CrossMark

Philosophical expertise under the microscope

Miguel Egler[1] · Lewis D. Ross[1]

Abstract Recent experimental studies indicate that epistemically irrelevant factors can skew our intuitions, and that some degree of scepticism about appealing to intuition in philosophy is warranted. In response, some have claimed that philosophers are experts in such a way as to vindicate their reliance on intuitions—this has become known as the 'expertise defence'. This paper explores the viability of the expertise defence, and suggests that it can be partially vindicated. Arguing that extant discussion is problematically imprecise, we will finesse the notion of 'philosophical expertise' in order to better reflect the complex reality of the different practices involved in philosophical inquiry. On this basis, we offer a new version of the expertise defence that allows for distinct types of philosophical expertise. The upshot of our approach is that wholesale vindications or rejections of the expertise defence are shown to be unwarranted; we must instead turn to local, piecemeal investigations of philosophical expertise. Lastly, in the spirit of taking our own advice, we exemplify how recent developments from experimental philosophy lend themselves to this approach, and can empirically support one instance of a successful expertise defence.

Keywords Philosophical methodology · Thought experiments · Expertise defence · Intuitions · Experimental philosophy

✉ Miguel Egler
    me49@st-andrews.ac.uk

[1]   Arché Research Centre, University of St Andrews, 17–19 College Street, St Andrews, Fife KY16 9AL, Scotland, UK

⚛ Springer

1 Introduction

Recent studies in experimental philosophy suggest that intuitions about philosophically interesting concepts are subject to surprising variation and marked susceptibility to cognitive biases. From these findings, some philosophers (henceforth 'Critics') conclude that appealing to intuitions for the purposes of philosophical inquiry is methodologically problematic, and that at least *some degree* of scepticism about this practice is in order. In response, others (henceforth 'Defenders') have claimed that philosophers are *experts* in such a way as to vindicate their reliance on intuitions—this has become known as the 'expertise defence'.[1]

Our paper explores the viability of this defence, and argues that there is empirical evidence to suggest that it can be partially vindicated. Central to our view is the proposal that extant discussions of the expertise defence are unsatisfactory because they do not account for important methodological distinctions amongst practices of philosophical inquiry. To set up our argument, we first briefly review the two most prominent approaches to formulating the expertise defence and show how they operate with a problematic assumption: namely, that the expertise defence either stands or falls *wholesale* (Sect. 2). We argue that the many ways in which intuitions are put to use in philosophical inquiry engender distinct *types* of philosophical expertise, and that wholesale defences or rejections of the expertise defence are thereby unwarranted–instead, we argue that local, piecemeal investigations of philosophical expertise are better suited to assess the expertise defence (Sect. 3). In more detail, we suggest that the expertise defence must be sensitive to the cognitive underpinnings of intuitions appealed to within philosophy (Sect. 3.1), and to different implementations of thought-experiments (Sect. 3.2), and argue that neglecting these distinctions significantly inhibits our thinking about the expertise defence (Sect. 3.3). Finally, in the spirit of taking our own advice, we demonstrate how recent developments in experimental philosophy (viz. *cognitive epistemology*) enable a piecemeal assessment of philosophical intuitions and, crucially, provides empirical support for at least one instance of a successful expertise defence (Sect. 4).

2 Current approaches to the expertise defence

The expertise defence is primarily motivated by the observation that most experimental critiques of intuitions rely on data gathered from studying non-philosophers. Defenders point out that we might naturally think philosophers are just better equipped than non-philosophers to engage in philosophical inquiry. If so, then the force of many experimental critiques rests on an illicit inference: that failures affecting the judgements of non-philosophers generalise to the judgements of *expert* philosophers.

---

[1] This is not the only way that philosophers have responded to Critics. Another prominent line of response can be found in Cappelen (2012), Deutsch (2015) who argue that intuitions play no important evidential role in philosophy and that, therefore, experimental challenges have little to no relevance for philosophical theorising. We set these critiques aside here (for a response: Nado 2016); we assume that intuitions do play a role in philosophy, and that defenders of appeals to intuitions in philosophy must grasp the nettle and respond to experimental critiques.

Although this response has a ring of plausibility, it requires considerable precisification. We will organise our discussion around two independent claims that any version of the expertise defence needs to establish:

1. Philosophers have *philosophical expertise*—adequate training in philosophy leads to improved performance in some task(s).
2. Philosophical expertise *defends* philosophers from distorting factors that Critics claim undermines the practice of appealing to intuitions in philosophy.

