

How not to test for philosophical expertise¹

forthcoming in *Synthese*

Regina A. Rini
NYU Bioethics
gina.rini@nyu.edu

Abstract Recent empirical work appears to suggest that the moral intuitions of professional philosophers are just as vulnerable to distorting psychological factors as are those of ordinary people. This paper assesses these recent tests of the ‘expertise defense’ of philosophical intuition. I argue that the use of familiar cases and principles constitutes a methodological problem. Since these items are familiar to philosophers, but not ordinary people, the two subject groups do not confront identical cognitive tasks. Reflection on this point shows that these findings do not threaten philosophical expertise - though we can draw lessons for more effective empirical tests.

Keywords: expertise defense, methodology, moral intuition, philosophical intuition

0. Introduction

Compared to ordinary people, philosophers spend a lot more time, and have a lot more experience, thinking about such topics as justice, knowledge, or reference. But does this training mean that philosophers are *better* at thinking about these things? Specifically, are philosophers’ intuitive reactions to these sorts of questions less susceptible to distortion than those of ordinary people? It seems well established that ordinary people display patterns of unreliability in their intuitions (Weinberg, Nichols, and Stich 2001, Alexander 2012). Could philosophical training provide some form of immunity to these effects? Until recently, this question was mostly only the target of speculation (Singer 1972, Ludwig 2007, Weinberg et al. 2010, Williamson 2011). But new papers ((Tobia, Buckwalter, and Stich 2013) Schwitzgebel and Cushman 2012) have put claims of philosophical expertise to empirical test, and have

¹ Thanks to Wesley Buckwalter, Eric Schwitzgebel, Kevin Tobia, Guy Kahane, Simon Rippon, and two anonymous referees for *Synthese* for helpful comments on drafts of this paper, and to Nora Heinzelmann, Shaun Nichols, and Steven Lukes and the NYU Sociology of Morality Working Group for discussion. This research received sponsorship from the VolkswagenStiftung’s European Platform for Life Sciences, Mind Sciences, and the Humanities (grant II/85 063).

returned negative results. There now seems to be empirical evidence that philosophical training does *not* improve the intuitions of professional philosophers in this way.

The purpose of this paper is to evaluate these recent empirical tests. It is my view that the studies in question do not establish their charge against philosophical expertise, because there are worries about the way in which they test for expertise. Specifically, I will argue that the use of stimuli likely to be *familiar* to philosophers generates a problem for comparing the responses of philosophers and non-philosopher subjects. Noticing this problem requires us to think carefully about our interpretation of the data discussed – and I argue that the available interpretations do not imply a significant challenge to the notion of philosophical expertise. This paper does not take any view on the status of philosophical expertise itself; I am not offering a positive argument on behalf of philosophical expertise. But I do aim to show that, so far at least, we have not put it to effective empirical test.

In the first section I provide a brief overview of the background debate concerning empirical studies of intuition. In the second section I introduce a framework for understanding this debate, clarifying the empirically-testable hypotheses it generates. The third section applies this framework to the recent empirical tests of philosophical expertise and briefly describes their findings. In the fourth section I then critique these findings, arguing that the authors' interpretation is not the best available. The status of philosophical expertise remains undecided, pending further empirical inquiry. The fifth section concludes by offering suggestions for how next to proceed.

1. Intuitions, Empirical Challenges, and the Expertise Defense

Philosophical practice typically relies on *intuitions*, by which I mean immediate mental states of judging that some concept properly applies to some object or situation. An intuition is not (simply) a perceptual

state; it involves the use of an intellectual category, like *justice* or *knowledge* or *reference*. There is no consensus on the nature of intuitions, though various proposals have great influence. It is, for instance, widely accepted that intuitions are ‘non-inferential’, where this means that they arise spontaneously in the mind, not as a result of deliberation or reasoning (Audi 2008). George Bealer (2000) refers to intuitions as ‘intellectual seemings’, or conscious episodes of applying a priori modal concepts to instances. Hilary Kornblith (1998) argues instead that intuitions are simply first-approximation reactions, drawn from prior experience of applying relevant concepts to other instances. Whatever the details, it is widely believed that intuitions are central to the contemporary practice of philosophy: in some way or other, philosophy involves constructing systematic theories from the content of intuitions.²

Recent attention has turned to understanding the mental operations underlying intuitions, and how these might or might not confer epistemic authority on the contents of intuitions. Williamson (2004) calls this the “psychologizing of philosophical method” and warns that it can lead to a deep sort of skepticism. Indeed, a series of recent papers and books have claimed to undermine the credentials of intuition on the basis of empirical findings about how people reach judgments concerning philosophical topics (Weinberg, Nichols, and Stich 2001; Machery et al. 2004; Sinnott-Armstrong 2008; Alexander 2012). Such arguments depend crucially on evidence from psychology and experimental philosophy, showing that people’s intuitions are sensitive to a host of things which seem irrelevant to the truth of philosophical theories. For instance, intuitions about moral permissibility and knowledge-possession seem to be sensitive to the order in which the intuiter encounters test cases (Petrinovich and O’Neill 1996; Swain et al. 2008; Lanteri et al. 2008; Lombrozo 2009; Liao et al. 2011; Wiegmann et al. 2012) or whether the intuiter imagines herself as a bystander or actor in a test case (Nadelhoffer and Feltz 2008). On the sensible assumption that such factors are irrelevant to the truth about moral permissibility or

² However, Cappelen (2012) argues that intuition’s role in philosophy is far less essential than has been portrayed. See also Williamson (2007) for the claim that appeals to intuition actually involve disguised employment of deductively valid arguments.

knowledge, these studies seem to show that philosophical theories constructed from intuition risk incorporating distorting elements.

However, all of these studies were conducted using non-philosopher subjects, people without extensive training in philosophical modes of thinking, and that fact has given rise to a particular counterargument. Can we really assume that professional philosophers exhibit the same vulnerability to such distorting factors as do non-philosopher subjects? We do not distrust the practice of professional mathematicians simply because ordinary people have unreliable intuitions about the cardinality of infinite series (Ludwig 2007, 148).³ And if professional philosophers are relatively immune to these effects, then distortion in folk intuitions simply has no relevance to the status of professional philosophical practice. This line of argument, as a response to empirically-fueled attacks on intuition, has come to be called the *expertise defense* (Weinberg et al. 2010).

Claims for philosophical expertise predate the present controversy. Peter Singer (1972) has argued that a professional ethicist can be a “moral expert”, a person “familiar with moral concepts and with moral arguments, who has ample time to gather information and think about it” (117). More recently - in opposition to empirical findings of distortion in folk intuition - several authors have suggested that philosophical training may sharpen skills necessary to the interpretation and utilization of intuitive test cases, including the abilities to recognize ambiguity and vagueness and to appreciate which details of a case are relevant to philosophical questions (Ludwig 2007, 152; Kauppinen 2007, 113; Grundmann 2010, 502; Horvath 2010, 467; Williamson 2011, 216). Experimentalist critics of intuition have replied, on the basis of evidence about expertise in other domains, that philosophical training is unlikely to provide the correct type of feedback to instill genuine, distortion-avoiding expertise (Weinberg et al. 2010).

³ Though see (Ryberg 2013) and (Rini 2014) for discussion of the analogy between moral expertise and expertise in fields like mathematics.

So far there has been much speculation about the status of the expertise defense, and a good deal of argument over the burden of proof. Weinberg and his coauthors (2010) think the expertise defense dubious enough that its proponents must provide evidence in its favor, while others (e.g. Grundmann 2010, 495; Williamson 2011) assert that the burden is instead on those who would challenge such an important aspect of philosophical practice on evidence about non-philosophers. Amid these disputes, what everyone *does* seem to agree upon is that this is ultimately an empirical matter, to be decided by directly testing professional philosophers' ability to avoid distorting influences in their intuitive practice.

