Philosophical Psychology

ISSN: 0951-5089 (Print) 1465-394X (Online) Journal homepage: http://www.tandfonline.com/loi/cphp20

The experimental critique and philosophical practice

Tinghao Wang

To cite this article: Tinghao Wang (2018) The experimental critique and philosophical practice, Philosophical Psychology, 31:1, 89-109, DOI: 10.1080/09515089.2017.1396310

To link to this article: https://doi.org/10.1080/09515089.2017.1396310

Published online: 20 Nov 2017.

Submit your article to this journal

Article views: 113

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=cphp20

PHILOSOPHICAL PSYCHOLOGY, 2018
VOL. 31, NO. 1, 89–109
https://doi.org/10.1080/09515089.2017.1396310

Routledge
Taylor & Francis Group

The experimental critique and philosophical practice

Tinghao Wang

Department of Philosophy, University of California, San Diego, CA, USA

ABSTRACT
Some experimental philosophers have criticized the standard intuition-based methodology in philosophy. One worry about this criticism is that it is just another version of the general skepticism toward the evidential efficacy of intuition, and is thereby subject to the same difficulties. In response, Weinberg provides a more nuanced version of the criticism by targeting merely the philosophical use of intuition. I contend that, though Weinberg's approach differs from general skepticism about intuition, its focus on philosophical practices gives rise to a new difficulty. Most extant experimental surveys investigate intuitions about particular cases through vignettes giving little contextual information. However, philosophical practices crucially depend on intuitions about general claims and typically provide more contextual background. I argue that, due to these two differences between surveys' and philosophers' appeals to intuition, Weinberg's critique lacks enough support from current experimental data. I conclude that experimental philosophers who engage in the negative program should pay more attention on testing philosophers' use of general intuitions and context-rich intuitions.

ARTICLE HISTORY
Received 9 October 2015
Accepted 21 July 2017

KEYWORDS
Experimental philosophy; intuition; metaphilosophy; philosophical methodology; thought experiments

1. Two interpretations of the experimental critique

A common distinction is made between the "positive program" and the "negative program" in experimental philosophy. According to Alexander and Weinberg (2007), while the negative program "challenges the usefulness" of the current intuition-based philosophical practice, the positive program takes it that "experimental philosophy is (at least an indispensable part of) the proper methodology for this practice" (p. 298). This paper will focus on the negative program, from which traditional philosophical methodology has received much criticism. I will refer to the philosophers engaging in the negative program as "experimentalists" and their criticism as the "experimental critique." Experimentalists conduct psychological research, mainly using survey methods, to investigate people's intuitive responses to thought experiments. Two early well-known findings suggested that

CONTACT Tinghao Wang tiw159@ucsd.edu, shawntinghaowang@gmail.com

© 2017 Informa UK Limited, trading as Taylor & Francis Group

non-Westerners surprisingly do not share Westerners' intuitions about Gettier cases (Weinberg, Nichols, & Stich, 2001) and Kripke's (1980) Gödel case (Machery, Mallon, Nichols, & Stich, 2004).[1] Given that demographic background is plausibly irrelevant to the truth of judgments in those cases, many experimentalists concluded that the intuitive disagreements stem from cultural bias. More recently, experimentalists have performed surveys which show that intuitive judgments vary as a function of other irrelevant factors like the subject's personality (e.g., Feltz & Cokely, 2009), age (e.g., Colaço, Buckwalter, Stich, & Machery, 2014), gender (e.g., Buckwalter & Stich, 2014), and the order in which cases are considered (e.g., Liao, Wiegmann, Alexander, & Vong, 2012; Swain, Alexander, & Weinberg, 2008).

According to experimentalists, the above survey data don't merely suggest that some particular philosophical views (e.g., the view that the Gettier case isn't a case of knowledge) are ill-grounded, but in one way or another present a serious challenge to the standard methodology of philosophy. On what I will call the "skeptical interpretation" of the experimental critique, widespread intuitive disagreements suggest that intuition is too unreliable to be a legitimate source of evidence.[2] It is then inferred that the use of intuition should be completely removed from philosophical practice. According to Liao (2008), for example, experimentalists deny that "there are intuitions to which we can sometimes appeal" (p. 254) and think that "we need to abandon the use of intuition altogether" (p. 256). Also, Alexander and Weinberg (2007) suggest that, for some experimental philosophers, "experimental evidence seems to point to the unsuitability of intuitions to serve as evidence at all" (p. 63). And Chudnoff (2013) claims that one aim of experimental philosophy is to argue for "skepticism about intuition," which is the view that "intuition experiences do not justify us in believing propositions" (p. 98).[3]

Under the skeptical interpretation, the experimentalists' argument is a version of "intuition skepticism," by which I mean any general skepticism about the evidential efficacy of intuitions, including both philosophical and non-philosophical intuitions. Consequently, the experimentalists' argument faces two well-known objections that any version of intuition skepticism must face. First, the argument leads to the unacceptable consequence that even ordinary or everyday intuitions are inadequate as evidence. Consider, for instance, a philosophy student's intuition that her friend in the history department does not know what Bayes' theorem is or one's intuition that "Barack Obama" refers to Barack Obama. As Williamson (2007) argues, such ordinary intuitions are based on our general capacity to apply concepts, and denying their justificatory status would result in an unsustainable form of "judgment skepticism." In other words, the experimental critique is at risk of overgeneralizing. Second, some have argued that arguments for intuition skepticism need to rely on some intuitions in order to justify their use of basic epistemic classifications (Bealer, 1992), reasoning (Bonjour, 1998), and epistemic norms (Pust, 2001). But this would make the skeptical reading of the experimental critique, which rejects the use of any intuition, self-defeating.

In the face of these difficulties, Weinberg (2007) has provided a more intricate interpretation of the challenge from experimentalists—though I will ultimately argue that this interpretation gives rise to a new difficulty. While the skeptical interpretation targets the epistemic status of intuition as a whole, Weinberg attacks the trustworthiness of a particular *practice*: "the current analytic philosophical practice of appealing to intuitions as evidence for philosophical claims" (2007, p. 320). Further, while the skeptical interpretation focuses on the *unreliability* of intuition, Weinberg contends that the real problem is the lack of *corrigibility* (or what he calls "hopefulness") in the philosophical practice of employing intuition.[4] That is to say, philosophers are short of appropriate methods to identify and correct errors in their practices. In particular, experimentalists' survey results indicate that intuitive responses to influential thought experiments can unexpectedly vary depending on irrelevant factors like cultural background. According to Weinberg, this strongly suggests that we overall have little knowledge about when people will agree or disagree in their intuitions; we are thus not in a position to use degree of intersubjective agreement as a guide to detection of errors. After arguing that its errors cannot be properly mitigated by other methods like external corroboration, Weinberg concludes that the philosophical practice of using intuition lacks corrigibility and is thus untrustworthy.[5]