So, what could the nature of the expertise be that philosophical training brings about, and how might it help philosophers avoid deleterious biases? In this section, we review and critique two natural responses to this question. Firstly, we outline and reject the 'expert intuitions' defence—the view that locates philosophical expertise in the superior intuitions of philosophers. Secondly, we outline and critique the more plausible 'expert practices' defence, which locates philosophical expertise in the superior methods employed by philosophers.

2.1 Expert intuitions

Call the suggestion that philosophical expertise yields superior intuitions the *expert intuition defence*. In line with the criteria outlined above, this amounts to the following proposal:

> **Expert Intuition Defence** Philosophical training leads philosophers to have intuitions that are substantially less susceptible to the distorting effects influencing the intuitions of non-philosophers.

There are several ways to flesh out this basic suggestion. For instance, Kauppinen (2007) proposes that philosophers bring a distinctive kind of *reflective* intuition to bear on philosophical cases. Ludwig (2007) argues that philosophers are apt to provide intuitions that stem from *conceptual competence* alone, rather than responses influenced by other factors (such as pragmatic implicatures). Both versions of Expert Intuitions contrast the intuitions at play in philosophical discourse with the quick, spontaneous responses provided by non-philosophers to experimental surveys. Their hope is that the more refined intuitions of philosophers respect fine-grained distinctions elided by the folk, whilst stripping out the influence of various irrelevant factors and routine inattentiveness to cases.

Although it might seem promising to deny the symmetry between the intuitions of philosophers and quick judgements participants provide in experimental surveys, Expert Intuitions Defence is a difficult position to maintain. Firstly, there is substantial empirical evidence that many of our judgements in a variety of domains are susceptible to errors due to the effect of cognitive biases.[2] If the influence of biases are a pervasive part of human psychology, susceptibility to bias in philosophical discourse—like in other domains—is a fair default assumption. Furthermore, Critics have already gathered a small but growing body of recent evidence suggesting that trained philosophers

---

[2] However, it is crucial to emphasise that not all biases have negative distorting effects (Antony 2016). In this paper we restrict use of the term 'bias' to those negative distorting effects.

display no particular immunity to cognitive biases (for reviews, see: Machery 2015; Mizrahi 2015). Without countervailing empirical evidence, it is hard to see how we could be justified in holding the sort of exceptionalism Defenders wish to attribute to professional philosophers.[3] Biases are prevalent in all sorts of judgements and specific measures are required to ameliorate their influence; believing against the empirical evidence that philosophers' intuitions are somehow immune to these biases seems to be an article of faith rather than a well-supported position.

2.2 Expert practices

The issues facing Expert Intuitions have led some to conclude that preoccupation with intuitions *simpliciter* is misguided. Jennifer Nado succinctly captures this worry:

> [The] focus on intuition is somewhat odd, given that it is in no way essential to the expertise defence; philosophical methodology obviously consists in much more than the gathering of intuitive judgements, and there are therefore many potential loci for philosophical expertise.
>
> (Nado 2014a, p. 1029)

Nado goes on to suggest that philosophers "display expertise in assessing, critiquing, and (when appropriate) rejecting or explaining away" the data gathered by the method of cases (Nado 2014a, p. 1041).[4] Williamson (2011) also shows dissatisfaction with the narrow focus on intuitions; in this sense, he characterises philosophical expertise as the improved performance in the *use* of thought experiments in philosophical inquiry—by which he means the evaluation, assessment, and deployment of them in theorising (Williamson 2011, pp. 224–225).[5] And there are interesting yet brief comments along similar lines to be found in Rini (2015, §3). We take these comments to suggest philosophical expertise is to be found in what philosophers *do* with intuitions. As an attempt to provide a rough locus for our discussion that remains neutral on various metaphilosophical positions, we will characterise the variety of philosophical practices that might be promising for the expertise defence as follows:

> **Philosophical Practices** Philosophers elicit, invoke, assess, and synthesise intuitions about philosophical cases.[6]

Attempts to locate philosophical expertise in what philosophers *do* with intuitions is a significant improvement on previous intuition-focused proposals of the expertise defence: we agree with Nado that there is more to philosophy than simple reliance on

---

[3] See also Weinberg et al. (2012).

[4] It is noteworthy that Nado endorses an account of philosophical expertise that does not support the expertise defence.