Indeed, empirical tests are necessary, and now some have been done. In Section 3 I will describe two papers purporting to give empirical evidence undermining the expertise defense. However, as I will argue in section 4, neither of these papers actually settles the matter; they are not adequate empirical tests. If my argument is correct, then we are back where we started. The expertise defense is an empirical matter, but we do not have evidence available to settle its status.⁴

2. Two empirical hypotheses

Before proceeding to the empirical evidence against the expertise defense, we should distinguish two ways that philosophical expertise might be formulated as an empirically testable hypothesis. Either of the following proposals, if found to withstand testing, would vindicate the use of intuitions in philosophical inquiry (regardless of evidence of distortion in *folk* intuitions).

⁴ An important qualification: much of the debate over the expertise defense appears to proceed on the assumption that philosophical intuition is all of one sort – that whatever we might say about intuitions in epistemology, we might also say about intuitions in metaphysics or ethics. There is some reason to be skeptical of this assumption (Nado 2012), but I will not be able to engage with it here. Still, it should be noted that the empirical studies discussed below deal solely with *moral* intuition, and there is a live question as to whether anything said about these findings can be generalized to other areas of philosophy.

(1) **Philosophers have *better intuitions* than do non-philosophers.** That is, the intuitions of philosophers are significantly less likely to be affected by distorting factors than are the intuitions of non-philosophers.

(2) **Philosophers make *better use of intuitions* than do non-philosophers.** That is, though the intuitions of philosophers may not be any less subject to distortion than are those of non-philosophers, philosophers who employ intuitions in theory construction do so in some way that ameliorates the effects of distortion.

Some participants in the debate over intuitions seem to assume that philosophical expertise is of the first sort: over time, philosophers become less likely to *have* distorted intuitions, so the intuitions they employ in constructing theories are therefore of higher quality. But couldn't philosophical expertise be of the second sort? Perhaps philosophers are no more likely than anyone else to have good intuitions, but are much better at judiciously using the ones they do have. Perhaps some intuitions are *recognized* by philosophers as distorted and so discarded. Those intuitions that survive are arranged carefully, so that the weaker ones sit alongside stronger counterparts.

The second proposal points to a long-standing distinction in moral theory, between simple reactions and 'considered judgments'. The latter are a special category of judgments, rendered under reliability-conducive circumstances. In his early 'Outline of a Decision Procedure for Ethics', John Rawls lays out some of the conditions for judgments to be, in this sense, 'considered'. Considered judgments (Rawls 1951, 181-183):

(a) respond to "not especially difficult" cases which are "likely to arise in ordinary life"⁵

⁵ Interestingly, Rawls also says that "all judgments on hypothetical cases are excluded" (Rawls 1951,182). He abandons this requirement in *A Theory of Justice*, though he continues to maintain that "in deciding which of our judgments to take into

- (b) are attentive to specific facts within cases
- (c) exhibit “certitude” – a feeling of confidence by the intuiter
- (d) are stable – there is agreement across time (within one judge) and between multiple judges
- (e) do not derive from explicit consultation of an ethical theory (hence are “intuitive”).

Notice that quite a few mental states we might call ‘moral intuitions’ are excluded by these criteria.

Most importantly: a fleeting, uncertain, unstable moral intuition will *not* count as a considered judgment. I wish to flag this early, because I will argue that the empirical results discussed below do not seem to be carefully targeted at considered judgments. On the best available interpretations, the distortion these studies purport to find in philosophers’ intuitive practice seems to be explained by fleeting, uncertain, or unstable reactions – not the sort of considered judgments that philosophers (aim to) use in constructing theory. (Of course, there is still a problem if it turns out that philosophers are bad at recognizing when they have actually reached a considered judgment, and are unwittingly employing unstable intuitions. I take it that this is the claim of some of the empirical studies, which I will discuss.) Before we get to all this, however, notice that the distinction between mere reactions and considered judgments is many decades old – it is not an ad hoc maneuver to avoid new empirical studies.⁶

Rawls specifies not only conditions for the *production* of considered judgments, but also further conditions for their responsible *use*:

account we may reasonably select some and exclude others. For example, we can discard those judgments made with hesitation, or in which we have little confidence” (Rawls 1971, 47).

⁶ It is worth noting that focusing on Rawlsian considered judgments does not involve what Weinberg and Alexander (2014) call a “thick” conception of philosophical intuition. That is, I will not suppose that the intuitions under discussion are a class with special conceptual or cognitive properties (see Ludwig 2007 and Kauppinen 2007). Weinberg and Alexander argue that the special properties of thick intuitions may make them untestable in experimental studies, and perhaps even undetectable in ordinary philosophical practice. But the cognitive restrictions implied by “considered judgments” are relatively pedestrian.

A reasonable man knows, or tries to know, his own emotional, intellectual, and moral predilections and makes a conscientious effort to take them into account in weighing the merits of any question. He is not unaware of the influences of prejudice and bias even in his most sincere efforts to annul them; nor is he fatalistic about their effect so that he succumbs to them as being those factors which he thinks must sooner or later determine his decision. (Rawls 1951, 179)

'Outline of a Decision Procedure in Ethics' is plausibly understood as the root of intuitive methodology in contemporary moral philosophy, and already there is explicit attention to the possibility that intuitions (including, presumably, those belonging to philosophers) may be biased. Although Rawls makes no claims about expertise, it does appear that he saw skill in moral method as partly a matter of *using* intuitions well – i.e. of picking out considered judgments from mere reactions – and not simply a matter of having better intuitions.

Discussion of the expertise defense therefore ought to distinguish these two proposals. Notice that selecting an appropriate empirical test of the expertise defense depends upon which proposal is at issue. The existing methodology used with non-philosopher subjects – manipulating a distorting factor and observing whether it influences intuitions – is an adequate test *only* of the first proposal. Finding that philosophers' intuitions are subject to distorting factors would challenge *only* the claim that philosophers have better intuitions than non-philosophers. But this finding is *consistent* with the claim that philosophers make better use of their intuitions, so it simply is not an appropriate empirical test. Some other empirical test must be employed to examine the second proposal. Fortunately, the studies I will discuss in the next section manage, collectively, to test both hypotheses – but we will need to keep the distinction in mind in order to evaluate them.

3. Empirical evidence against the expertise defense

If the expertise defense can be formulated as two distinct hypotheses, then we will have to use two distinct means of testing it. To test the first hypothesis, we will have to see if philosophers' intuitions are any less susceptible to distorting influences than are folk intuitions. To test the second hypothesis, we will have to see if philosophers are better than non-philosophers at using intuitions in a way that mitigates whatever distortions may arise. I will now discuss two recent studies which, between them, appear to test both expertise hypotheses – and appear to show that both hypotheses fail.⁷

First: Tobia, Buckwalter, and Stich (2013) (hereafter 'TBS'), who tested only the first hypothesis, regarding the susceptibility of philosophers' intuitions to distorting factors. In this case, the distorting factor was an Actor-Observer effect. A well-documented finding in social psychology (e.g. Jones and Nisbett 1971), the Actor-Observer effect is a tendency to form divergent assessments of one's own actions and identical actions of other people. A recent study by Nadelhoffer and Feltz (2008) showed an Actor-Observer effect in folk moral judgments regarding the famous Trolley Switch case. (The Trolley case concerns the moral rightness of allowing a runaway trolley to kill five innocent people, versus pushing a switch to divert the trolley to a side track where it will kill only one (Foot 1967).) Nadelhoffer and Feltz showed that non-philosopher subjects were more likely to think it morally permissible to flip the switch when they imagined another person as the actor than when they imagined themselves as the actor.

⁷ These are not the only empirical studies with some relevance to the expertise defense. For instance, Schulz et al. (2011) appear to show that professional philosophers exhibit a (presumed distorting) link between personality traits and views on free will and moral responsibility. Similarly, a series of behavioral studies by Eric Schwitzgebel and colleagues (Schwitzgebel 2009; Schwitzgebel et al. 2012) appear to show that professional moral philosophers are no better morally behaved than ordinary people. I leave these studies to the side because they make assumptions (about how to measure expertise or about the relationship between knowledge and behavior) that require separate discussion.

