williamson focuses largely on judgments, really it is not a type of mental activity *per se*, but rather a *method*—a set of practices—that is both the positive target of Williamson's arguments, and the negative target of mine and other critics.[2] The powers of human judgment may be fine for some purposes, and not for others, and most interesting methodological questions here can only be framed in terms of what we might want to *do* with these judgments, and in what practices[3] we might look to deploy them.

I will be considering here a kind of challenge to those armchair practices. Let's call it *the experimentalist's challenge*, and those who are making the challenge *experimentalists*, as it is based on a growing body of experimental work that suggests that judgments of the sort that philosophers rely upon so centrally in this practice display a range of inappropriate sensitivities. That is, there is some evidence that the judgments vary systematically with factors that one would not expect to track the relevant philosophical truths. Most of this work can be divided into four categories: demographic differences; order effects; framing effects; and environmental influences. For example, judgments about knowledge, reference, and morality[4] have all been found to differ somewhat depending on whether the agent offering the judgment is of Western or Asians descent, even, in some cases, where both groups are native-English-speaking American college undergraduates (though I agree with Williamson that it is not at all obvious at this time how best to explain this variation (p. 190)). The order in which thought-experiments are considered also seems capable of influencing judgments about morality and knowledge.[5]

Petrinovich and O'Neill (1996) in a study of trolley cases discovered that small differences in wording could exploit framing effects along the lines of those famously studied by Tversky and Kahneman (1981); for one group of participants, the action being considered was described as throwing "the switch which will result in the death of the one innocent person on the side track." For another group, the

---

[2] It follows that I will not be trafficking at all in "the psychologization of evidence" (pp. 211ff). It is fine for my purposes if one considers only the judged propositions, when known, to be evidence. The challenge can be framed in Williamson-friendly terms that many propositions commonly taken to be evidence, in fact are not such.

[3] There are of course many distinctions to be made within these practices of taking intuitions/judgments as evidence for philosophical claims, and many of these issues look different when philosophers aim at something like a conceptual analysis, and when they take themselves to be investigating an extra-mental reality. In the context of this Symposium, given Williamson's views, my arguments can be taken as focusing on the latter sort of practices.

[4] Weinberg et al. (2001); Machery et al. (2004); Doris and Plakias (2008).

[5] Haidt and Baron (1996); Swain et al. (2008).

action was described as throwing "the switch which will result in the five innocent people on the main track being saved." The difference in wording had a significant effect on participants' judgments despite the fact that, in the context of the trolley problem vignette, they are obviously describing the same action.[6]

Perhaps most unexpectedly, in many cases people's judgment are influenced by features of the physical or social situation in which the judgment is elicited. These influences, as with order effects and framing, are typically covert. Those affected usually have no idea that they are being influenced and, short of doing or reading about carefully controlled empirical studies, they have no way of finding out. For example, psychologists[7] asked subjects to make moral judgments on a range of vignettes. Some of the subjects performed the task at a clean and tidy desk. Others did it at a desk arranged to evoke mild feelings of disgust. There was a dried up smoothie and a chewed pen on the desk, and adjacent to the desk was a trash container overflowing with garbage that included a greasy pizza box and dirty-looking tissues. They found that the judgments of the subjects in the gross setting were substantially more severe.

These sorts of empirical findings indicate that armchair practice with thought-experiments may be inappropriately sensitive to a range of factors that are psychologically powerful but philosophically irrelevant. Unwanted variation[8] in any source of evidence presents a *prima facie* challenge to any practice that would deploy it. Once they recognize that a practice faces such a challenge, practitioners have the intellectual obligation to see whether their practice can meet that challenge. Once challenged, practitioners incur an obligation to (i) show that their practice's deliverances are immune to the unwanted noise; (ii) find ways of revising their practice so that it is immune; or (iii) abandon the practice. "Immune" here of course should not be read as requiring anything like infallibility—just a reasonable insulation of the conclusions produced by the practice from the unwanted variation that may afflict its evidential sources.[9]