[5] Neither Nado nor Williamson phrase their accounts in terms of 'intuitions', arguing that intuition-talk is theoretically inert and often leads to confusion. However, our use of 'intuitions' here is roughly equivalent to 'judgements elicited from philosophical cases', which both Nado and Williamson refer to in their accounts.

[6] We borrow this description of how philosophers use intuitions in philosophy from Fischer and Collins (2015, p. 11).

intuition, and that this broader nexus of practices must be taken into account when examining the nature of philosophical expertise.

However, we argue that extant formulations of this approach remain problematic. Specifically, we suggest that these attempts to refine the locus of philosophical expertise are problematically *monolithic*. By this, we mean that current approaches imply that the philosophical expertise relevant to **Philosophical Practices** consists of: (i) *a single package of* (ii) *general philosophical skills* that are (iii) *shared by all/most philosophers*, and are (iv) *apt to be successfully deployed across a range of philosophical contexts*.[7] This monolithic view has important ramifications for how we think about the expertise defence which, recall, is roughly composed of two claims: (1) Philosophers have *philosophical expertise*—adequate training in philosophy leads to improved performance in some task(s); and (2) Philosophical expertise *defends* philosophers from distorting factors that critics claim undermines the practice of appealing to intuitions in philosophy. The monolithic account of philosophical expertise currently implied in the literature leaves us with two basic possibilities:

> **Expert-Practices Defence** Philosophers display improved performance in the practices of eliciting, invoking, assessing, and synthesizing intuitions for purposes of philosophical inquiry. Moreover, *this suffices* to shield philosophers from mistakes and problems afflicting non-philosophers when appealing to intuitions in philosophical inquiry.
> **No Expert-Practices Defence** Philosophers display improved performance in the practices of eliciting, invoking, assessing, and synthesizing intuitions for purposes of philosophical inquiry. However, *this does not suffice* to shield philosophers from mistakes and problems afflicting non-philosophers when appealing to intuitions in philosophical inquiry.

These two options capture the tendency—prevalent in the literature—to either defend or refute the expertise defence *simpliciter*. That is, both sides in the debates about the expertise defence advance accounts which say that, whatever the relevant form of philosophical expertise consists in, it either *univocally does* or *does not* support the expertise defence (Williamson 2007, 2011; Kauppinen 2007; Ludwig 2007; Weinberg et al. 2010; Tobia et al. 2013; Nado 2014a; Machery 2015; Rini 2015).

We propose that any such extreme pessimism or optimism is unwarranted. Specifically, we will argue that this way of construing the expertise defence rests on an inadequate conception of philosophical expertise that neglects important distinctions

---

[7] There has been some recognition in the literature of the need to distinguish between types of 'philosophical expertise' when investigating the expertise defence. For instance, Weinberg et al. (2010, p. 335) recognise that there is an important distinction to be drawn between improved performance in "close analysis of texts, or the critical assessment of arguments" and improved performance "at conducting thought experiments". In a similar vein Nado (2014a, p. 1041) distinguishes expertise in the "construction of thought experiments" and reflecting on philosophical theories, viz. "teasing out their consequences, determining their compatibility with other positions, and so forth." After drawing such distinctions, Nado and Weinberg are concerned with asking whether these types of expertise might plausibly shield philosophers from bias; on this front, they find little grounds for optimism. However, when philosophical expertise is put in this way, the expert practices defence begins to look rather unpromising. After all, even if philosophers are better than the average person at analysing texts, or coming up with intriguing thought-experiments, this doesn't seem apt to protect them against the multiplicity of biases identified by Critics.

among tasks encompassed by **Philosophical Practices**. We advance an alternative option, not identified in the literature, that aims to be more faithful to the complex reality of such practices. The main thrust of our proposal is the idea that the relation between philosophical expertise and the expertise defence is considerably more intricate than the binary choice between **No Expert-Practices Defence** and **Expert-Practices Defence**. More precisely, we argue for the following:

> **Expert-Practices Defences** *Some* philosophers display improved performance in *certain sets of practices* in eliciting, invoking, assessing, and synthesizing intuitions for purposes of philosophical inquiry. Moreover, this suffices to shield *those* philosophers from *specific* mistakes and problems afflicting non-philosophers when appealing to intuitions in philosophical inquiry.