TBS regard the Actor-Observer effect as *distorting* in the moral domain because “whether an action in a moral scenario is framed in first or third person terms is almost always irrelevant to a moral judgment about the action” (2012, 3). They used the Switch case and Bernard Williams’ ‘Jim and the Indians’ case (Smart and Williams 1973) as stimuli.⁸ They tested Actor and Observer variants of the cases on non-philosopher subjects and professional philosophers.⁹ And they indeed found a significant Actor-Observer effect among *both* non-philosopher subjects and philosopher subjects.¹⁰ Since professional philosopher subjects were just as likely as non-philosopher subjects to display a distorting influence of this sort, the first formulation of the expertise defense seems to be experimentally undermined. That is, the hypothesis that professional training allows philosophers to have *better* (or at least less distortion-prone) intuitions appears to have failed this test.

⁸ The ‘Jim and the Indians’ case concerns an innocent man, Jim, who is given the opportunity to save a number of defenseless South American Indians from a sadistic paramilitary force, if he will only agree to pull the trigger and kill one of the condemned Indians himself (Smart and Williams 1973, 98).

⁹ TBS used undergraduates for non-philosopher subjects (borrowing data from Nadelhoffer and Feltz (2008) for the Switch case). Philosopher subjects were APA Pacific conference attendees with PhDs in philosophy (Tobia et al. 2012, 3-5).

¹⁰ Curiously, the direction of the effect reversed between the two groups: non-philosopher subjects were more likely to find the proposed action in Switch or Jim and the Indians morally obligatory in the Observer condition, while philosophers were more likely to find it obligatory in the Actor condition (Tobia et al. 2012, 4-5). It is intriguing that non-philosophers and philosophers displayed the Actor-Observer effect in opposed directions – but, for the present, this directional difference doesn’t matter. That an Actor-Observer effect constitutes a distortion of moral intuition does not depend on the direction of the effect – any difference between Actor and Observer responses is considered evidence of distortion. The reversal is, however, interesting for interpretation of Nadelhoffer and Feltz’s original data. In their paper, they suggest that Actor subjects were trying to avoid the aversive experience of imagining hitting the switch: “If you are asked to imagine yourself to be in the position of having to decide whether it would be permissible for you to hit the switch, one easy way of keeping yourself from having to make such a hard decision is to simply judge it to be impermissible!” (Nadelhoffer and Feltz 2008, 141). It is unclear how to understand philosophers’ greater willingness to endorse action in the Actor condition on this interpretation.

There remains, however, the second formulation of the expertise defense: that even if philosophers' intuitions are no less distorted than those of non-philosophers, philosophers are able to *use* their intuitions in some way that mitigates this distortion. A second study, Schwitzgebel and Cushman (2012) (hereafter 'SC'), offers a test of this formulation of the expertise defense. Previous research (Petrinovich and O'Neill 1996; Lombrozo 2009; Liao et al. 2011; Wiegmann et al. 2012) had found that non-philosopher subjects' moral judgments were sensitive to *order effects*: their intuitive responses to test cases depended upon the order in which these cases were presented. It seems clear that order effects are a form of distortion; presumably the moral valence of a particular action does not depend upon whether the intuiter has just thought about some other action.

SC's subjects were people recruited over the internet, placed into three expertise categories based upon (self-reported) educational background: philosophers (with a Masters or doctorate in philosophy), other academics (with an advanced degree in some other field), and non-academics (non-philosopher subjects – those without an advanced degree in any academic field). The researchers also tracked an important subset of philosophers: those with a PhD and specialization in ethics, the most plausible expert moral intuiters. All subjects read and responded to a large number of scenarios, the details of which are not necessary to the present point (see Schwitzgebel and Cushman 2012, 138-9). For simplicity, I will focus on 'double effect' scenarios: cases like the Trolley case and the Trolley's famous cousin, the Footbridge case.¹¹ Typically, most people (philosophers and ordinary people) treat Trolley and Footbridge as exhibiting a moral difference, generally regarding the action in Footbridge as morally worse than that in Trolley. The non-equivalence of the two cases is central to their role in the philosophy literature. So SC tested whether their subjects rated Trolley-type and Footbridge-type cases *equivalently*. They varied

¹¹ The Footbridge case (Thomson 1976) resembles the Trolley case in that it involves bringing about the death of one person to save five, but it differs in that the agent must physically push the one person in front of the oncoming vehicle, rather than flipping a switch to direct danger toward a person.

which of the two case-types was presented *first*, and tested whether this ordering had an effect on equivalence ratings. And indeed it did: among all expertise categories, non-philosopher through ethics PhDs, subjects were more likely to judge the case types equivalently if they encountered Footbridge-type cases first.

Whatever one thinks about the Footbridge and Trolley cases, whether or not they are morally equivalent presumably does not depend upon the order in which one reads about them – so these results appear to show a distorting order effect on intuitions. Since philosophers were not immune to the distortion, this result appears to further undermine the first formulation of the expertise defense: philosophers apparently do not have better intuitions. So far, then, SC have merely provided additional evidence for the claims made by TBS. However, SC go beyond this, testing how their subjects *used* intuitions in evaluating general philosophical views. Hence these next results can be understood as a test of the second formulation of the expertise defense.

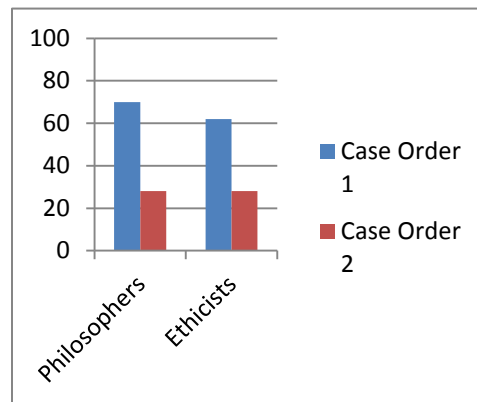
After providing intuitive reactions to particular cases, SC's subjects were asked to indicate their agreement with several general moral principles. Importantly, the principles were related to the cases presented earlier, in that the cases (or others like them) are typically invoked by moral philosophers in constructing or challenging the principles. So, for instance, the Doctrine of Double Effect (which morally distinguishes bringing about harm as an intended means versus a foreseen side-effect) has been widely discussed as an explanation for differing responses on Trolley-type and Footbridge-type cases.¹²

SC were interested in seeing whether subjects' endorsement of moral principles would be influenced by the order in which they encountered cases. And this is what they found. Here we can look at detailed results for one of these tests (there are others). Some subjects saw an ordering of cases like this: a

¹² See, among many, Foot (1967), Thomson (1976), Quinn (1989), Unger (1996), Kamm (2000), Thomson (2008), and Liao (2009).

Switch-type case followed by a Push-type case, then a Bad Moral Luck Case followed by a Good Moral Luck case. (Good and Bad Moral Luck cases relate to another principle, not discussed here.) Other subjects saw an ordering reversed within each of these pairs: Push-Switch, Good-Bad. Schwitzgebel and Cushman compared the willingness of these two groups of subjects to endorse the Doctrine of Double Effect.¹³ Among philosophers in the first group, about two-thirds indicated agreement with the Doctrine. But among the other group, less than one-third indicated agreement with the Doctrine. I've reproduced this data in the graph below:

Fig. 1 Percent endorsing Doctrine of Double Effect (Schwitzgebel and Cushman 2012)



Case Order 1: Switch-type case followed by Push-type case, then Bad Moral Luck followed by Good Moral Luck. Case Order 2: Push-type case followed by Switch-type case, then Good Moral Luck followed by Bad Moral Luck. 'Philosophers' are people with a Masters or PhD in Philosophy. 'Ethicists' are a subset of Philosophers: Philosophy PhDs

¹³ More precisely, what subjects actually responded to was the following question: "Sometimes it is necessary to use one person's death as a means to saving several other 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?" SC interpreted responses of 'worse' as endorsing the Doctrine. (Schwitzgebel and Cushman 2012, 138-140)

reporting a specialization in Ethics. As reported in Schwitzgebel and Cushman (2012, 146).