As a quick illustration, consider human color perception, which displays a number of the sorts of sensitivities discussed above. For some practices, these sensitives just do not pose a threat, because the required discriminations are sufficiently robust; there is little risk of confusing red and green traffic signals, for example, which is presumably part of why those colors are used for such signals. Where more delicate discriminations are required, perhaps for getting subtle readings from a litmus strip, then the practice of using the naked eye by itself can be revised by supplementing it with, say, external aids like a color chart for comparison. (An excellent real-world example here is the adoption of double-blind

---

[6] Baron (1994) and Sunstein (2005) have argued that the distorting influences of framing are widespread in the ethical judgments of philosophers, jurists and ordinary folk. See also Sinnott-Armstrong (2008).

[7] Schnall et al. (2008).

[8] Strictly speaking, unwanted *lack* of variation can be as much of a problem as unwanted variation; a thermometer that always gave the same readings even as the temperature changes is an epistemically unsuccessful thermometer.

[9] One appropriate line of response would be to deny that the variation is *unwanted*, by defending a form of relativism, contextualism, or the like. I will not address such responses here, though see Swain et al. (2008) for a brief discussion.

methods in many of the sciences. Revising experimental practices in this way eliminates unwanted sensitivity to the beliefs and desires of experimenters and their subjects, not by overcoming it where it appears, but by preventing it from creeping into the results in the first place.) And for some purposes, no practice based on human color perception can be made sufficiently trustworthy, and we rely instead on spectrometers or pH meters and the like.

So, if the preliminary empirical findings canvassed above hold up—a nontrivial "if"!—then philosophers' armchair practices face the experimentalist's challenge. The existence of a challenge does not, by itself, tell us whether that challenge will turn out to be already met, or can be met with some appropriate changes to the practice, or whether it will prove fatal to the target practice. Although I think that an argument can ultimately be made that the target philosophical practice is not (currently) trustworthy, I am not looking to press that point here.[10] My objective here is just first to frame the challenge, and then contend that armchair-friendly philosophers like Williamson cannot afford to dismiss it without serious consideration of how, or whether, it can be met.

2 The experimentalist's challenge versus judgment skepticism

For one strategy of dismissal would be to assimilate this challenge to skepticism, in hopes that it can thereby be disregarded as hyperbolic, outside the requirements of responsible cognition. And one important part of Williamson's project in *PoP* is his critique of skepticism about our ordinary capacities for judgment. For example, "judgment skeptics" may look to argue for the metaphysically revisionary claim that there are no such things as mountains. Though Williamson does not directly address experimental philosophy when he offers examples of judgment skeptics, it would be easy for readers familiar with that research to get the impression that it is among his targets. (This impression would also be strengthened by his choice of cultural variation in knowledge judgments as an example of a judgment-skeptical hypothesis (pp. 221–222).) If this impression were correct, then maybe philosophers would have grounds to dismiss the challenge out of hand, or to take it up as a more purely academic endeavor as with other forms of traditional skepticism. Yet this impression is not correct, and contrasting the experimentalist's challenge with judgment skepticism will be useful in illuminating the former.

On Williamson's account, the judgment skeptics deploy skeptical hypotheses of a sort familiar in more traditional forms of skepticism, such as skepticism about the external world, and their strategy is one in which "they try to put defenders of a piece of common sense into the position of arguing for it over the judgment skeptical scenario from a starting point neutral between the two alternatives, just as skeptics about the external world do…" (pp. 222–223). The judgment skeptic's strategy is thus one that starts with a fairly anemic premise, asserting the mere possibility of a certain state of affairs. Indeed, though the judgment skeptics often *style* themselves as scientific, Williamson is right that that predicate is merited with

---

[10] See Weinberg (2007).

scarequotes at best, as "[t]he 'scientific' flavor of their alternative scenario disguises the resemblance to more traditional forms of skepticism" (p. 222). This slender premise is then coupled with another, rather more radical premise about the ensuing dialectic: the premise of "Evidence Neutrality", such that "in a debate over a hypothesis **h**, proponents and opponents of **h** should be able to agree whether some **p** constitutes evidence without first having to settle their differences over **h** itself" (p. 210). As a result, the judgment skeptic's sincere dissent from, say, the ordinary judgment that there are mountains, is enough to neuter the epistemic standing of that ordinary judgment.