In the following section, we motivate **Expert-Practices Defences** by arguing that the diversity of practices described under **Philosophical Practices** engender distinct *types* of philosophical expertise. In this sense, we suggest that philosophical expertise in such practices consists of: (i) *a diverse package* of (ii) *specific philosophical skills* that are (iii) *shared by some philosophers*, and are (iv) *apt to be successfully deployed only in a narrow range of philosophical contexts*. To articulate this view, we detail two methodologically important distinctions that are overlooked in current discussions of philosophical practices: first, between the types of cognitive processes underlying philosophical intuitions, and second, among the different ways in which thought-experiments can be put to use in philosophical inquiry. We then demonstrate why eliding these distinctions fails to enable an informative evaluation of the expertise defence.

3 Finessing the expertise defence

3.1 Cognitive processes

In recent years, there has been a growing consensus that intuitions do not form a homogenous class of mental states.[8] This claim is rather plausible: there do seem to be significant differences between, for instance, intuitive judgements about whether a subject knows some proposition, and whether pushing people in front of trains for the greater good is morally permissible. Given the superficial dissimilarities between these intuitions, it is quite natural to wonder whether the practices subsumed under **Philosophical Practices** will fare equally well with respect to each of these intuitions. We will argue that they do not, and that discussion of philosophical practices would benefit from taking into account the differences between distinct intuitions.

Recent developments in experimental philosophy provide empirical evidence suggesting that philosophical inquiry should distinguish between different types of philosophical intuitions. The studies in question employ techniques and findings from the empirical sciences with the aim of identifying the particular *cognitive processes*

---

[8] Nado (2014b) provides a very clear discussion of these points.

underlying our philosophical intuitions.[9] For current purposes, we need only detail developments from two such lines of inquiry.

The first concerns recent proposals that seek to trace back a certain class of epistemic intuitions to the operation of a set of specific cognitive processes. The epistemic intuitions in question concern intuitive knowledge-attributions: judgements about what other people may or may not know under certain circumstances—e.g., whether people can know the time from a stopped clock, even if it is only by sheer luck that the clock indicates the correct time at that very moment. In a set of recent papers, Nagel (2012a, b), Boyd and Nagel (2014) offers a psychological explanation of these intuitive knowledge-attributions. Making use of a large body of findings from cognitive psychology, Nagel proposes that these intuitions stem from the exercise of the ordinary psychological capacities for 'mind-reading'—i.e., a set of automatic cognitive processes that humans routinely rely on in order to understand and predict the mental states of others.

The second strand of research we wish to call attention to concerns recent attempts to identify the cognitive underpinnings of intuitions about whether inflicting harm can ever be morally permissible, if doing so would bring about some greater good. Most empirical inquiry into the nature of these intuitive judgements seeks to pinpoint the specific processes driving people's responses to the so-called 'Trolley Problem'—i.e., cases where people are asked if it would be morally permissible to in some way alter the course of a run-away trolley in order to save many people, at the cost of sacrificing the life of one person (for a discussion: Thomson 1985). A number of empirical investigations into the nature of these intuitions provide robust neuroscientific evidence suggesting that the intuition according to which it would *not* be morally permissible to alter the course of the trolley—thus sacrificing one to save the many—are driven by a *negative emotional response* to this case (for a review: Greene 2015, pp. 700–706). Thus, these studies suggest that our verdicts in response to the 'Trolley Problem' are shaped by the affective processes we bring to bear when considering it.

These empirical findings provide evidence that there are in fact significant dissimilarities between the many types of intuitions that philosophers appeal to in philosophical inquiry. Specifically, they illustrate how different intuitions can be driven by distinct cognitive processes. Now, one may ask why any of this is relevant for an evaluation of the expertise defence. Here's why: these discoveries reveal how different intuitions are susceptible to distinctive kinds of *biases*. To explain, let's first turn to empirical evidence of the biases afflicting intuitive knowledge-attributions.