This result challenges the *use* formulation of the Expertise Defense. If professional philosophers are better than non-philosophers at using intuitions in some way that reduces distorting effects, then the order in which these cases were presented should *not* have affected their endorsement of the Doctrine. Yet this seems to be what happened.

In fact, philosophers' endorsement of moral principles appears to be *more* sensitive to ordering effects than that of non-philosophers. According to Schwitzgebel and Cushman, the willingness of *non-philosopher* subjects to endorse the Doctrine was *not* affected by the order in which they read the case. This is an interesting finding: case-order affected both non-philosophers' and philosophers' *intuitions*, but affected only philosophers' principle-endorsement. It seems that philosophers' professional training allowed them to appreciate the *relationship* between the cases and principles they considered, in a way that non-philosophers did not, but it did not allow them to avoid distorting order effects. According to Schwitzgebel and Cushman, this is a sign that philosophical expertise may simply be a matter of *rationalization* (149). Unlike non-philosophers, professional philosophers are especially disposed and especially well-equipped to make sure that their case judgments and principle endorsements *appear* to be internally consistent. Unfortunately, since their reactions to cases are partly driven by irrelevant order effects, philosophers' consistency serves only to unwittingly systematize the influence of these irrelevant factors.

Apparently, then, there is evidence against both formulations of the expertise defense. As shown in both the TBS and SC studies, philosophers' intuitions are often just as vulnerable to distortion as those of non-philosophers. And the SC study also shows that philosophers do not make use of their intuitions

in some way that eliminates this distortion. If these findings are correct, the expertise defense appears to be in serious danger of empirical refutation.

4. Problems with testing philosophical expertise

In the rest of this paper, I will argue that the findings just described are much less threatening to the expertise defense than it might appear. If my argument is successful, the result will be to restore the status quo ante: the validity of the expertise defense will remain an empirical matter, and we will once again be in a state lacking decisive empirical evidence either way. My aim is not entirely negative, however. Seeing how *not* to test for philosophical expertise can help us refine our empirical techniques, and perhaps come up with better ways to test for it.

4.1 The trouble with familiarity

Recall the dialectical situation: the aim of these studies is to demonstrate that philosophical expertise provides no great immunity against distorting factors in the production and use of moral intuitions. This aim is implemented by comparing the responses of philosophers and non-philosophers and showing that they make similar sorts of mistakes. This side-by-side comparison appears to directly undermine speculative claims about the intuition-using superiority of philosophers.

Yet there is something odd about this arrangement, something that has not been much noticed in existing discussion. It is problematic to compare philosophers and non-philosophers directly on these tasks, because the two groups do not engage with the stimuli in the same way. For non-philosopher subjects, the cases and principles employed are likely to be quite unfamiliar; intuitive judgments and reasoning about them are likely to be experienced as relatively novel cognitive projects. But for philosophers, this is quite unlikely to be true. The case-types and principles employed as stimuli are well-

known to philosophers. Philosophers will have memories of previously encountering the stimuli, and how they have responded in the past. So when philosophers react to these stimuli, they are doing something different from non-philosopher subjects, a different sort of cognitive task.

In a moment I will explain why this is a problem for the experiments; I will argue that we have hypothesis-independent reason to expect philosophers' responses to familiar cases to be unaffected by the experimental manipulations. If I am right about this, then we have reason to be skeptical that these experiments show what they are claimed to show. First, in case there is any doubt, I will establish that the stimuli used in these studies are likely to be highly familiar to professional philosophers. The Trolley and Jim and the Indians cases used by TBS are, as already noted, taken directly from very famous works in moral philosophy (Foot 1967, Smart and Williams 1973). The cases used by SC, though not all verbatim duplicates of published thought experiments, were deliberately designed to mimic the logic of Trolley- and Footbridge-type scenarios, in a way that would surely be apparent to trained philosophers.¹⁴ Similarly, the moral principles tested by SC are likely to be very familiar to philosophers. The Doctrine of Double Effect is an extremely well-known, central piece of moral philosophical theory. In one form or another, it dates (at least) to Thomas Aquinas. It is explicitly discussed in many contemporary papers in normative ethics, including the modern classics from which the Trolley problem itself originates (Foot 1967; Thomson 1976).¹⁵ It is very likely to appear in any introductory course taught by professors of normative ethics; it features prominently in standard texts.¹⁶

¹⁴ One of the scenarios involved a boxcar, rather than a trolley, moving under a footbridge occupied by a familiarly large man. Some of the scenarios used in other parts of the study were *directly* taken from well-known literature, such as those testing intuitions about moral luck (see Williams 1982, Nagel 1979).

¹⁵ See also other citations in note 12 above.

¹⁶ See, for example, Fischer and Ravizza (1992, 162-198), Kagan (1997, 103), Driver (2006, 128-135), Shafer-Landau (2009, 206) and Gensler (2011, 156).

Given these facts, it is likely that *very few* of the professional philosophers, at least those with a specialization in Ethics, were unaware of the cases or the Doctrine before encountering them in this survey. Schwitzgebel and Cushman are likely to concede this point; they remark on how striking the ordering effect finding is, “considering how familiar and widely discussed the doctrine is within professional moral philosophy (149).”

So it is very likely that the task presented to philosophers (responding to *familiar* cases and principles) was not the same as the one presented to non-philosopher subjects (responding to *unfamiliar* cases and principles). Of course, this difference is not in itself problematic for the claim that these studies challenge expertise. After all, on many uses of the term, expertise in moral philosophy *requires* familiarity with standard cases and principles in the literature; it is hardly an objection to the design of these experiments to point out that the philosophers subjects were indeed experts in this sense. Further, the fact that philosophers confronted a somewhat different task than did non-philosophers might be taken to *heighten* the threat these results pose to the expertise defense. It appears that even with the benefit of familiarity, philosopher subjects were *still* subject to distorting effects. Doesn't this show how empirically implausible the expertise defense must be?

I do not think so. Instead, I think it points to something amiss with the experimental paradigm. The fact that *familiar* cases and principles generated these results give us independent reason – that is, independent of affirming the expertise defense – to question whether these experiments really tested what they were meant to test. I'll now explain my skepticism.

Imagine you are one of the professional philosopher subjects in these studies. You read a familiar case from the moral philosophy literature, or a familiar principle, and are asked to give your opinion. Since these cases and principles are familiar, you must have thought about them before. You've probably taught them to your students. You may even have published articles or books defending certain views

about them. Since you've thought about them before, the *most reasonable response* to being asked about them yet again is to report whatever opinion you have previously come to, before you took this particular survey. Why do otherwise? These are the same old cases and principles; there is nothing (apparently) special about them this time around. Why give any answer other than the familiar one?

If the preceding reflection sounds right, then we have independent reason – independent of the expertise defense – to *expect* that professional philosophers would *not* be affected by the Actor-Observer and presentation order manipulations used in these studies. This is because philosophers will have come to their familiar views *before* participating in these surveys and encountering these manipulations. Hence we would not expect the manipulations of these studies to have *any* effect, since philosophers are likely merely to report their familiar views, rather than forming new intuitions.

Note that this argument holds *even if* the expertise defense is false. That is, even if philosophers' use of intuitions is just as vulnerable to distortion as that of non-philosophers, we should not expect to see that in *this* study. It might very well be that the first time (perhaps long ago) that philosophers encountered these cases and principles, they were affected by distortions. But by now the cases, principles, and accompanying responses are familiar. We should not expect any effect at all in this once-again iteration.