Williamson notes that these skeptics, unlike most traditional skeptics, want to claim that their preferred skeptical scenario really obtains—not just that there might not be such things as mountains, but that there really aren't. But this point of contrast plays no role in the skeptical argument itself. Presumably that claim is one that the judgment skeptic wants to defend in part *by means of* their skeptical argument to disable common-sense judgments. Thus it is not a claim that they would be looking to deploy in those arguments themselves. But it does leave them open to the objection that they cannot consistently preserve the pieces of knowledge that they want to, while attacking ordinary judgment in the way they do.

For the skeptics' argument turns out to be problematic, even on their own terms. Judgment skeptics "present their views as superior to 'common sense' judgments in compatibility with the results of the natural sciences. They take for granted that those results have some positive epistemic status. Indeed they often treat them as scientific knowledge. They feel a crisis of confidence in common sense, not in scientific method … (p. 222)". Yet this asymmetry is not one that the skeptics are entitled to: "how can such skeptics prevent their arguments for skepticism from applying as far as the sciences themselves? … Although in practice judgment skeptics are often skeptical about only a few judgments or concepts at a time, the underlying forms of argument are far more general. We may suspect that judgment skepticism is a bomb which, if it detonates properly, will blow up the bombers and those whom they hope to promote together with everyone else" (pp. 223–224).

The judgment skeptics' target is very ambitious, while their initial premise for reaching that target is rather thin. Thus they need the epistemically explosive further premise of Evidence Neutrality. And that premise that they need is the very one that leads to their downfall. So far, I agree with Williamson's reasoning. However, if the same charge of self-detonation can be lodged against the experimentalist's challenge, then philosophers would have good reason to ignore that challenge. So I had better argue that no such charge properly adheres to the experimentalist.

First, although classical skeptical arguments often start from a possibility of error that may be little more than conceived, the experimentalist's challenge rests rather on actual evidence that such errors are a very real, empirically plausible threat. After all, the studies described above have predominantly looked at judgments of a type very similar to the ones typically used by philosophers in the target method. But no real scientific investigation could show that mountains do not exist. ("Just a conspiracy of cartographers, then?", as Stoppard has Guildenstern ask Rosencrantz.) In contrast, human judgment's inappropriate sensitivity to philosophically extraneous factors when considering hypotheticals of a sort typically deployed in

armchair philosophical practice—that's an eminently testable hypothesis, and furthermore one that has been tested, at least a little bit.

That difference would be merely cosmetic, however, if experimentalists were to reach next for a version of Evidence Neutrality as a further premise. But we have no such commitment to starting from a neutral, epistemically-impoverished starting point. Quite the opposite—our strategy is based on a standpoint that is epistemically-enriched compared to the restrictions of the armchair, so why on earth would we want to reduce down to a neutral starting point? Certainly we are not at all suggesting that no armchair judgments can be appealed to during this debate over philosophical methodology. Indeed, we do not even require that the whole class of philosophical judgments in question here be treated as epistemically neutered *during* the debate, even if some experimentalists might think such a devaluing is a possible (or even likely) *outcome* of the debate. Such judgments cannot really do much work on behalf of the challenged practice, precisely because the questions at hand just are not on the whole amenable to armchair treatment. Some empirical claims may be knowable from the armchair. But many are not. And the sorts of claims at issue here—regarding what will or will not make a given practice safe from unwanted variation—are largely in the "not" category. So, while skeptics essentially need to disarm their opponents as soon as battle is joined, experimentalists have no similar need. If our interlocutors want to bring a comfy chair to a lab-bench fight, we say: let them.[11]