One hotly debated topic in contemporary epistemology concerns whether the standards for knowledge are sensitive to the error possibilities that are made salient in a given context (DeRose 1992; Cohen 1999). A central issue driving these debates is the observation that people are usually more willing to attribute knowledge to others when no possibilities of error are made salient; however, once such possibilities are raised, most people feel the temptation to retract such attributions. This amounts to a startling pattern of judgements, given that the orthodoxy has it that merely mentioning the *possibility* of error should not shift the standards for knowledge. Building

---

[9] Indeed, as Knobe (2016) suggests, most contemporary work in experimental philosophy is practised as cognitive science.

on psychological explanations of intuitive knowledge attributions in terms of 'mind-reading' capacities, Nagel (2012b) makes a move for the orthodox view. Specifically, she argues that these patterns of intuitive judgement are due to a specific *psychological bias* known as *epistemic egocentrism*, or the *curse of knowledge*. Simplifying, this bias amounts to a tendency to wrongly predict the mental states of other subjects who stand in epistemic positions that are more naive than one's own. In these situations, most people erroneously represent others' perspectives as if they shared their own privileged epistemic standing—and, moreover, they tend to penalise others (e.g., retract knowledge attributions) for not responding to the evidence as they think appropriate from their privileged epistemic standing. Subsequent empirical investigations have borne out the predictions of this psychological explanation, providing additional evidence that changes in intuitive knowledge-attributions in the light of possibilities of error can indeed be explained in terms of the egocentrism bias (Alexander et al. 2015).

Let us now turn to a different set of debates: namely, discussions about the Doctrine of Double Effect (DDE)—i.e., the principle which states that it is sometimes permissible to bring about harm as a *side-effect* of aiming to bring about some good, even though it would be impermissible to bring about the same harm as a *means* of bringing about the same good. One of the central lines of support for the DDE is that it neatly predicts typical responses to certain complex morally charged scenarios. For instance, this principle nicely captures one common response-pattern to two variants of the Trolley-Problem:[10]

> **Lever** A runaway trolley is headed toward five innocent people who are on the track and who will be killed unless something is done. Patrick may pull a lever, which will redirect the trolley onto a second track, saving the five people. However, on this second track is an innocent bystander (Susy), who will be killed if the trolley is turned onto this track.
>
> **Push** A runaway trolley is headed toward five innocent people who are on the track and who will be killed unless something is done. Patrick can push an innocent bystander (Susy) in front of the trolley. The runaway trolley would be stopped by hitting (and killing) Susy, but would thereby save the five people on the track.

According to the DDE, it would (i) be permissible to stop a train from killing five people by pulling a lever to *divert* the train onto a second track, leading to the death of one person, but (ii) it would not permissible to *push* one person onto the track—thereby killing them, but stopping the train—in order to prevent the train from killing five people. Indeed, this is precisely the pattern of responses most people have when they evaluate these cases. But here's the rub: studies show that people can be primed to give different judgements to **Lever** and **Push**, simply by modifying the order in which these cases are presented to them (Schwitzgebel and Cushman 2012, 2015). For instance, if one is presented with the **Push** case first then you are more likely to judge the subsequent **Lever** case to be impermissible. One prominent explanation for why people are susceptible to such order effects is the operation of an affective bias: one's judgements are being prompted by an unreflective consideration of one's emotional state. By using one's immediate emotional reaction as a benchmark, we become

---

[10] These formulations are modified from Liao et al. (2012).

disposed to changes in intuitive response depending on whether we are presented with an emotionally 'potent' variation of the thought-experiment (such as pushing a person to their death) in the first instance or later on (for a discussion: Greene 2015).

What is the upshot of all this? It shows that particular biases and errors afflicting certain intuitions are not likely to arise for others. And as a result, different skills will be required to successfully appeal to an expertise defence within some lines of philosophical inquiry than in others. For instance, an expertise defence for philosophers working specifically on the Trolley-problem will require them to deploy practices that can shield them from certain affective biases[11] (in order to either prevent the biases, or to ameliorate their effects on their intuitions) whilst another philosopher working on the relationship between knowledge and stakes will have to be cautious regarding biases such as epistemic egocentrism. These are skills that they require in virtue of their lines of inquiry engaging different cognitive processes.[12] Therefore, asking whether expert philosophical practices suffice for an expertise defence cannot be carried out in an intuition-general way; the sorts of biases one is susceptible to in a given line of inquiry, and the skills and practices required to offset these biases, will depend on the cognitive processes producing one's intuitions.

To take stock, we have briefly documented recent developments in experimental philosophy indicating that our intuitions are underwritten by very distinct kinds of cognitive processes. Furthermore, we explained how these psychological explanations of our intuitions reveal them to be vulnerable to a variety of distinct biases. We then argued that these distinctions are methodologically significant, insofar as they suggest that distinct skills are required to offset biases afflicting different intuitions. These findings provide an initial attack on the monolithic interpretation of philosophical expertise—which holds that expertise in the tasks under **Philosophical Practices** amount to one set of general skills.