How can Weinberg's approach deals with the two worries about the skeptical interpretation of the experimental critique? Since he does not challenge ordinary uses of intuitions, the over-generalization problem appears to be avoided. Weinberg states that appeals to intuitive judgments in "most ordinary cases that some particular object or event falls under a particular concept" (2007, p. 335) are to a great extent trustworthy. People can often spot and rectify errors in such practices with proper methods, such as appeals to intersubjective agreement. The difference between ordinary uses of intuition and the philosophical use, Weinberg claims, is that philosophers frequently rely on intuitions about hypothetical cases, but usually "set no constraints on how esoteric, unusual, far-fetched, or generally outlandish any given case may be" (2007, p. 321). Weinberg seems to think that, because people are much less susceptible to bias and more prone to agreement on ordinary intuitions than on intuitions concerning far-fetched imaginary scenarios, intersubjective agreement can better mitigate errors in ordinary appeals to intuition than in philosophical appeals.[6]

Weinberg's strategy also seems to manage the self-defeat problem. He takes it that philosophers' employment of intuition should be *substantially restricted*; he aims to exclude the use of a significant number of intuitions, rather than all intuitions, from philosophical practice. Consequently, even if Weinberg's argument needs to depend on some intuitions, it is not self-refuting as long as he does not rule out using the intuitions that he relies on. Indeed, Weinberg admits that his argument depends on two intuitions about epistemic norms. Yet, he claims that his own reliance on intuition, unlike standard philosophical practices of using intuition, is corrigible; the intuition that epistemic norms should "do a good job

of producing true beliefs," for instance, can be calibrated by the observation from the history of science that "truth has played an important role in guiding norm selection" (2007, p. 340).

Though Weinberg's approach seems to avoid intuition skepticism, his focus on practices raises the question of whether the judgments being studied by experimentalists genuinely reflect philosophical practices. I will argue that experimental surveys' use of intuition differs from philosophers' use in the following two respects. First, experimentalists mostly investigate *case intuitions*, that is, intuitions about whether a notion is applicable in a given particular case; but philosophical practices crucially depend on *general intuitions*, that is, intuitions about the truth of a general principle or about a connection between abstract notions.[7] Second, when philosophers do treat case intuitions as evidence, they typically provide more contextual information than what experimentalist surveys have thus far provided. These two differences, as I shall argue, show that Weinberg's argument lacks enough empirical support from existing data. Along the way, I will contrast my position with two recent responses to experimentalists: the response that philosophical evidence involves only trained philosophers' intuitions (e.g., Devitt, 2011; Horvath, 2010; Ludwig, 2007) and the response that philosophers do not rely on any intuitions as evidence at all (e.g., Cappelen, 2012; Deutsch, 2009, 2010, 2015).[8] I stress that my purpose is not to defend "armchair" philosophical methodology. Indeed, if experimentalists design surveys in ways that more accurately represent philosophical practices, they might be able to generate data favoring Weinberg's conclusion. I aim to show only that most current experiments are not conducted in such ways and thereby cannot justify Weinberg's criticism of standard philosophical methodology.

2. Case intuition and general intuition

For a proponent of the skeptical approach, intuitions are treated as a single problematic class, and therefore any study of intuition is potentially relevant. By contrast, Weinberg's version of the argument crucially relies on the assumption that the practices examined in experimental surveys are representative of the philosophical practices involving uses of intuition as evidence. Yet, I suspect that they are fairly unrepresentative. In virtually all their survey designs, experimentalists seek to prompt respondents' intuitions about particular cases.[9] They ask questions like "Does Bob really know that Jill drives an American car, or does he only believe it?" (Weinberg et al., 2001, p. 443) or measure respondents' degree of agreement with statements like "John decided to kill his wife of his own free will" (Feltz & Cokely, 2009, p. 345). Contending that such investigations should motivate dramatic revisions in philosophers' appeals to intuition, Weinberg seems to presume that philosophical practice involves primarily reliance on specific-case intuitions.

A closer look at philosophical practice, however, reveals that the alleged epistemic centrality of case intuitions is an exaggeration. Philosophers often grant

significant evidential roles to *general intuitions*—intuitions about general principles or about connections between abstract philosophical notions. For example, epistemologists almost universally share the intuition that knowledge requires truth, and nearly all proposed theories of knowledge depend on this general intuition. Also, Sosa takes the main support for "formal foundationalism" to be the intuitive plausibility of the idea that "epistemic justification is subject to the supervenience that characterizes normative and evaluative properties generally" (1980, p. 15). In philosophy of mind, Chalmers dismisses "Type-A Materialism" because it makes a "highly counterintuitive claim" (2010, p. 114) that consciousness does not need further explanation once all the functions are explained. And Sommers (2010) highlights the role of the "transfer of non-responsibility principle" in debates concerning free will and moral responsibility. These examples could be easily multiplied. The use of general intuition constitutes an important aspect of philosophers' intuition-based methodology; this aspect, however, has been overlooked in most experimental surveys.

At this stage, Weinberg might retreat and claim that experimentalists aim to criticize only philosophical practices involving *thought experiments*. It might be suggested that, although philosophers frequently rely on general intuitions, their thought-experimental judgments are mainly based on case intuitions. This weaker claim, however, is still dubious. One reason to suspect it comes from recent work by Cappelen (2012) and Deutsch (2009, 2010, 2015). They suggest that thought-experimental judgments are often supported by *philosophical arguments* instead of intuitions. For example, according to Deutsch (2010), Gettier's (1963) judgment that Smith lacks knowledge is primarily based on arguments like the following: Smith does not know because his belief is true due to a lucky coincidence. Both Cappelen and Deutsch go further and claim that, since experimentalists wrongly assume that thought experiments rely on intuitive evidence, experimental surveys have no philosophical significance. I will refer to this line of response to experimentalists as the *anti-Centrality response*.[10] This response, if correct, seems to undercut Weinberg's criticism, but many remain unconvinced by anti-Centrality arguments. In what follows, I will present a more plausible response to Weinberg and contrast it with the anti-Centrality response.

I agree with Cappelen and Deutsch that philosophers usually give strong evidential weight to arguments concerning thought experiments. However, I do not endorse their further view that thought experiments do not rely on *any* intuitions as evidence. Instead, I maintain that the arguments provided in thought experiments are often grounded in general intuitions—they start from intuitively plausible premises about principles or about connections between abstract concepts. As both Ichikawa (2013) and Brogaard (2014) have pointed out, supporters of the anti-Centrality response have difficulty explaining our epistemic access to the premises of philosophical arguments. For example, Gettier's argument mentioned above relies on the following premise: if one's belief is true as a matter of luck, then one does not know. This premise is most naturally understood as being

supported by a general "anti-luck" intuition; other possible interpretations, such as that Gettier stipulates the premise without any evidence or that he relies on further reasons for accepting the premise without stating them, seem uncharitable.[11]

A further reason to believe that the anti-luck intuition has a central epistemic status in the Gettier literature is as follows. Note that epistemologists—even those who defended the standard theory of knowledge as justified true belief before Gettier published his paper—almost unanimously agree that Gettier successfully refuted the JTB theory.[12] This sudden shift of opinion seems abnormal, not least because people commonly have a psychological tendency to disregard or underweight evidence that could disconfirm their old views. Philosophical practices are commonly described as involving a mutual adjustment of theories and data. If case intuitions are central evidence for Gettier judgments, it is hard to explain why nearly every philosopher has chosen to revise the theory rather than to revise judgments about Gettier cases. Intuitions are typically regarded as providing merely prima facie evidence; it is unclear why case intuitions alone are thought to almost uncontroversially override all the theoretical advantages (e.g., simplicity) of the JTB account.