Instead of deploying the dialectically (over-)powerful premise of Evidence Neutrality, the experimentalists rest their challenge on a set of substantial, armchair-inaccessible claims about the world. Which practices do or do not face these challenges; of those that do, whether they can or cannot meet such challenges in their current forms; of those that cannot, whether sufficient revisions can be found that would enable them to do so—these are the questions upon which the experimentalist's challenge turns, and they are ones that cannot typically be answered from the armchair. Moreover, they are all questions for which different practices may be shown upon investigation to have different answers. My considered view is that the target philosophical practice is in deep trouble in this regard, whereas scientific practices are generally either unchallenged or have the wherewithal to meet such challenges; and our ordinary epistemic practices are

---

[11] But cannot the defender of the armchair then just *beg off* of the challenge? For example, Westerners and East Asians disagree concerning the Gettier case, for example—so can't we (Westerners) just appeal to our judgments that we (Westerners) are right, and the East Asians are wrong, so there's just no challenge to worry about?

No. First, bootstrapping arguments of this sort have proved generally unpopular with armchair-judging epistemologists, for whom the near-consensus seems to be that the problem of easy "knowledge" is indeed a *problem* (Cohen (2002); Elga (2007) is a nice example of taking this as an established view. I should note that there are at least some key dissenters, e.g., van Cleve (2003), Bergmann (2004).) Nonetheless this is not a route that the defender of the armchair can easily pursue.

Second, even while armchair judgments are *generally* left as admissible as we explore whether and how armchair philosophy might meet this challenge, it is nonetheless probably appropriate to consider the *particular* cases for which there is evidence of inappropriate sensitivity as facing an actual defeater. This defeater may of course potentially be removed by further investigation, but it does stand until such investigations are concluded successfully. So there is no quick-and-dirty evasion of the challenge to be had here. (I am particularly grateful to Adam Leite for discussion of this point.)

something of a mixed bag. But that is not a case that can be made briefly or easily. My purpose here is to indicate that the experimentalist's game is not a skeptical one, so all I need to contend is that one cannot discern from the armchair whether any overgeneralization problem or self-undermining problem applies. Moreover, even if this challenge may impose parallel intellectual obligations on scientists and philosophers alike, the former may well turn out, on investigation, to be in much better shape for meeting that challenge than the latter.

This last point also helps illuminate both the promise and the limitation of some of the protestations Williamson makes in *PoP* in the face of some of the experimentalist evidence. For there is one passage in which Williamson responds to the studies indicating demographic variability in philosophical judgments:

> Most intellectual disciplines have learned to live with significant levels of disagreement between trained practitioners, concerning both theory and observation: philosophy is not as exceptional in this respect as some pretend. Notoriously, eye-witnesses often disagree fundamentally in their descriptions of recent events, but it would be foolish to conclude that perception is not a source of knowledge, or to dismiss all eye-witness reports. To ignore the evidence of thought experiments would be a mistake of the same kind, if not of the same degree. Disagreement can provide a reason to be somewhat more cautious than we might otherwise have been, in our handling of both eye-witness reports and of thought experiments; such caution is commonplace in philosophy. There is no need to be panicked into more extreme reactions. (pp. 191–192)

Now Williamson is certainly right that many forms of inquiry have had to deal with demographic variability and all sorts of other unwanted variation, and that the presence of such is not a reason to throw out the whole form of inquiry. But we cannot stop at that observation. For these other inquirers have had to take the troublesome phenomena *seriously*, and to take positive steps to *learn* how to deal with it, often changing their practices significantly. Learning how to live with problematic sensitivities is not just a matter of putting up with it without changing what one does. Williamson has perhaps underestimated here just how demanding that learning process can be. Fields that have successfully accommodated unexpected sensitivities have done so by making adjustments to their norms of inquiry. Sometimes this has been fairly painless, such as learning to use double-blind methods. Sometimes it is more disruptive, as we see in contemporary legal debates on what actually to *do* to accommodate the findings on the unreliability of eyewitness testimony. But sometimes no changes are available that can save the method, and it must be abandoned, as with the death of introspectionism nearly a century ago.