In what follows, we underscore yet another important methodological distinction that the monolithic interpretation fails to account for—viz., between different ways of implementing thought-experiments in philosophical inquiry—and outline its relevance for assessing the expertise defence. We then conclude this section by discussing an alternative approach to delineating philosophical expertise that can accommodate for these distinctions, and which allows for a more informative evaluation of the expertise defence.

---

[11] Of course, this isn't to say the ethical judgements should be free from affective influences *tout court*. We only make the uncontroversial point that, since order effects are clearly undesirable, ethicists must ensure that affective influences do not leave them vulnerable to such a deleterious bias.

[12] Again, this distinction cannot neatly be drawn by looking at *areas* of philosophy. For instance, it is noteworthy that some studies have found that considerations with moral valence can influence knowledge attributions (Beebe and Buckwalter 2010). This implies that it is at least possible that affective processes may shift epistemic attributions, too. These initial results further support our suggestion that our focus should be on the specific cognitive processes our inquiries engage, not on the philosophical domain we take ourselves to be working in.

3.2 Different uses of thought-experiments

Philosophers use thought-experiments in a variety of different ways. However, these distinct uses have not received equal scrutiny in the methodological literature. Here are three rather different ways that

Standard Uses Firstly, and most prominently, philosophers use thought-experiments to *elicit judgements which are used to defend or refute claims* about concepts, folk theories, and/or the phenomena such concepts and folk theories are about. These uses of thought-experiments, what we are calling **Standard** uses, are widespread and familiar within the philosophical literature. The canonical examples are Gettier-cases. For most people, considering the details of Gettier-like scenarios cues a particular judgement—namely that subjects lack knowledge in certain 'lucky' conditions—and this judgement, in turn, is frequently used to defend a negative answer to the philosophical question 'is knowledge reducible to justified true belief?' Other famous examples of **Standard** uses of thought-experiments include Kripke's (1980) use of the Gödel case against descriptivist theories, Parfit's (1984) use of fission cases to argue against certain theories of personal identity, and Jackson's (1982) use of the Mary case to argue against Physicalism. In all these cases, the application of the thought-experiment seeks to elicit specific judgements, and then uses these judgements to *answer* some philosophical question.

Aporetic Uses Secondly, some thought-experiments are used in order to elicit *conflicting judgements* about the same scenario. We will call these **Aporetic** uses of thought experiments, to indicate that they are meant to give rise to a kind of *puzzlement* that helps to refine inquiry into their subject-matter.[13] A simple example is the case of the ship of Theseus: "is a ship that has had all of its components gradually replaced on the course of a voyage the same ship on arrival as on departure?" Most people feel the pull of both the positive and negative answers: in one sense, it is indeed the same ship; but, in another, it seems that by substituting all of its parts we now have an entirely different ship (for discussions: Smart 1972; Scaltsas 1980; Lowe 1983). This thought-experiment thus gives rise to a sense of puzzlement about the notion of identity, which can then be put to use in refining inquiry. For instance, if it turns out that the intuitive pull of each answer corresponds to a different concept of identity, then philosophers will have to hone their inquiry accordingly. Newcomb's problem is another instance of an **Aporetic** use of a thought-experiment. This is a puzzle that elicits two plausible yet inconsistent judgements about rational attitudes (Nozick 1969).[14] Reflection on this problem has led to animated (and ongoing) discussions about the source of the puzzlement; it has promoted philosophers to draw novel distinctions and formulate questions

---

[13] One philosopher who has discussed the paradoxical nature of certain thought-experiments is Sorensen (1998).

[14] As a recent poll suggests, people tend to diverge almost evenly on both judgements: from 31,854 votes, 53.5% chose one judgement and 46.5% chose the other (https://www.theguardian.com/science/alexs-adventures-in-numberland/2016/nov/30/newcombs-problem-which-side-won-the-guardians-philosophy-poll).

that attempt make inquiry into rational attitudes more precise. (For instance, New-comb's paradox shows that there are cases where the decision-theoretic principle of dominance gives advice that conflicts with the principle of expected utility-maximization; this is a surprising result as it dissociates our judgements about rationality from the predictions of standard decision-theory.) In sum, **Aporetic** uses of thought-experiments significantly improve philosophical inquiry, to the extent that they tease out important distinctions that help us to better understand central philosophical notions.