Yet there *is* an effect. Seemingly, philosophers *are* affected by Actor-Observer and presentation order manipulations, *even though* the cases and principles are highly familiar. This requires an explanation. Since philosophers' likely familiarity with the stimuli – independent of any claims about expertise – gives us reason to predict we would not get this effect, then we should be able to provide some explanation for why it did occur. If we cannot give such an explanation, then we should be skeptical about these findings. As in any science, if we encounter anomalous results from an experimental technique – that is,

results we have reason to predict against, independent of the hypothesis being tested – then this gives us reason to doubt the reliability of the experimental technique.¹⁷

This is what I am calling the *familiarity problem*. We need an explanation for why philosopher subjects did not simply respond to familiar cases and principles with familiar answers. If we cannot provide such an explanation, then we should doubt the findings. Now, I think that we *can* provide such an explanation; we *can* explain why familiar stimuli produced unfamiliar responses. But I do not think we can do so in a way that preserves the authors' intended interpretation of these studies.

In the rest of this section, I will consider several possible explanations. Each one does away with the anomaly – it allows us to see why distortion emerged despite familiarity. However, I will argue, none of these explanations produce a clear challenge to the expertise defense. That is, the best available explanations for these findings do not involve clear evidence against the claim the philosophers have better intuitions or make better use of their intuitions.

Before proceeding to the candidate explanations, it is important to be clear about the magnitude of what it is we are trying to explain. The studies discussed above did *not* show distortion in the intuitions of all or even most professional philosophers. The results presented by TBS suggest that only 25-27% of philosophers are affected by the Actor-Observer manipulation.¹⁸ Similarly, SC report that only 42% of all

¹⁷ For example: Bennett et al. (2010) conducted an fMRI comparison of social processing in the brains of healthy human beings... and the brain of a dead fish. There was obviously good reason to predict that the dead fish's brain would *not* respond selectively to images of social situations, yet seemingly it did! The purpose of this study, of course, was to point out that certain investigative techniques (in this case, inadequate statistical correction for multiple comparisons) lead to unreliable results. There is a similar logic in LeBel and Peters' (2011) critique of social psychological research methods in the wake of Bem's (2011) infamous demonstration of 'precognition'.

¹⁸ I get this figure by taking the difference between percentages of respondents who approved of an action in the Actor condition with those who approved in the Observer condition. For the Trolley case, TBS report 36% (Actor) versus 9%

philosophers, and only 34% of Ethics specialists, exhibited order effects on their endorsement of the Doctrine of Double Effect (Schwitzgebel and Cushman 2012, 149). So our explanation need only account for the behavior of *some* philosophers, somewhere more than a quarter and less than half of the total. We can allow that the larger fraction, who did not show the effect, did simply report familiar responses to familiar cases and principles (and are apparently not susceptible to these distortions).¹⁹

4.2 Explanation I: Not really experts

One explanation for why (some) philosopher subjects did not merely report familiar responses is that the cases and principles were *not* familiar to these subjects. For purposes of these studies, expertise status was based upon self-reported qualifications: respondents indicated whether or not they possessed a doctorate in Philosophy and (in SC's study) a research specialization in Ethics. Especially in SC's study, whose participants were people on the internet, it is possible that some fraction of those claiming qualifications misrepresented their background. (This is less likely for TBS, who surveyed attendees at a Philosophy conference.) Even among those whose claim of the relevant training was

(Observer) acceptance; the figures for Trolley are 89% and 64% (Tobia et al. 2012, 4-5). The assumption is that this percentage represents the fraction of subjects who *would have* responded differently had they been assigned to the other experimental condition. If one would not have responded differently, then one is not susceptible to the effect.

¹⁹ One might think that the fact that only a minority of philosophers apparently exhibited distorting effects is itself a point in favor of the expertise defense. Couldn't expertise defense proponents simply insist that not all so-called philosophical 'experts' (those with doctorates in the field) really are expert in the relevant sense? In that case, expertise defense proponents can accept these results at face value; they need only admit that we will have some difficulty picking out the genuine experts. But this is not a good position for the expertise defense proponent, because each individual philosopher must wonder whether she is in the group affected by distortion. One cannot know the answer to this introspectively, and the data do appear to show that a sizable fraction of those with philosophical training remain affected by distortion. Hence epistemically responsible practice would seem to require having oneself empirically checked for distorting effects – exactly the sort of 'psychologizing' of philosophy resisted by expertise defense proponents.

truthful, there may be some who are simply not familiar with the cases and principles employed in these studies.

This explanation solves the familiarity problem: there is no familiarity problem if there is no familiarity.

But it also presents no threat to the expertise defense. People who falsely claim professional training are obviously not experts in the relevant sense. It is only slightly more controversial to say that Philosophy PhDs who are unfamiliar with these cases and principles are also not experts in the relevant sense. If one underwent years of training but never encountered the Trolley Problem or the Doctrine of Double Effect, then one was very probably not trained in the discipline meant by the designation 'moral philosophy' in this debate. So if the familiarity problem is explained by unfamiliarity, then distortion appears to be accounted for by non-experts, and the expertise defense is unscathed.

4.3 Explanation II: Lack of attention

Another possibility is that some philosophers did not give familiar responses to familiar stimuli because they failed to notice that the stimuli were familiar. In many domains it is possible to interact, perhaps in complex ways, with a familiar stimulus without noticing its familiarity. For instance, I had the experience of pointedly stepping around a woman at a coffeeshop who happened to be blocking the aisle while talking on her phone. This maneuver had several fairly complex components: I had to spatially process her position in order to adjust my own, and I had a vague but relatively sophisticated social representation of her as a person being slightly inconsiderate. It was only after I'd walked away that I realized the woman was actually a friend of mine, someone I knew fairly well. I had been distracted (I was *also* talking on my phone) and so I had failed to notice the familiarity of the person, even though I employed some sort of representation of her in my social and spatial processing.

Could some of the philosopher subjects in these experiments have a similar relationship to the stimuli?

They interacted with them sufficiently to generate an intuitive response, but not sufficiently as to

recognize their familiarity? Let us grant that this is possible. Sometimes people – even professional philosophers – fill out surveys quickly, without thinking very hard about their answers. If so, this solves the familiarity problem: philosophers did not give familiar responses because they did not process the stimuli as familiar.

However, if this explanation is correct, then these results are no challenge to the expertise defense. Proponents of the expertise defense are certainly not committed to the view that trained philosophers have an ability to always, everywhere, under any circumstances, avoid intuition-distorting effects. Recall Rawls' notion of considered judgments, which are generated only under suitable circumstances. Considered judgment requires a relatively high level of cognitive engagement, which is very unlikely to obtain when one is so distracted as to fail to notice that cases or principles are familiar. So proponents of the expertise defense can happily accept this explanation: they need only insist that the distracted philosophers in this study were not engaged with the stimuli in the careful, thoughtful way that philosophers (presumably) do when seriously constructing philosophical theories, rather than quickly responding to surveys.²⁰

4.4 Explanation III: Lack of familiar responses

A third explanation for why some philosophers do not simply report familiar responses to familiar stimuli is that, though they recognize the stimuli as familiar, they do not have familiar *responses*. Perhaps they have so far refrained from forming judgments about these cases or principles. Hence they

²⁰ Sosa (2007; 2010) claims that experimental findings involving even non-philosopher subjects (not professional philosophers) can be explained in a similar way. He argues that the apparent differences of opinion tracked by the studies may be caused by purely verbal ambiguity, rather than genuine difference. In particular, he says, “verbal reports by rushers-by on the street corner are hard to take seriously as expressive of considered views with full understanding of the issues under dispute” (Sosa 2010, 422). For related points, see Cullen (2010) and Bengson (2013).

would be unable to simply report their prior reactions, because they had none. If this so, then the familiarity problem can be explained: we would not expect mere reporting of familiar responses if the subjects have until this point not formed familiar responses to them.