Learning what will or will not work for a given practice of inquiry is itself something that is a matter of inquiry, and rarely can it be predicted in advance what fruits different methodological measures will yield. Philosophy has just begun to take the initial step of taking these kinds of worries seriously, and only time—and sustained empirical investigation – are likely to reveal what changes, if any, there are that philosophers must yet learn to implement. And that "if any" may cut in two

directions. Further investigation *might* reveal that philosophical practice already possesses the wherewithal to shield its results from unwanted sensitivities.[12] Yet it might also indicate that no set of revisions that yield a recognizably armchair method at the end can achieve such insulation.

3 Should philosophers try to meet the challenge?

At this point, an obvious question to ask is: which philosophical judgments are *likely* to be undermined by such phenomena? And my answer at this point is: we just don't know. I certainly do not claim that *all* of the judgments that have played a role in philosophical discussions are influenced by demography, order, framing, or context. Indeed, some of the research cited above report cases in which demographic or order effects were looked for and *not* found.[13] So it is *possible* that these problems recounted afflict just a modest and unimportant subset of philosophical practice. Even if philosophers cannot simply shrug off our challenge as mere skepticism, nonetheless they may well wonder: can they not just shut their eyes, cross their fingers, and hope for the best?

But if results like those in §1 do hold up, then philosophers cannot presume such a happy ending likely. For it is striking that so far, when researchers have gone looking for trouble concerning such judgments, more often than not they have found it. Indeed, some of those results were found by investigators who *weren't* even looking for them in the first place.[14] Rather than the culmination of years of painstaking labor that finally eked out a few troubling findings, experimentalists have stumbled over them as soon as we thought of stepping through the door marked "Empirical Studies of Philosophical Judgment". Although it is early days yet for this research, there is good reason to think that there is plenty more of philosophical interest for it yet to uncover. The experimentalists do not claim, as Williamson seems to worry they might, that "philosophy can nowhere usefully proceed until the experiments are done" (p. 6). But philosophy will need to take an active interest in the conduct of such experiments, and in their results, and to think carefully through the implications they might have for our armchair practices.

Although experimentalists challenge a practice that Williamson is at pains to defend, there is nonetheless a crucial point of convergence. Methodological worries

---

[12] One popular such claim is that philosophers have a kind of *expertise* that should keep their judgments from displaying these sorts of sensitivities, and Williamson gestures at such a move (p. 191). The truth of such claims is a complicated matter, and there is a large and complex literature on the psychology of expertise that could be brought to bear. But it is perhaps enough here that such a literature *needs* to be brought to bear. One implication of that literature is that our folk theory of expertise is not actually very good at discerning what sorts of training will or will not produce real expertise. Whether or not philosophers are, in the sense appropriate here, experts in their judgments, is just one more philosophically-important question that cannot be addressed well from the armchair. See, e.g., Shanteau (1992) and the papers in Charness et al. (2006).

[13] For example, the "Coinflip" cases in Weinberg et al. (2001) and Swain et al. (2008).

[14] For example, both the Gettier results in Weinberg et al. (2001) and the order effects of Swain et al. (2008) had not been initially sought out, but revealed surprisingly in early pilot studies, then confirmed in further investigation.

do call for methodological reforms, and experimentalists can agree with Williamson's vigorous exhortations in his Afterword, "Must Do Better", about the intellectual obligations of disciplinary improvement. Williamson clearly thinks that some ways that philosophy has recently been done—in particular, some philosophers' under-formalized theorizing about such matters as realism and truth—are ways it just ought not be done. Williamson is keenly attentive to the ways in which current philosophical practices may, through careless imprecision of language and sloppiness with particulars, lead philosophers down blind alleys. We are admonishing philosophers to be no less attentive to the ways in which our philosophical practices are at a real risk of likewise failing through careless invocation of poorly-disguised empirical generalizations and an unnoticed susceptibility to undesirable sensitivities.