Critical Uses A third way to use thought-experiments is to invoke particular intu-itions in order to scrutinise which aspects of the specified case elicit them. Insights into which features of a case are responsible for eliciting certain intuitions can then be used in assessing the evidential value of those intuitions. **Critical** uses of thought-experiments are not predominant in the philosophical literature, yet they are arguably amongst the most discerning. One such example is Gendler's (2010) close scrutiny of a range of thought-experiments about personal identity. From this detailed analysis, she concludes that some of the most prominent thought-experiments used in the literature are just too far divorced from the situations in which we ordinarily apply the concept of personal identity. Her central contention is that our intuitions about such cases are thereby unlikely to track features essen-tial to the concept of personal identity, making them useless for philosophical theorising. Another example of a **Critical** use is Jackson's (2003) *later* appeal to the case of Mary the Neuroscientist, where he attempts to diagnose *why* this case tends to elicit the intuition that Mary does indeed learn something new when she sees the red rose for the first time. Jackson invokes this intuition as a foil for his argument, which aims to show that this judgement turns on a mistaken conception of the nature of perceptual experience—viz., one which erroneously characterises experience in terms of *what it's like*. On these grounds, he proposes that we ought to reject this intuition, and that the case of Mary the Neuroscientist is inadequate for philosophical purposes.[15]

These three different ways to conduct philosophical inquiry have not—as far as we know—been separated out in the literature on the expertise defence. Critics have over-whelmingly focused on **Standard** uses, marshalling experimental evidence in order to argue that the practice of appealing to intuitions in this way is massively susceptible to truth-irrelevant factors. However, separating out different uses of thought-experiments reveals there to be a number of *different* practices that philosophers use in dealing with intuitions about cases. In other words, there is no *single* practice of eliciting, invoking, assessing and synthesising intuitions. The distinctions between **Standard**, **Aporetic**, and **Critical** illustrates at least some differences in how these practices can be put to use in philosophical inquiry.

The relevance of these distinctions is that these practices may have different epis-temic profiles with respect to the influence of deleterious biases. To see this, consider

---

[15] Another prominent example of a **Critical** use of thought-experiments is Dennett's (1991, pp. 437–440) critical examination of Searle's Chinese Room scenario. Additionally, the recent literature in experimental philosophy provides examples of empirically-informed **Critical** uses of thought-experiments: e.g., Nagel (2011), Fischer (2014) and Alexander et al. (2017).

how different uses of thought-experiment are used to inquire into the same philosophical issue—the nature, and perhaps the varieties of, personal identity. For example, Parfit's fission cases can be used as an argument for a particular view within debates about personal identity, and Gendler's use of various thought-experiments was used to reject the relevance of certain intuitions for theorising about personal identity. All of these turn out to be cases of eliciting, invoking, assessing and synthesising judgements about cases, but each for rather different purposes. Even if it were the case that Critics could adduce empirical evidence to demonstrate that **Standard** usages within this inquiry are susceptible to the influence of truth-irrelevant factors (e.g. order effects, affective biases, etc.), this would not be evidence that the **Critical** use of this thought-experiments are similarly susceptible. After all, it is plausible that **Critical** implementations of these thought-experiments could pick up on exactly these biases.

Although this in itself is an interesting hypothesis for the dialectic surrounding the expertise defence, more will be required to push back against the experimental Critic. Specifically, it will need to be shown that certain uses of thought-experiments can in fact *shield* philosophers from the effect of biases and truth-irrelevant factors. In the final section, we will discuss recent empirical evidence that suggests that a **Critical** use of a particular thought-experiment, applied to one particular line of inquiry, dos indeed succeed in uncovering and shielding philosophers from the effect of bias and truth-irrelevant influences that have been found to affect the **Standard** use of that same thought-experiment in that line of inquiry.

3.3 Discussion

In this section, we have called attention to two important methodological distinctions that have been neglected in evaluations of the expertise defence. In particular, we suggest that the dominant monolithic interpretation of philosophical expertise does not adequately accommodate for distinctions between the various uses of thought-experiments in philosophical inquiry, nor the particular cognitive processes underlying philosophical intuitions.[16]

In line with the arguments in this section, we suggest that **Expert-Practices Defences** is a better way to think about the relation of support between philosophical expertise and the expertise defence. Recall, this is proposal that *some* philosophers display improved performance in a *particular set* of practices in eliciting, invoking, assessing, and synthesising intuitions, and that this suffices to shield *those* philosophers from *particular* deleterious effects on philosophical theorising. This proposal better accommodates the fact that philosophical expertise with respect to **Philosophical Practices** can stand in a variable relation of support to the expertise defence. Recall, the reason for this is that the many distinctions amongst the **Philosophical Practices** engender distinct types of philosophical expertise, some of which may shield philosophers from biases, whilst others do not.