But, once again, this explanation does not generate a challenge to the expertise defense. If these philosopher subjects have until now refrained from forming judgments on the cases and principles, why would they bother to do so *now*? And if they only now bother to do so, does this have much bearing on their ordinary philosophical practice? Note that subjects were not given the option of declining to answer the questions: in the TBS experiments, subjects were forced to choose ‘yes’ or ‘no’. Similarly, the only response options available in SC’s study were effectively endorsement or rejection of the Doctrine of Double Effect.²¹

To get a baseline on philosopher’s actual readiness to form judgments about these issues, we can look at the Phil Papers internet survey of professional philosophers, which allowed a range of other answers (Bourget and Chalmers forthcoming). That survey did not ask about the Doctrine of Double Effect, but it did ask about the Trolley case. Only 47.9% of target professional philosophers (faculty at certain universities) chose simple options of accepting switching or not switching; the other 52% of subjects chose responses ranging from “leaning toward” an option to “there is no fact of the matter”. Similarly, when asked their views on familiar normative theories like consequentialism, deontology, and virtue ethics, only 25.2% of philosophers selected simple endorsement of any of the three. In general, it seems, philosophers’ actual moral judgments do not correspond well to forced choice binary answers.

²¹ To be precise: although Schwitzgebel and Cushman gave their subjects a forced choice in endorsing moral principles, they did use a Likert scale for responses to cases. However, their main results come from binary coding pairs of responses to cases as either ‘equivalent’ or ‘inequivalent’ (Schwitzgebel and Cushman 2012, 140-141).

If someone is not disposed to give binary answers, and has until this point refrained from forming a judgment about a case or principle, then a forced choice between endorsement and rejection is unlikely to reveal much genuine commitment. At best, it is likely only to indicate a weak inclination and a desire to please the experimenters by saying *something*. It is very unlikely to tell us anything about this person's considered judgments or her ordinary use of intuitions in coming to decisions about philosophical principles and theories, since this person's ordinary philosophical practice apparently involves *refraining* from forming these judgments. So if this explanation is correct, these results do not bear on ordinary philosophical practice, and therefore do not bear on the expertise defense.²²

4.5 Explanation IV: Diachronic instability

A final candidate explanation: some philosophers failed to report familiar responses because they do not have *diachronically stable* familiar responses. That is, although they recognized these cases and principles as familiar *and* they have previously formed judgments about them, they were not able to reproduce those responses. Perhaps they have even tried to do this, but failed. Perhaps they misremembered their previous responses: they produced immediate responses to the cases presented in the survey (and so were affected by the Actor-Observer and ordering distortions) and then falsely remembered these as consistent with earlier reactions to these cases and principles.

This explanation solves the familiarity problem: if some philosophers do not have diachronically stable responses to the cases and principles, then they cannot merely report familiar responses (even if they think they have). And, unlike the previous three explanations, this one *does* seem to be threatening to the expertise defense. If (some) philosophers' intuitions are diachronically unstable, then their intuitions are not reliable, despite philosophical training. To put it another way: the expertise defense implies that

²² Thanks to Guy Kahane and Simon Rippon for suggestions on how to express this point – and special thanks to the latter for suggesting the Phil Papers survey as a comparison.

philosophers' intuitions are relatively immune to diachronic instability. On this explanation, these studies *do* appear to undermine that claim.

However, I will now argue that this appearance is misleading. At best, it is *possible* that this explanation undermines the expertise defense, but this would require further empirical claims which are not in evidence. The question ultimately turns on whether or not philosophers are able to introspectively detect instability in their judgments, and whether the unstable intuitions in these studies belong to those philosophers.

Consider: some philosophers will *acknowledge* that their judgments about certain issues are diachronically unstable. In discussing the research targeted in this paper, I have witnessed several philosophers indicating that they are not surprised by the influence of order effects or similar irrelevant factors. These philosophers point out that the issues in question are exactly the sort about which people find themselves internally conflicted. The Trolley Problem is a knotty one; the Doctrine of Double Effect can sometimes seem plausible and sometimes not.²³ Given the internal conflict underlying such philosophers' reactions to these cases and principles, they say, it should be no surprise that manipulations like order of presentation can impact which reaction is experienced on any particular occasion.

So are these philosophers conceding that the expertise defense is mistaken? Not necessarily.

Remember, once again, that the expertise defense is only invoked for philosophical practice involving considered judgments. But if I *know* that my judgments about certain cases or principles are diachronically unstable, then I will not regard *these* as considered judgments. Considered judgments are specifically those that are *not* "made with hesitation" or "in which we have little confidence" (Rawls

²³ Thanks to both Shaun Nichols and Guy Kahane for correctly predicting that I would find philosophers ready to admit to diachronic instability on these issues if I asked around. (Though note that I did not necessarily say *they* are among that group).

1971, 47). Proponents of the expertise defense are not committed to the view that philosophical intuitions are reliable no matter how weakly or variably experienced (see Kauppinen 2007, 103-104; Williamson 2011, 219).

Some philosophers will acknowledge that their intuitions about certain cases or principles are diachronically unstable – and will not claim these intuitions as protected by the expertise defense! The question, then, must be: do the distorted intuitions detected in these studies belong to *those* philosophers? If so, the studies present no challenge to the expertise defense. But if not, a genuine challenge remains. In that case, it would appear that some philosophers who *think* they are employing considered judgments are unwittingly employing diachronically unstable intuitions, against which their expert training appears to have provided no defense.

So which is it? Recall that only a minority of philosophers were affected by the manipulations of these studies. As Schwitzgebel and Cushman concede, their results are “consistent with the possibility that a majority of philosophers adhere consistently to principles” (150). Were these minority of philosophers the ones who are *aware* of the instability of their intuitions (and who would not claim expertise in this case) or were they philosophers who deny instability? We do not know the answer. The subjects in these studies were not asked to indicate their confidence in these judgments, or whether they represented stable views.

But the question is not a complete empirical blank. There is some evidence that people *are* able to introspectively assess the reliability of their intuitions. In a recent study, Jennifer Cole Wright (2010) found that non-philosopher subjects’ confidence in their judgments about test cases tracked the stability of those judgments. Wright used an order effect manipulation similar to the one employed by Schwitzgebel and Cushman, and found that when subjects reported high confidence in a particular judgment, that judgment was less likely to be affected by order of presentation. Judgments in which the

subjects had low confidence were more likely to be affected. She interprets this as evidence that “people are able to introspectively track – and thus potentially protect against – their vulnerability to (at least some forms of) bias,” by monitoring their own confidence levels (Wright 2010, 500). Hence it is not implausible to think that philosophers are aware (indirectly at least) of their vulnerability to diachronic instability.²⁴

So this explanation is threatening to the expertise defense *only* if we can rule out the possibility that the fraction of distorted intuitions among the philosophers subjects come from philosophers who acknowledge the instability of their intuitions and thus would not claim the immunity of expertise (on these issues). Or, to put it the other way: the expertise defense is threatened only if it can be shown that distorted intuitions take place among philosophers who *deny* diachronic instability in their intuitions. At the moment, we have no such evidence.

4.6 Explaining away the findings

I have argued that the findings reviewed in section three are puzzling, because we have hypothesis-independent reason to expect not to find them. Specifically, it seems puzzling that philosophers would respond to familiar cases and principles in any way other than merely reporting their familiar responses, formed before the study’s manipulations could have any effect. I have allowed that this puzzle can be resolved, and the findings accepted, if we can offer some explanation for the familiarity problem. I have considered four such explanations. The first three do not present any threat to the expertise defense. Only the fourth allows such a threat – but requires further evidence we currently lack, and in fact goes against what evidence (Wright 2010) we have.

²⁴ Indeed, Wright herself suggests that “most moral philosophers and scientists already do this, treating clear/strong intuitions (especially their own) more severely than unclear/weak ones” (Wright 2010, 500). Grundmann (2010, 501) applies Wright’s data to the expertise debate in a similar way. But see also Zamzow and Nichols (2009, 374) for a word of caution on this point.