To explain this abnormality, I suggest that philosophers' central justification for their Gettier judgments comes from the general anti-luck intuition, an intuition which epistemologists had been using even before Gettier's paper. Importantly, epistemologists relied on the anti-luck intuition as evidence for the justification condition of the JTB theory. As Pritchard notes, the "standard response" (2012, p. 247) to the question of why knowledge has to be justified is that mere true beliefs can be formed as a result of luck. The JTB theory is bankrupt, not only because of Gettier's counterexamples, but because the counterexamples are novel applications of the anti-luck intuition, which was previously thought of as a main reason for accepting the JTB account. Since the JTB theory is undermined by the very general intuition that was once standardly thought of as supporting it, it is to be expected that epistemologists would suddenly abandon this theory in the face of Gettier's counterexamples.

Another case in point is Lehrer's (2000) "Truetemp" case against externalist theories of knowledge. In this thought experiment, scientists insert into Mr. Truetemp's brain an accurate temperature recording device. The device reliably causes him to possess and accept thoughts about the temperature, but Mr. Truetemp himself knows neither that the device has been inserted nor that his thoughts about the temperature are reliable. Cappelen (2012) observes that Lehrer provides several arguments to support his judgment that Mr. Truetemp does not know what the temperature is. The central idea of those arguments, I think, is that Mr. Truetemp does not know because "the correctness of the thought is opaque to him" (Lehrer, 2000, p. 187). This argument relies on what I will call the "opacity principle" that, if the correctness of a thought is opaque to a person, then she does not know the content of that thought. That is to say, for a thought to count as knowledge, the thinker needs to have at least some rough background information about why her

thought is correct, such as information about the underlying cognitive mechanism or about the truth frequency of her thought processes. In the absence of any such background information, according to Lehrer, Mr. Truetemp fails to know what the temperature is.

Cappelen claims that the Truetemp case is not based on any intuition. Again, I disagree with this claim, for the opacity principle is most obviously understood as expressing a general intuition regarding the nature of knowledge. I take it that Lehrer's main evidence against externalism consists in this general intuition. This can be seen from his reaction to modified accounts of externalism, which exclude the Truetemp case as a case of knowledge. Lehrer responds that "the fundamental difficulty remains" (2000, p. 188) and then turns to another example to illustrate this difficulty. In this example, a person is told that Mr. Haller is in her office, yet she has no idea whether the person telling her this is reliable or not. Lehrer then uses an argument based on the opacity principle to defend his conclusion that the person does not know that Mr. Haller is in her office. However, the lack of knowledge in the Mr. Haller case appears far less intuitive than it is in the Truetemp case. Individuals trust the testimony of strangers all the time, for instance, when they get lost and ask for directions. It thus seems unlikely that most people will deny testimonial knowledge in the Mr. Haller case, at least if they are not already thinking of the general opacity principle. Therefore, any intuitive plausibility of Lehrer's judgment about the Mr. Haller case to a great extent comes from the general intuition. The fact that Lehrer maintains that the fundamental difficulty remains in this example strongly indicates that his argument against externalism relies more on the general intuition about the opacity principle than on specific-case intuitions.

I have argued above that general intuitions play a substantial evidential role in both Gettier cases and the Truetemp case. This constitutes a serious challenge to Weinberg's formulation of the experimental critique, for a series of influential experimental studies on those cases (e.g., Swain et al., 2008; Weinberg, Alexander, Gonnerman, & Reuter, 2012; Weinberg et al., 2001) all focus on intuitions about particular cases while ignoring general intuitions. Similar emphasis on general intuitions can be found in many other philosophical thought experiments, like Block's (1978) China-brain case and Kripke's (1980) cases against descriptivism. In general, experimentalists have exaggerated case intuitions' role in philosophers' use of intuitions. Their experimental surveys generally target unrepresentative samples of intuitions; thus, they have in effect been testing a different sort of practice than the one that occurs in philosophy. Note that this is a more serious problem for Weinberg's account than for the skeptical account. A skeptic about intuition regards all intuitions as a class and rejects them wholesale; as a result, she can appeal to any study of intuition for support. By contrast, Weinberg tries to reject only philosophical uses of intuition; therefore, only studies of intuition that resemble philosophical practices closely can lend support to his critique.

In response to this challenge, experimentalists might adopt one of the following four strategies. First, they might claim that case intuitions and general intuitions are generated by the same mental capacity and are thus likely to be subject to the same biases. As a result, it might be suggested, survey data indicating biases in case intuitions are indirect evidence for biases in general intuitions. However, the "homogeneity" assumption that the same psychological mechanism is responsible for both case intuitions and general intuitions is a substantial empirical hypothesis—one which experimentalists have gathered little evidence for. In fact, current psychological evidence seems to count against the homogeneity assumption. For example, Schwitzgebel and Cushman (2012) report that, while the order of presentation influences non-philosophers' intuitions concerning particular moral scenarios, it shows no statistically significant impact on their intuitions concerning moral principles. Further, Nado presents a wide range of psychological evidence casting doubt on the homogeneity thesis, contending that intuitions are "generated by several fundamentally different sorts of mental processes" (2014, p. 15).[13]

Experimentalists might appeal to a second strategy at this stage. They might concede that intuitions are generated by a variety of distinct psychological processes. However, it might be suggested, few psychological models would treat case intuition and general intuition as so wildly different that factors influencing one only fairly rarely influence the other. Current experimental data concerning case intuitions thus constitute indirect evidence for variation in general intuitions, even if the homogeneity hypothesis is false.[14] As a response, I suspect that the above objection is based on a misunderstanding of what experimentalists need to show. Weinberg's conclusion isn't that some general intuitions are variable to a worrisome degree; rather, he needs to show that *the general intuitions used by philosophers* are variable to a worrisome degree. Indeed, he might grant that general intuition as a kind is, though sometimes variable, a generally trustworthy source of evidence. But he would contend that those general intuitions that philosophers frequently appeal to are not trustworthy. As I have argued, the paradigmatic case intuitions studied by experimentalists so far—such as the Gettier intuition and the Truetemp intuition—are treated as a weaker source of evidence than the relevant general intuitions. Given this context, experimentalists' earlier response is in effect claiming that variation in what philosophers treat as weaker evidence (e.g., the Gettier intuition) constitutes indirect evidence for variation in what philosophers treat as stronger evidence (e.g., the general anti-luck intuition). But it remains unclear how the former is indirect evidence for the latter, unless we read "indirect" in a way that makes the available evidence too weak to give enough support to Weinberg's conclusion.

The reason is as follows. Generally speaking, in order to establish a strong criticism of the methodology of some discipline, it doesn't suffice to reveal problems with a method that scholars in this discipline treat as providing merely weak support for their theories and hypotheses. Rather, one needs to defend the existence of problems with what scholars in the discipline treat as strong evidence.