The experimentalist's challenge is offered in much the same spirit of Williamson's own warnings of the importance of methodological self-examination:

> Philosophers who refuse to bother about semantics, on the grounds that they want to study the non-linguistic world, not our talk about the world, resemble scientists who refuse to bother about the theory of their instruments, on the grounds that they want to study the world, not our observation of it. Such an attitude may be good enough for amateurs; applied to more advanced inquiries, it produces crude errors. Those metaphysicians who ignore language in order not to project it onto the world are the very ones most likely to fall into just that fallacy, because their carelessness with the structure of the language in which they reason makes them insensitive to subtle differences between valid and invalid reasoning. (pp. 284–285).

But—as Williamson does note—there is more than one way in which philosophy "must do better", and one cannot allow a commitment to being anti-skeptical mask the need for improvement, when and where such a need may manifest. The experimentalist can offer the following line of argument in parallel:

> Philosophers who refuse to bother about the empirically-discoverable workings of our minds, on the grounds that they want to study the extramental world, not our thought or concepts about that world, resemble scientists who refuse to bother about the theory of their instruments on the grounds that they want to study the world, not our observation of it. Such an attitude may be good enough for amateurs; applied to more advanced inquiries, it produces crude errors. Those metaphysicians who ignore the empirical in order to preserve the ideal of methodological self-sufficiency are the very ones most likely to fall into error, because their carelessness of the structure of the human mind with which they reason makes them insensitive to subtle differences between accurate and inaccurate observations.

Given that Williamson is surely right that a "few … errors easily multiply to send inquiry into completely the wrong direction (p. 288), unsystematic guesswork about the empirical should be no more acceptable than unsystematic guesswork about logic. Surely Williamson is not being a judgment skeptic when he tells us the value of being careful about logic and language. And the experimentalist is likewise just

as non-skeptical when she advocates being careful about the psychology of human judgment. Both offer warnings that philosophers would be reckless not to heed—even if it means giving up on a conception of philosophy as self-sufficiently residing in its armchair.

References

Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences, 17*, 1–42.

Bergmann, M. (2004). Epistemic circularity: Malignant and benign. *Philosophy and Phenomenological Research, 69*, 709–727.

Charness, N., Feltovich, P. J., Hoffman, R. R., & Ericsson, K. A. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. Cambridge: Cambridge University Press.

Cohen, S. (2002). Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research, 65*, 309–329.

Doris, J. M., & Plakias, A. (2008). How to argue about disagreement: Evaluative diversity and moral realism. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Volume 2: The cognitive science of morality: Intuition and diversity*. Cambridge, MA: MIT Press.

Elga, A. (2007). Reflection and disagreement. *Noûs, 41*(3), 478–450.

Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology, 26*, 201–218.

Machery, E., Mallon, R., Nichols, S., & Stich, S. (2004). Semantics, cross-cultural style. *Cognition, 92*, B1–B12.

Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology, 17*, 145–171.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096–1109.

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes, 53*, 252–266.

Sinnott-Armstrong, W. (2008). Framing moral intuitions. In W. Sinnott-Armstrong (Ed.), *Moral psychology: The cognitive science of morality*. Cambridge, MA: MIT Press.

Sunstein, C. (2005). Moral heuristics. *Behavioral and Brain Sciences, 28*, 531–542.

Swain, S., Alexander, J., & Weinberg, J. (2008). The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research, 76*, 138–155.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458.

van Cleve, J. (2003). Is knowledge easy—or impossible? Externalism as the only alternative to skepticism. In S. Luper (Ed.), *The skeptics: Contemporary essays* (pp. 45–59). Aldershot, Hampshire: Ashgate.

Weinberg, J. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy, 31*(1), 318–343.

Weinberg, J., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*(1 & 2), 429–460.

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