5. How to test for philosophical expertise

If the studies discussed above have not succeeded in testing the expertise defense, where does the debate over philosophical expertise stand? In my view, it remains an unsettled empirical matter. If we want to determine whether trained philosophers produce better intuitions, or make better use of intuitions, we still need to do experimental testing of these claims.

In effect, I claim that the substantive lesson of this paper is to restore the status quo ante, before the SC and TBS studies were published. Importantly, there is a live debate over what that status might be. Some philosophers (Horvath 2010, Williamson 2011) have argued that the burden of proof is on those who deny philosophical expertise, and therefore there is a presumption that the intuitive practice of philosophers remains reliable until clearly proven otherwise. On the other side, Weinberg and colleagues (2010) argue that there are theoretical reasons to doubt that philosophical training would generate the appropriate sort of expertise, and therefore expertise claims require positive empirical demonstration.

I do not take a position on these assertions about the burden of proof. This paper aims only to show that, however the burden is understood, the findings presented by TBS and SC do not resolve it. In this sense, the aim of this paper has not been to defend claims of philosophical expertise, but only to raise points of psychological best practice as applied to this debate.

Therefore I will conclude with some forward-looking suggestions for how we might continue to empirically investigate the expertise defense. A central lesson of my arguments has been this: to be confident that we are testing for philosophical expertise, we will have to be careful to avoid problems

with familiarity in our stimuli. It will not work to test philosophers and non-philosophers side-by-side on the same stimuli, if these stimuli are ones likely to be familiar to professional philosophers.²⁵

One seemingly obvious solution is to use *unfamiliar* stimuli – cases and principles philosophers will not recognize. Unfortunately, this is likely to be quite difficult, and may be impossible. In order for a psychological study to really test reasoning in a target domain, the stimuli must be organized according to the conceptual logic of that domain (Kahane and Shackel 2010). In effect, this means that the cases or principles must be constructed so as to respond to the logical points at issue in ongoing philosophical debates. For instance, the scenarios used will have to be isomorphic to standard cases in the philosophy literature, turning on such logical joints as means versus side-effect. (SC's cases are isomorphic to standard Trolley cases in exactly this way.) If they did not have this logical correspondence, then the psychological test would not have much relevance to philosophers' actual practice. However, if they *do* have this correspondence, then trained philosophers are quite likely to notice it, and so to treat the cases as familiar.

So if we cannot avoid familiar stimuli, what can we do? I suggest that the next step should be to focus on evidence for or against diachronic instability in intuitions. As I argued in section 4.5, the expertise defense is threatened only if it can be shown that philosophers' intuitions (a) change over time, (b) do so without philosophers' knowledge, and (c) belong to those philosophers who believe themselves to have expertise in the relevant domain.

The best way to establish that philosophers' intuitions change over time would be to conduct a longitudinal study, showing that the *same* philosophers, surveyed at two different times, exhibited differences in their reactions to test cases or principles. Longitudinal studies are a good deal harder to conduct than

²⁵ Grundmann (2010, 503) also objects to using familiar stimuli to test philosopher subjects, though on the different grounds that their responses will not be theory-neutral.

simple surveys, but a claim as ambitious as one questioning basic philosophical methodology certainly merits such care.

In addition, philosophers (and non-philosophers) responding to these studies should be asked to report their confidence in these judgments, as in Wright (2010). Response options should include “not sure” or “no settled view”, so that responses that are clearly not considered judgments can be excluded from analysis. Subjects should be *explicitly* asked if they believe their intuitions exhibit diachronic instability. They might even be asked if they believe that philosophical training provides immunity to distortion, and if they would claim this immunity themselves.²⁶

Combining these improvements would provide far stronger evidence regarding the expertise defense. A longitudinal study that demonstrated diachronic instability in the confidently-claimed considered judgments of professional philosophers would provide *exactly* the evidence that opponents of philosophical expertise require. If such evidence cannot be provided, then the expertise defense appears to be in quite good shape. Thus far (to my knowledge) no one has attempted to collect this data – and we have no way of knowing which way the evidence goes. So let the experiments continue!

References

²⁶ Thanks to Simon Rippon for helpful discussion of these ideas. A further option (quite different from that discussed in the text) would be to follow Schulz et al. (2011) and examine distorting effects that take place *outside* the experimental context. These studies did not *manipulate* subjects' intuitions, so the familiarity problem does not arise. Instead they tested for already-existing personality traits and correlated these to philosophical views (see also Arvan 2013). On the assumption that the truth of philosophical views does not depend upon a philosophers' personality traits, these findings would appear to undermine claims of philosophical expertise. But I leave this approach to the side here, as it is not as obvious that personality traits are 'distorting' in the same way as Actor-Observer and order effects. (See Zamzow and Nichols 2009 for reasons to tolerate or even favor personality-linked differences in philosophers' views.)

- Alexander, Joshua. 2012. *Experimental Philosophy: An Introduction*. 1st ed. Polity.
- Arvan, Marcus. 2013. "Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits: A Correlational Study." *Neuroethics* 6 (2): 307–18.
- Audi, Robert. 2008. "Intuition, Inference, and Rational Disagreement in Ethics." *Ethical Theory and Moral Practice* 11 (5): 475–92.
- Bealer, George. 2000. "A Theory of the a Priori." *Pacific Philosophical Quarterly* 81 (1): 1–8211.
- Bem, Daryl J. 2011. "Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect." *Journal of Personality and Social Psychology* 100 (3): 407–25. doi:10.1037/a0021524.
- Bengson, John. 2013. "Experimental Attacks on Intuitions and Answers." *Philosophy and Phenomenological Research* 86 (3): 495–532. doi:10.1111/j.1933-1592.2012.00578.x.
- Bennett, Craig M., Abigail A. Baird, Michael B. Miller, and George L. Wolford. 2010. "Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction." *Journal of Serendipitous and Unexpected Results* 1 (1): 1–5.
- Bourget, David, and David J. Chalmers. forthcoming. "What Do Philosophers Believe?" *Philosophical Studies*
- Cappelen, Herman. 2012. *Philosophy Without Intuitions*. Oxford University Press.
- Cullen, Simon. 2010. "Survey-Driven Romanticism." *Review of Philosophy and Psychology* 1 (2): 275–96. doi:10.1007/s13164-009-0016-1.
- Driver, Julia. 2006. *Ethics: The Fundamentals*. 1st ed. Wiley-Blackwell.
- Fischer, John Martin, and Mark Ravizza. 1992. *Ethics: Problems and Principles*. 1st ed. Harcourt Brace Jovanovich.
- Foot, Philippa. 1967. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* 5: 5–15.
- Gensler, Harry J. 2011. *Ethics: A Contemporary Introduction*. Taylor & Francis.
- Grundmann, Thomas. 2010. "Some Hope for Intuitions: A Reply to Weinberg." *Philosophical Psychology* 23 (4): 481–509. doi:10.1080/09515089.2010.505958.
- Horvath, Joachim. 2010. "How (not) to React to Experimental Philosophy." *Philosophical Psychology* 23 (4): 447–80. doi:10.1080/09515089.2010.505878.
- Jones, E. E., and R. E. Nisbett. 1971. "The Actor and the Observer: Divergent Perceptions of the Causes of Behavior." In *Attribution: Perceiving the Causes of Behavior*, edited by E. E. Jones, D. E. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins, and B. Weiner, 79–94. General Learning Press.
- Kagan, Shelly. 1997. *Normative Ethics*. Westview Press.
- Kahane, Guy, and Nicholas Shackel. 2010. "Methodological Issues in the Neuroscience of Moral Judgement." *Mind & Language* 25 (5): 561–82. doi:10.1111/j.1468-0017.2010.01401.x.
- Kamm, Frances M. 2000. "The Doctrine of Triple Effect and Why a Rational Agent Need Not Intend the Means to His End." *Aristotelian Society Supplementary Volume* 74 (1): 21–8211.
- Kauppinen, Antti. 2007. "The Rise and Fall of Experimental Philosophy." *Philosophical Explorations* 10 (2): 95–118.
- Kornblith, Hilary. 1998. "The Role of Intuition in Philosophical Inquiry: An Account with No Unnatural Ingredients." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Theory*, edited by Michael R. Paul and William Ramsey, 129–41. New York: Rowham and Littlefield.
- Lanteri, Alessandro, Chiara Chelini, and Salvatore Rizzello. 2008. "An Experimental Investigation of Emotions and Reasoning in the Trolley Problem." *Journal of Business Ethics* 83 (4): 789–804. doi:10.1007/s10551-008-9665-8.