It is not enough to present problems with the weaker method and claim that they constitute indirect evidence for similar problems in the stronger method, especially when the stronger method is employed precisely with the purpose of compensating the insufficiency of the weaker method. Even if there is "indirect" evidence for potential problems with the stronger method in a fairly weak sense of "indirect," the available evidence won't be enough to propose a methodological objection to this discipline. Philosophers, even without engaging in psychological experiments, might agree and expect that their weaker evidence is variable to a worrisome degree; indeed, this is exactly why they find it necessary to present stronger evidence to compensate this worry. Therefore, variation in case intuitions alone doesn't suffice to establish a refutation of traditional philosophical methodology.

Further, it is unclear how experimentalists can establish a plausible objection against traditional philosophical methodology by making discoveries that can be well expected by traditional philosophers. An example is Block's (1978) China-brain thought experiment, in which a billion people in China communicate with one another in ways that are functionally equivalent to a human mind. After appealing to the case intuition that the China-brain lacks mentality, Block notes that this provides merely prima facie doubt against functionalism. He emphasizes that he will "not rest on this appeal to intuition" because this kind of reliance on case intuition is "notoriously fallible" and "far from decisive" (1978, p. 278). He then argues that the content of the intuition has "a rational basis" and that this basis provides "a good reason for doubting that Functionalism is true" (1978, p. 278). This rational basis rests on the general intuition that "mentality depends crucially on psychological and/or neurophysiological processes and structures" (1978, p. 282). Block would agree and expect that the China-brain case intuition is, when considered alone, variable to a worrisome degree. However, the general intuition about mentality both functions as Block's primary evidence and strengthens the evidential status of the case intuition by giving it an underlying general rationale. For Block, it is to be expected that there is significant variation in the China-brain case intuition, and the role of the relevant general intuition is exactly to address this worry. It is thus unclear how experimentalists can propose a plausible objection to Block's argument by merely presenting psychological biases that Block himself would expect to occur.

The third strategy that experimentalists might appeal to involves the claim that, even if general intuitions play crucial evidential roles in philosophical practice, case intuitions surely play *some* evidential role. They might maintain that current experimental results should at least lead to substantial revisions in philosophical practices involving specific-case intuitions. In response, I agree that case intuitions at least sometimes play important evidential roles in philosophy. However, the above response seriously limits the scope and the methodological significance of the experimental critique. Further, I will argue in the next section that experimentalists' data also fail to cast doubt on philosophical practices involving case

intuitions. If my arguments there are successful, then the above strategy won't be sufficient to defend even a seriously limited form of the experimental critique.

Given the problems with the above three strategies, experimentalists might adopt a fourth strategy in response to the above challenge. They might concede that current psychological results don't suffice to establish Weinberg's conclusion that philosophers' use of intuition is untrustworthy. However, they might treat Weinberg's conclusion as a promising hypothesis, and attempt to test it by designing more experiments relevant to the general intuitions that philosophers frequently rely on. I regard this as the best approach in responding to my challenge above. Indeed, there have been a few experimental studies on general philosophical intuitions, such as Nichols and Knobe (2007), Nahmias et al. (2007), and Schwitzgebel and Cushman (2012). Unfortunately, these studies haven't provided sufficient evidence for Weinberg's version of the experimental critique. The first two studies concern abstract intuitions about free will and moral responsibility, but they make conflicting conclusions regarding whether the relevant abstract intuitions are trustworthy; it is thus too hasty to draw any methodological lesson from those studies. And Schwitzgebel and Cushman (2012) didn't discover any statistically significant variation in general intuitions of non-philosophers.[15] Nevertheless, these are just a few experiments, and extensive bias in general intuitions might be found in other philosophical practices or by further investigation.

Admittedly, empirical tests on general intuitions might be more difficult to design than tests on case intuitions. This is partially because ordinary people might find it hard to evaluate certain philosophical principles if they are presented in the ways typically presented in academic contexts. For example, ordinary people might have no idea how to assess the double effect doctrine that "it is worse to harm a person as a means of saving others than to harm a person as a side-effect of saving others" (Schwitzgebel & Cushman, 2012, p. 137). Thus, one might need paraphrase or simplify such principles, but it can be challenging to paraphrase in a way that both (i) preserves the original meaning of the principle, and (ii) remains neutral on different philosophical theories concerning the principle.[16] Schwitzgebel and Cushman attempt to make such a modification to the double effect doctrine by the prompt as follows:

> Sometimes it is necessary to use one person's death as a means to saving several more people—killing one helps you accomplish the goal of saving several. Other times one person's death is a side-effect of saving several more people—the goal of saving several unavoidably ends up killing one as a consequence. Is the first morally better, worse, or the same as the second? (2012, p. 139)

Though one might question the accuracy of the above paraphrase, it serves as an exemplar of what tests on general intuitions might be like. Indeed, experimental philosophy can help enrich philosophical practice by exploring better ways to describe philosophically relevant principles to ordinary persons in ways that they could easily understand and evaluate. The experimentalists' project is a perfectly

legitimate one, as long as they don't take themselves as already establishing a strong criticism against standard philosophical methodology.

3. Context-poor intuition and context-rich intuition

In the last section, I have argued that experimentalists overestimate the epistemic importance of specific-case intuitions. In spite of this, I think case intuition still plays some evidential role in philosophical practice. As I've said in the last section, at this point, Weinberg might maintain that current experimental results should at least lead to substantial revisions in philosophical practices involving specific-case intuitions. However, there is still a key difference between philosophers' and surveys' use of case intuitions: philosophical texts usually provide more contextual information than experimental surveys do. Throughout this section, I will limit my discussion to intuitions about particular cases.

To get clear on my proposal, it is useful to compare it with a recently popular "expertise defense" of standard philosophical methodology. This defense says that philosophers only treat as evidence intuitions of those who possess a certain degree of philosophical training or expertise. Proponents of this approach have postulated the existence of different types of philosophical expertise, such as being better at understanding and interpreting descriptions of scenarios (Horvath, 2010), being better at making judgments based solely on conceptual competence (Kauppinen, 2007; Ludwig, 2007), or an expertise analogous to expertise in other disciplines like physics and psychology (Devitt, 2011; Williamson, 2011). They speculate that such expertise can reduce or eliminate the cognitive biases found in experimentalists' surveys, most of which investigate intuitive judgments of non-experts.

In order to test this speculation about philosophical expertise, experimental philosophers have conducted a number of surveys directly on philosophers. Contrary to what the expertise defense predicts, philosophers are found to be no less susceptible to intuitive bias than laypeople. Their so-called "expert" intuitions vary dramatically according to factors like personality (Schulz, Cokely, & Feltz, 2011), order (Schwitzgebel & Cushman, 2012), and the subject's linguistic background (Vaesen, Peterson, & Van Bezooijen, 2013). Though such results are far from decisive, they do seem to constitute a strong challenge to proponents of the expertise defense, who have provided little experimental data in support of their thesis.

Supporters of the expertise defense allege that experimentalists' early surveys are mistaken in what *individuals* they should test. By contrast, I think the more serious problem for most surveys, whether of laypeople or of professional philosophers, is that they are mistaken in what *practices* they should test. Note that this is a problem especially pressing for Weinberg's account of the experimental critique, which is aimed at challenging the philosophical practice of using intuition, instead of intuitions themselves. Weinberg needs to show not only that

intuitions are unstable, but that they are unstable in practices similar enough to philosophers' appeals to intuition. However, current experimental surveys mostly generate intuitions through vignettes that provide little contextual information, or in a "context-poor" way; thus, they are not testing the philosophical employment of case intuition, which typically involves more substantial contextual information, or is "context-rich."

By contextual information, I mean the information that philosophers provide in the text surrounding a thought experiment.[17] For instance, this can involve explicitly calling attention to a particular aspect of the scenario, reasoning about the relevant thought-experimental judgment, or comparing different cases. Such information can perform many functions; for instance, Cullen (2010) has argued that surveys' lack of explicit "conversational contexts" can lead to misunderstanding of both the vignette and the question.[18] In this section, I will focus on another function of contextual information that will be particularly important for the purpose of evaluating Weinberg's argument: contextual information frequently highlights the ordinary aspects of a far-fetched imaginary scenario.[19]

Take Lehrer's Truetemp case as an example. Though this thought experiment primarily relies on a general intuition, the independent fact that the particular case intuition is in agreement with Lehrer's views provides some extra evidential support. In eliciting this case intuition, Lehrer explicitly draws readers' attention to the fact that Mr. Truetemp "has no idea whether he or his thoughts about the temperature are reliable" (2000, p. 187). The reasoning from the opacity principle highlights the same fact: one main reason to believe that the correctness of the thought is opaque to Mr. Truetemp is exactly that he does not know whether his thoughts are reliable or trustworthy. This aspect of the case is made still more evident in comparisons between different cases. For instance, a later appearance of the Truetemp case in Lehrer's book is immediately followed by an ordinary case, in which one reads an accurate thermometer at a gas station but has no idea whether the thermometer is trustworthy or not.[20] Surveys on this thought experiment, however, do not stress the above fact. Take Swain and colleagues' (2008) Charles case, which is modeled after the original Truetemp case, as an example. While the vignette provides information like that "Charles is unaware that his brain has been altered" and that "apart from his estimation, he has no other reasons to think it is 71 degrees" (Swain et al., 2008, p. 154), it does not emphasize the specific fact that Charles has no idea about the reliability or trustworthiness of his thoughts. This aspect of the case is neither explicitly called to attention nor highlighted by reasoning or comparison between cases.[21]

The aspect that Lehrer emphasizes is ordinary, in the sense that it is more familiar to most people than the Truetemp case itself. While no one encounters anyone exactly like Mr. Truetemp in real life, most people frequently experience cases where one believes a certain proposition without knowing the reliability of one's source. For instance, readers of tabloid newspapers frequently believe what is said in the papers without giving a thought to the publication's reliability.

Lehrer's ordinary case about the thermometer is another example.[22] Since most important philosophical cases are unusual to a certain extent, it is to be expected that philosophers frequently provide contextual information to lay stress on more usual features of the esoteric scenarios. Given that most experimental surveys conducted thus far have largely left out such contextual information, it is unclear how their results can have a bearing on the standard philosophical methodology.

One might reply that variation in context-poor intuitions provides indirect evidence for similar variation in context-rich intuitions. For, it might be suggested, there is no good reason to think that contextual information can sufficiently reduce the degree of variation found in context-poor intuitions. In response, I agree one cannot assume that, generally speaking, contextual information helps reduce cognitive bias. However, this is not the assumption that my objection to Weinberg is based on. My assumption is rather that philosophers' use of contextual information in their appeals to intuitions helps reduce cognitive bias to the degree that can secure the trustworthiness of traditional philosophical methodology. The assumption ought to be a default one for reasons similar to the ones I presented in the previous section. In discussing the methodology of any discipline, the default position ought to be that what scholars in this discipline treat as stronger evidence is indeed stronger and more trustworthy than what they treat as weaker evidence, especially when the stronger evidence is presented precisely with the purpose of addressing the problems with the weaker evidence. A strong criticism of the methodology of this discipline needs to generate sufficient evidence against this default position. In philosophers' use of case intuitions, they treat context-rich intuitions as stronger evidence than the corresponding context-poor intuitions. Thus, experimentalists need arguments showing that the particular ways that philosophers use contextual information fail to sufficiently strengthen their methodology. This can be seen more clearly by the following analogy. According to Weinberg, the experimentalists' argument wouldn't be convincing if they merely presented data about intuitive bias generally; he thinks that experimentalists need to find variation in the intuitions that philosophers use. Similarly, he shouldn't be satisfied with reasons or empirical data against the general claim that contextual evidence reduces bias; rather, he should think experimentalists need specific data showing that the contextual information that philosophers use fails to sufficiently reduce intuition variation. But experimentalists haven't presented enough empirical data indicating this failure.

Further, it is important to note that the earlier response to my objection is not one that Weinberg can appeal to. Remember that, in order to avoid the over-generalization problem, Weinberg claims that ordinary uses of intuition are more corrigible and trustworthy than the philosophical use. Since one main function of contextual information is to make salient the ordinary sides of thought experiments, philosophers' appeals to intuition turn out to be more "ordinary" than Weinberg supposes. As a result, by Weinberg's own criteria, there is good reason

to think that philosophers' appeals to intuition are more trustworthy than the context-poor uses of intuition that experimentalists have tested. For example, consider Lehrer's ordinary thermometer case, where a person reads an accurate thermometer at a gas station without giving a thought to the reliability of the thermometer. We might in everyday life make intuitive judgments about such a person, and our practices would surely be an "ordinary" use of intuition. However, it is unclear on what grounds Weinberg can maintain that this practice is trustworthy but Lehrer's use of the Truetemp intuition is not, given that Lehrer takes both cases as having the same philosophically significant aspects. The more clearly we see philosophers' extensive use of rich contexts in generating intuitions, the less clear Weinberg's distinction between "ordinary" and "philosophical" usage of intuitions will be.