- LeBel, Etienne P., and Kurt R. Peters. 2011. "Fearing the Future of Empirical Psychology: Bem's (2011) Evidence of Psi as a Case Study of Deficiencies in Modal Research Practice." *Review of General Psychology* 15 (4): 371–79. doi:10.1037/a0025172.
- Liao, S. Matthew. 2009. "The Loop Case and Kamm's Doctrine of Triple Effect." *Philosophical Studies* 146 (2): 223–31. doi:10.1007/s11098-008-9252-y.
- Liao, S. Matthew, Alex Weigmann, Joshua Alexander, and Gerard Vong. 2011. "Putting the Trolley in Order: Experimental Philosophy and the Loop Case." *Philosophical Psychology* 25 (5): 661–71.
- Lombrozo, Tania. 2009. "The Role of Moral Commitments in Moral Judgment." *Cognitive Science* 33 (2): 273–86. doi:10.1111/j.1551-6709.2009.01013.x.
- Ludwig, Kirk. 2007. "The Epistemology of Thought Experiments: First Person Versus Third Person Approaches." *Midwest Studies In Philosophy* 31 (1): 128–59. doi:10.1111/j.1475-4975.2007.00160.x.
- Machery, Edouard, Ron Mallon, Shaun Nichols, and Stephen P. Stich. 2004. "Semantics, Cross-Cultural Style." *Cognition* 92 (3).
- Nadelhoffer, Thomas, and Adam Feltz. 2008. "The Actor–Observer Bias and Moral Intuitions: Adding Fuel to Sinnott-Armstrong's Fire." *Neuroethics* 1 (2): 133–44.
- Nado, Jennifer. 2012. "Why Intuition?" *Philosophy and Phenomenological Research*, n/a–n/a. doi:10.1111/j.1933-1592.2012.00644.x.
- Nagel, Thomas. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.
- Petrinovich, Lewis, and Patricia O'Neill. 1996. "Influence of Wording and Framing Effects on Moral Intuitions." *Ethology & Sociobiology* 17 (3): 145–71. doi:10.1016/0162-3095(96)00041-6.
- Quinn, Warren S. 1989. "Actions, Intentions, and Consequences: The Doctrine of Double Effect." *Philosophy and Public Affairs* 18 (4): 334–51.
- Rawls, John. 1951. "Outline of a Decision Procedure for Ethics." *The Philosophical Review* 60 (2): 177–97. ———. 1971. *A Theory of Justice*. 1st ed. Cambridge, MA: Harvard University Press.
- Rini, Regina A. 2014. "Analogies, Moral Intuitions, and the Expertise Defence." *Review of Philosophy and Psychology* 5 (2): 169–81. doi:10.1007/s13164-013-0163-2.
- Ryberg, Jesper. 2013. "Moral Intuitions and the Expertise Defence." *Analysis* 73 (2): 3–9. doi:10.1093/analys/ans135.
- Schulz, Eric, Edward T Cokely, and Adam Feltz. 2011. "Persistent Bias in Expert Judgments About Free Will and Moral Responsibility: a Test of the Expertise Defense." *Consciousness and Cognition* 20 (4): 1722–31. doi:10.1016/j.concog.2011.04.007.
- Schwitzgebel, Eric. 2009. "Do Ethicists Steal More Books?" *Philosophical Psychology* 22 (6): 711–25. doi:10.1080/09515080903409952.
- Schwitzgebel, Eric, and Fiery Cushman. 2012. "Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers." *Mind and Language* 27 (2): 135–53.
- Schwitzgebel, Eric, Joshua Rust, Linus Ta-Lun Huang, Alan T. Moore, and Justin Coates. 2012. "Ethicists' Courtesy at Philosophy Conferences." *Philosophical Psychology* 25 (3): 331–40. doi:10.1080/09515089.2011.580524.
- Shafer-Landau, Russ. 2009. *The Fundamentals of Ethics*. Oxford University Press, USA.
- Singer, Peter. 1972. "Moral Experts." *Analysis* 32 (4): 115–17. doi:10.2307/3327906.
- Sinnott-Armstrong, Walter. 2008. "Framing Moral Intuition." In *Moral Psychology, Vol 2. The Cognitive Science of Morality: Intuition and Diversity*, 47–76. Cambridge, MA: MIT Press.
- Smart, J. J. C., and Bernard Williams. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.
- Sosa, Ernest. 2007. "Experimental Philosophy and Philosophical Intuition." *Philosophical Studies* 132 (1): 99–107.

- . 2010. "Intuitions and Meaning Divergence." *Philosophical Psychology* 23 (4): 419–26. doi:10.1080/09515089.2010.505859.
- Swain, Stacey, Joshua Alexander, and Jonathan M. Weinberg. 2008. "The Instability of Philosophical Intuitions: Running Hot and Cold on Trolleytemp." *Philosophy and Phenomenological Research* 76 (1): 138–55.
- Thomson, Judith Jarvis. 1976. "Killing, Letting Die, and the Trolley Problem." *The Monist* 59 (2): 204–17.
- . 2008. "Turning the Trolley." *Philosophy and Public Affairs* 36 (4): 359–74.
- Tobia, Kevin, Wesley Buckwalter, and Stephen Stich. 2013. "Moral Intuitions: Are Philosophers Experts?" *Philosophical Psychology* 26 (5): 629–38. doi:10.1080/09515089.2012.696327.
- Unger, Peter. 1996. *Living High and Letting Die: Our Illusion of Innocence*. Oxford University Press.
- Weinberg, Jonathan M., and Joshua Alexander. 2014. "The Challenge of Sticking with Intuitions Through Thick and Thin." In *Intuitions*, edited by A. Booth and D. Rowbottom, 187–212. Oxford University Press.
- Weinberg, Jonathan M., Chad Gonnerman, Cameron Buckner, and Joshua Alexander. 2010. "Are Philosophers Expert Intuiters?" *Philosophical Psychology* 23 (3): 331–55. doi:10.1080/09515089.2010.490944.
- Weinberg, Jonathan M., Shaun Nichols, and Stephen Stich. 2001. "Normativity and Epistemic Intuitions." *Philosophical Topics*, 29 (1-2): 429–60.
- Wiegmann, Alex, Yasmina Okan, and Jonas Nagel. 2012. "Order Effects in Moral Judgment." *Philosophical Psychology* 25 (6): 813–36. doi:10.1080/09515089.2011.631995.
- Williams, Bernard. 1982. *Moral Luck*. Cambridge University Press.
- Williamson, Timothy. 2004. "Philosophical 'Intuitions' and Scepticism About Judgement." *Dialectica* 58 (1): 109–8211.
- . 2007. *The Philosophy of Philosophy*. Wiley-Blackwell.
- . 2011. "Philosophical Expertise and the Burden of Proof." *Metaphilosophy* 42 (3): 215–29. doi:10.1111/j.1467-9973.2011.01685.x.
- Wright, Jennifer Cole. 2010. "On Intuitional Stability: The Clear, the Strong, and the Paradigmatic." *Cognition* 115 (3): 491–503. doi:10.1016/j.cognition.2010.02.003.
- Zamzow, Jennifer L., and Shaun Nichols. 2009. "Variations in Ethical Intuitions." *Philosophical Issues* 19 (1): 368–88. doi:10.1111/j.1533-6077.2009.00164.x.