Finally, recent attempts to investigate philosophers' use of context-rich intuition don't appear to support Weinberg's conclusion. A few experimental studies (e.g., Nagel et al., 2013; Starmans & Friedman, 2012; Turri, 2013) have started to test more context-rich intuitions about Gettier cases. Importantly, these studies highlight the element of luck in Gettier cases. Yet, none of them report any statistically significant variation as a function of factors such as age, gender, and ethnic background; on the contrary, subjects are shown to have a broad consensus on knowledge attribution.[23] Still, I think these studies provide good exemplars for what an empirical project regarding context-rich intuitions should be like. For example, Starmans and Friedman (2012) describe to subjects the following Gettier-style scenario:

> Peter is in his locked apartment reading, and is about to have a shower. He puts his book down on the coffee table, and takes off his black plastic watch and leaves it on the coffee table. Then he goes into the bathroom. As Peter's shower begins, a burglar silently breaks into the apartment. The burglar takes Peter's black plastic watch, replaces it with an identical black plastic watch [a dollar bill], and then leaves. Peter is still in the shower, and did not hear anything. (p. 274)

They then present a series of comprehension questions immediately following the above prompt:

(1) Is there a watch [book] on the table? (Yes/No)
(2) How did the watch [book] get on the table? (Peter put it there/The burglar put it there)
(3) Would Peter say there is a watch [book] on the table? (Yes/No)
(4) Why would Peter say there is a watch [book] on the table? (Because Peter put a watch [book] on the table/Because a burglar put a watch [book] on the table)
(5) Peter ____ that there is a watch on the table. (Really knows/Only thinks)
(6) How confident are you about your answer to Question 5 (above)? (1 – Not at all confident to 10 – Completely confident) (Starmans & Friedman, 2012, p. 274)

One purpose of these comprehension questions, according to Starmans and Friedman, is exactly to "closely approximate the logical steps that philosophers have deemed relevant for the attribution of knowledge" (2012, pp. 276–277). Though these questions are presented after the description of the case, they nonetheless help enrich the context by emphasizing (i) the causal disconnection between what justifies Peter's belief and what makes his belief true, and (ii) the element of luck in the formulation of Peter's true belief. This is indeed similar to how philosophers provide contexts in their practices. They usually first briefly describe the case at issue, and then stress the features of the case that they regard as important.

As another example, Turri (2013) attempts to highlight the same elements in the Gettier case by dividing the story into three stages. He first presents to subjects paragraph I, which, when read alone, describes a standard non-Gettier case of knowledge where the protagonist has a justified true belief. Subjects are then presented paragraph II, which adds an element of bad luck to paragraph I, so that the protagonist's justified belief would normally be false. Finally, Turri presents paragraph III, which adds an element of good luck to paragraphs I and II so that the justified belief turns out to be true anyway. The motivation for this three-step method of eliciting intuition is to "dramatize the tripartite structure of Gettier cases, guiding participants to notice the intersection of evidence, truth, and luck, and highlighting that the bad luck's source differs conspicuously from the good luck's source" (Turri, 2013, p. 5). The above two studies thus provide us at least two methods of empirically testing philosophical practices involving context-rich intuitions: to present a series of comprehension questions regarding philosophically important aspects of the relevant case, and to explicitly separate the relevant case into several parts.

Experimentalists might suggest that there is an additional method to test context-rich intuitions: to examine trained philosophers' intuitions. Assuming that philosophers are those most familiar with the contexts of philosophically relevant scenarios, one might think, variation in their intuitions would be strong evidence for variation in context-rich intuitions. It might be added that, since we already have some empirical data about variation in philosophers' intuitions (e.g., Schulz et al., 2011; Schwitzgebel & Cushman, 2012; Vaesen et al., 2013), they constitute strong evidence that philosophical practices involving context-rich intuitions are not trustworthy.[24] However, I don't think this line of argument provides enough support for Weinberg's conclusion. I agree that testing trained philosophers' intuitions can be one indirect way to test context-rich intuitions, as long as (i) the experiments use the cases that philosophers frequently appeal to in their practices, and (ii) the philosophers tested are familiar with these particular cases. But I suspect that (i) isn't true of many of the above experiments; they fail to test the intuitions that philosophers actually appeal to. For example, Schulz and colleagues (2011) test philosophers' intuitions about the "compatibility question"—whether a person, as described in a particular scenario, can be free and morally responsible

in a deterministic world. However, as Sommers (2010) argues, this is the wrong question to ask; though philosophers engaging in the compatibility question do use some key intuitions about both principles and cases, they don't rely on the intuition about the compatibility question itself. Further, instead of eliciting intuitions frequently discussed in philosophical practice, Vaesen and colleagues (2013) elicit philosophers' intuitions about a series of new cases concerning knowledge. The above experiments are designed in order to show whether philosophers' moral intuitions or epistemic intuitions *in general* are more stable than non-philosophers'. This explains why they don't attempt to test the intuitions that philosophers use. But if one wants to test my hypothesis regarding the stability of those context-rich intuitions used in philosophical practice, then one should examine the intuitions that actually play important evidential roles in philosophers' theorization and argumentation.

4. Conclusion

The skeptical interpretation of the experimental critique targets intuition as a general kind. In a sense, any particular finding of intuitive bias adds some degree of confirmation to this criticism, though we have some reasons to think that its conclusion is ultimately too strong to be successful. By contrast, Weinberg's interpretation has a more modest conclusion, for he criticizes only the philosophical employment of intuition. However, Weinberg's approach faces a different problem: his conclusion cannot be justified by extant survey data. Experimentalists' surveys generally test context-poor intuitions about particular cases, but philosophers more frequently appeal to general intuitions and context-rich intuitions as evidence. Though the surveys might show that intuitions are sometimes variable, they have not shown that philosophers' appeals to intuitions are problematically sensitive. Still, I think Weinberg's interpretation outlines a more promising empirical project than the skeptical interpretation. The message from this paper is not that we should disregard the negative program of experimental philosophy altogether, but that those who work on this program must modify their research methodology. They need to more carefully examine how philosophers actually use intuition in the evaluation of theoretical claims. More specifically, they should pay more attention on testing philosophical practices involving general intuitions and context-rich intuitions.

Notes

1. Weinberg and colleagues' finding fails to be replicated by more recent experimental work (e.g., Kim & Yuan, 2015; Machery et al., 2015; Nagel, Juan, & Mar, 2013; Seyedsayamdost, 2015). By contrast, cross-cultural variation of intuitions in the Gödel case has been more robustly replicated (see, e.g., Machery, Sytsma, & Deutsch, 2015; Machery et al., 2010; Sytsma, Livengood, Sato, & Oguchi, 2015).

2. Note that one can have a skeptical stance on intuition that does not invoke unreliability. For example, Cummins (1998) defends a general skepticism toward intuition by arguing that intuition cannot both have independent calibrations and remain useful.
3. It is worth noting that, though the skeptical interpretation is commonly made, it gains little support from textual evidence. Experimentalists rarely explicitly endorse a complete rejection of intuition. Rather, they usually overtly challenge merely a specific kind of intuition, such as epistemic intuition or semantic intuition. That being said, some early work by experimentalists (e.g., Machery et al., 2004; Weinberg et al., 2001) might be easily read as tacitly suggesting a rejection of intuition across the board. In any case, the skeptical reading has become an influential one, especially among critics of experimental philosophy.
4. In another paper, Alexander and Weinberg (2014) suggest that there is an ambiguity about "reliable": it can be used as a synonym for either "trustworthy" or "highly predictable." With this more precise terminology, one can say that the skeptical interpretation invokes the sense of reliability as being highly predictable, while Weinberg's (2007) approach concerns the trustworthy sense of reliability.
5. In some sense of the term, Weinberg might be said to be "skeptical" of the current intuition-based philosophical practice. But his approach still differs from the skeptical interpretation of the experimental critique, because I've been using the "skeptical interpretation" to refer to the particular view that challenges the evidential efficacy of intuition and rejects the use of intuition across the board.
6. Williamson (2007) argues that, since thought-experimental judgments are based on the same capacities as other forms of counterfactual thinking, general skepticism toward thought-experimental judgments will lead to skepticism about many other judgments made in philosophical and ordinary discourse as well. However, Weinberg's view is not a general skepticism about thought-experimental judgments. Rather, it is only aimed at the particular thought-experimental judgments that philosophers rely on. As Machery (2011) contends, even if thought-experimental judgments in general are based on fairly reliable cognitive capacities, there might still be reasons to think that philosophers' use of thought experiments are beyond the proper domain of those capacities.
7. See Pust (2000, p. 32) for a similar distinction. There are multiple ways to spell out this rough classification between the two kinds of intuition. For example, one might distinguish them in terms of the generality of the propositional content of an intuition or in terms of whether an intuition concerns concrete or abstract matters. I will focus on paradigmatic cases of case intuition (e.g., the response to a particular Gettier case) and general intuition (e.g., the endorsement of anti-luck for knowledge); so far as I can tell, nothing in this paper hangs on how the distinction is exactly made.
8. In this paper, I will stay neutral on the issue of what intuitions are. Some philosophers have drawn a distinction between "thin" and "thick" notions of intuition. According to Weinberg and Alexander (2014), for instance, the "thin" conceptions "identify intuitions as merely instances of some fairly generic and epistemological uncontroversial category of mental states or episodes," while the "thick" conceptions "add to this thin base certain semantic, phenomenological, etiological, or methodological conditions" (p. 189). This paper doesn't attempt to decide between these two different kinds of theories of intuition.
9. Schwitzgebel and Cushman's (2012) study is an exception, which elicits participants' judgments about abstract moral principles. Yet, even in this study, specific-case judgments are still the main focus; it includes 17 questions concerning particular scenarios but only five questions concerning moral principles. There are a few more

investigations on general intuitions in experimental philosophy, such as Nichols and Knobe (2007) and Nahmias, Coates, and Kvaran (2007); nonetheless, these studies are aimed at neither discovering intuitive bias nor developing the experimental critique.

10. I borrow the term "Centrality" from Cappelen, who defines it as the claim that "contemporary analytic philosophers rely on intuitions as evidence (or a source of evidence) for philosophical theories" (2012, p. 3).
11. An anonymous referee suggests that, in order to reach the conclusion that Smith lacks knowledge, one needs to rely on the case intuition that Smith's belief is true as a matter of luck. I take it that this case intuition supports the minor premise in Gettier's argument, while it is the general anti-luck intuition that supports his major premise. This is consistent with my thesis that general intuition plays a more important evidential role than case intuition in the Gettier case. It is easy to infer from Gettier's description of the scenario to the minor premise that Smith's belief is due to luck; indeed, one might take this premise to be just part of the description of the scenario, given that Gettier writes that Smith's belief is true "by the sheerest coincidence" (1963, p. 123).
12. Weatherson (2003) is a notable exception.
13. Though my main target here is Weinberg's interpretation of the experimentalists' argument, it is worth noting that the homogeneity assumption also seems to underlie the skeptical interpretation. Indeed, Nado (2014) argues that, since intuitions are fairly heterogeneous, the entire experimentalist project of evaluating intuition seems misguided.
14. Thanks for an anonymous referee for pressing this objection.
15. An anonymous referee suggests that Feltz's (2013) study indicates variation in general intuitions about free will and moral responsibility, since Feltz examines a premise of Pereboom's Four Case Argument. However, I think this study examines neither a general intuition nor even a premise that Pereboom uses at all. Feltz focuses on the premise that "there is no relevant difference between a manipulated agent and an agent in a deterministic world" (2013, p. 55). He attempts to test our intuition about this general premise by eliciting survey responses about a series of cases regarding manipulation and determinism. He seems to assume that, if we have a general intuition about the above premise, then we will have the same intuitions about cases of manipulation and cases of determinism. But this oversimplifies Pereboom's argument. Pereboom can well expect that our pre-theoretical intuitions about these cases are different. His premise states that, however, once one reflects on these cases and realizes how similar they are in their philosophically relevant aspects, they will largely agree that there is no relevant difference between them. But Feltz's study doesn't show variation concerning this general intuition.
16. Thanks to an anonymous referee for pointing out this difficulty.
17. That is to say, contextual information includes both what is explicitly stated and what is tacitly implied in the text.
18. There is an important difference between Cullen's argument and mine. Cullen complains that, since experimentalists do not apply the correct survey methodology, their findings demonstrate variations in survey responses but not in intuitions. By contrast, I am neutral on whether experimental studies elicit intuitions; I claim only that the practices they study are significantly different from philosophical practices.
19. In fact, in a response to Cappelen, Weinberg makes a similar point that Lehrer's arguments work as "textual cues," which steer us "towards what he takes to be the proper viewing conditions for the case" (2014, p. 552). However, as I argue in this section, this view is not friendly to Weinberg's own account of the experimental

critique, for current experimental studies seldom present useful textual cues to respondents.

20. Using such contextual information, Lehrer also means to bring the reader's attention to the general intuition about the opacity principle—the intuition which, as I argued in the last section, constitutes Lehrer's main evidence against externalism.
21. One might respond that a reader could infer this information from the vignette; it is a simple inference from "Charles is unaware that his brain has been altered" to "Charles does not know that he is reliable at temperature estimation." However, even simple inferences can be easily neglected when they are not made salient to a reader.
22. This case might appear unintuitive. However, Lehrer seems to think that it is an intuitive case of non-knowledge and uses it to stress the ordinary aspects of the Truetemp case. The purpose here is not to defend Lehrer's use of the Truetemp case, but to illustrate one important role of contextual information that is common in philosophy but is often missing from experimental research.
23. One might think that these null results are evidence that cultural bias exists in the context-poor Gettier intuition but not in the context-rich Gettier intuition. However, I am not committed to this view for the following two reasons. First, it is difficult to infer from null results to null effects (see, e.g., Aberson, 2002). Second, the null results might be due to the fact that there is no cultural variation even in context-poor Gettier intuitions, as more recent studies on Gettier cases have shown (see, e.g., Machery et al., 2015). I am thankful to an anonymous referee for suggesting both points.
24. Thanks to an anonymous referee for suggesting this response.

Acknowledgments

For helpful comments on earlier versions of this paper, thanks to Jennifer Nado, Max Deutsch, Darrell Rowbottom, David Colaço, Shen-yi Liao, Dana Nelkin, Shyam Nair, Dan Marshall, Daniel Lim, Alexander Ehmann, and Chris Atkinson.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Aberson, C. (2002). Interpreting null results: Improving presentation and conclusions with confidence intervals. *Journal of Articles in Support of the Null Hypothesis, 1*(3), 36–42.

Alexander, J., & Weinberg, J. M. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass, 2*(1), 56–80.

Alexander, J., & Weinberg, J. M. (2014). The "unreliability" of epistemic intuitions. In E. Machery & E. O'Neill (Eds.), *Current controversies in experimental philosophy* (pp. 128–145). New York, NY: Routledge.

Bealer, G. (1992). The incoherence of empiricism. *Aristotelian Society Supplementary Volume, 66*, 99–144.

Bonjour, L. (1998). *In defense of pure reason*. Cambridge: Cambridge University Press.

Block, N. (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science, 9*, 261–325.

Brogaard, B. (2014). Intuitions as intellectual seemings. *Analytic Philosophy, 55*(4), 382–393.

Buckwalter, W., & Stich, S. (2014). Gender and philosophical intuition. In J. Knobe & S. Nichols (Eds.), *Experimental philosophy* (Vol. 2, pp. 307–346). New York, NY: Oxford University Press.

Cappelen, H. (2012). *Philosophy without intuitions*. New York, NY: Oxford University Press.

Chalmers, D. J. (2010). *The character of consciousness*. New York, NY: Oxford University Press.

Chudnoff, E. (2013). *Intuition*. New York, NY: Oxford University Press.

Colaço, D., Buckwalter, W., Stich, S., & Machery, E. (2014). Epistemic intuitions in fake-barn thought experiments. *Episteme, 11*(02), 199–212.

Cullen, S. (2010). Survey-driven romanticism. *Review of Philosophy and Psychology, 1*(2), 275–296.

Cummins, R. C. (1998). Reflection on reflective equilibrium. In M. DePaul & W. Ramsey (Eds.), *Rethinking intuition* (pp. 113–128). Lanham, MD: Rowman and Littlefield.

Deutsch, M. (2009). Experimental philosophy and the theory of reference. *Mind & Language, 24*(4), 445–466.

Deutsch, M. (2010). Intuitions, counter-examples, and experimental philosophy. *Review of Philosophy and Psychology, 1*(3), 447–460.

Deutsch, M. (2015). *The myth of the intuitive: Experimental philosophy and philosophical method*. Cambridge, MA: MIT Press.

Devitt, M. (2011). Experimental semantics. *Philosophy and Phenomenological Research, 82*(2), 418–435.

Feltz, A. (2013). Pereboom and premises: Asking the right questions in the experimental philosophy of free will. *Consciousness and Cognition, 22*(1), 53–63.

Feltz, A., & Cokely, E. T. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition, 18*(1), 342–350.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis, 23*, 121–123.

Horvath, J. (2010). How (not) to react to experimental philosophy. *Philosophical Psychology, 23*(4), 447–480.

Ichikawa, J. J. (2013). Review of philosophy without intuitions. *International Journal of Philosophical Studies, 21*(1), 111–116.

Kauppinen, A. (2007). The rise and fall of experimental philosophy. *Philosophical Explorations, 10*(2), 95–118.

Kim, M., & Yuan, Y. (2015). No cross-cultural differences in the Gettier Car Case intuition: A replication study of Weinberg et al. 2001. *Episteme, 12*(03), 355–361.

Kripke, S. A. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Lehrer, K. (2000). *Theory of knowledge* (2nd ed.). Boulder, CO: Westview Press.

Liao, S. M. (2008). A defense of intuitions. *Philosophical Studies, 140*(2), 247–262.

Liao, S. M., Wiegmann, A., Alexander, J., & Vong, G. (2012). Putting the trolley in order: Experimental philosophy and the loop case. *Philosophical Psychology, 25*(5), 661–671.

Ludwig, K. (2007). The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy, 31*(1), 128–159.

Machery, E., Deutsch, M., Mallon, R., Nichols, S., Sytsma, J., & Stich, S. P. (2010). Semantic intuitions: Reply to Lam. *Cognition, 117*(3), 361–366.

Machery, E., Mallon, R., Nichols, S., & Stich, S. P. (2004). Semantics, cross-cultural style. *Cognition, 92*(3), B1–B12.

Machery, E. (2011). Thought experiments and philosophical knowledge. *Metaphilosophy, 42*(3), 191–214.

Machery, E., Stich, S., Rose, D., Chatterjee, A., Karasawa, K., Struchiner, N., … Hashimoto, T. (2015). Gettier across cultures. *Noûs, 51*(3), 645–664.

Machery, E., Sytsma, J., & Deutsch, M. (2015). Speaker's reference and cross-cultural semantics. In A. Bianchi (Ed.), *On reference* (pp. 62–76). New York, NY: Oxford University Press.

Nado, J. (2014). Why intuition? *Philosophy and Phenomenological Research, 89*(1), 15–41.
Nagel, J., Juan, V. S., & Mar, R. A. (2013). Lay denial of knowledge for justified true beliefs. *Cognition, 129*(3), 652–661.
Nahmias, E., Coates, J., & Kvaran, T. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy, 31*, 214–242.
Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuition. *Nous, 41*, 663–685.
Pritchard, D. (2012). Anti-luck virtue epistemology. *Journal of Philosophy, 109*(3), 247–279.
Pust, J. (2000). *Intuitions as evidence*. New York, NY: Garland Press.
Pust, J. (2001). Against explanationist skepticism regarding philosophical intuitions. *Philosophical Studies, 106*(3), 227–258.
Schulz, E., Cokely, E. T., & Feltz, A. (2011). Persistent bias in expert judgments about free will and moral responsibility: A test of the expertise defense. *Consciousness and Cognition, 20*(4), 1722–1731.
Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language, 27*(2), 135–153.
Seyedsayamdost, H. (2015). On normativity and epistemic intuitions: Failure of replication. *Episteme, 12*(01), 95–116.
Sommers, T. (2010). Experimental philosophy and free will. *Philosophy Compass, 5*(2), 199–212.
Sosa, E. (1980). The raft and the pyramid: Coherence versus foundations. *Midwest Studies in Philosophy, 5*, 3–26.
Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition, 124*(3), 272–283.
Swain, S., Alexander, J., & Weinberg, J. M. (2008). The instability of philosophical intuitions: Running hot and cold on Truetemp. *Philosophy and Phenomenological Research, 76*(1), 138–155.
Sytsma, J., Livengood, J., Sato, R., & Oguchi, M. (2015). Reference in the land of the rising sun: A cross-cultural study on the reference of proper names. *Review of Philosophy and Psychology, 6*(2), 213–230.
Turri, J. (2013). A conspicuous art: Putting Gettier to the test. *Philosophers' Imprint, 13*(10), 1–16.
Vaesen, K., Peterson, M., & Van Bezooijen, B. (2013). The reliability of armchair intuitions. *Metaphilosophy, 44*(5), 559–578.
Weatherson, B. (2003). What good are counterexamples? *Philosophical Studies, 115*(1), 1–31.
Weinberg, J. M. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy, 31*(1), 318–343.
Weinberg, J. M. (2014). Cappelen between rock and a hard place. *Philosophical Studies, 171*(3), 545–553.
Weinberg, J. M., & Alexander, J. (2014). The challenge of sticking with intuitions through thick and thin. In A. Booth & D. Rowbottom (Eds.), *Intuitions* (pp. 187–212). New York, NY: Oxford University Press.
Weinberg, J. M., Alexander, J., Gonnerman, C., & Reuter, S. (2012). Restrictionism and reflection. *Monist, 95*(2), 200–222.
Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*(1), 429–460.
Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell.
Williamson, T. (2011). Philosophical expertise and the burden of proof. *Metaphilosophy, 42*(3), 215–229.