---

[16] We should note that we do not argue that philosophical expertise should be carved only at the joints of these distinctions; rather, we merely illustrate some ways that current accounts fail to accommodate the possibility that different types of philosophical expertise.

The central upshot of **Expert-Practices Defences** is that sweeping conclusions of the Critics' and Defenders' arguments are unwarranted: they are illicitly motivated by the simplistic way that philosophical expertise has been related to the expertise defence. Acknowledging the variable relationship between philosophical expertise and the expertise defence renders invalid any inference taking us from a small sample of evidence for the expertise defence in one set of cases to a wholesale rejection or acceptance of it in all cases. Given these considerations, it is possible to remain open-minded about the expertise defence in advance of close study into the specific area and practice to which it is being applied.

However, an important objection to such agnosticism is in the offing: even if Critics cannot automatically draw large-scale negative metaphilosophical conclusions from their empirical studies, they might claim that these studies—along with the lack of countervailing evidence—justifies a strong presumption against the viability of any version (local or global) of the expertise defence. In order to give some succour to Defenders, we will now outline one example where robust empirical evidence *supports* a successful instance of the expertise defence. We think that this approach of engaging with Critics on empirical grounds is more edifying than arguing over where the burden of proof lies in metaphilosophical debates; ideally Defenders and Critics will settle their metaphilosophical differences by gathering and evaluating targeted bodies of empirical evidence, not by attempting to claim temporary victories through dialectical fiat.

4 A (very) narrow account of the expertise defence

4.1 Cognitive epistemology

One of the most promising recent developments in experimental philosophy is the 'cognitive epistemology' movement.[17] Proponents of this movement seek to develop psychological explanations of our intuitions that can help to assess their evidentiary value. Recent studies in cognitive epistemology develop this approach by first tracing intuitions back to the particular cognitive processes that give rise to them, and then identifying vitiating circumstances in which our intuitions can lead us astray (e.g., Nagel 2010, 2012b; Fischer 2014; Alexander et al. 2015; Gerken and Beebe 2016; Fischer and Engelhardt 2016, 2017b). Ultimately, the hope is to provide an 'epistemic profile' of the cognitive processes underlying philosophical intuitions that allows us to assess their evidential merit in specific types of philosophical inquiry.

A primary advantage of cognitive epistemology is that it facilitates a piecemeal approach to investigations of the nature and epistemology of intuitions. In line with our previous discussions, we suggest that this fine-grained approach to investigating the epistemology of intuitions is fitting to assess the viability of the expertise defence. In this sense, cognitive epistemology provides resources which can help to articulate the two claims constitutive of the expertise defence:

---

[17] This term is coined by Fischer (2014).

1. Philosophers have *philosophical expertise*—adequate training in philosophy leads to improved performance in some task(s).
2. Philosophical expertise *defends* philosophers from distorting factors that Critics claim undermines the practice of appealing to intuitions in philosophy.

If cognitive epistemology pinpoints biases and their deleterious effects in philosophical theorising, then presumably we can also use these findings to evaluate whether practices philosophers employ makes them less susceptible to these biases. That is, we can ask: do philosophers happen to do the sorts of things that shields them from the problematic effects of biases? We now argue that, for the application of one philosophical practice to one specific philosophical puzzle, that the answer is 'yes'.

4.2 Philosophical expertise and ordinary language analysis

To demonstrate the relevance of recent empirical work in vindicating a narrow version of the expertise defence, we will look at the practices subsumed under the approach of Ordinary Language Analysis (OLA)—for helpful reviews, see Hansen (2014) and Baz (2016). Roughly put, OLA is the methodology which suggests that it can be philosophically fruitful to pay careful attention to how we use particular words in *everyday contexts*: the motivating idea is that this can help us to identify and resolve philosophical difficulties that arise from the mismatch in how these words are used in ordinary contexts, and in how philosophers use them in philosophical inquiry.[18] In what follows, we argue that there is empirical evidence showing that the practices of OLA, when applied within a particular line of inquiry, can give rise to a kind of philosophical expertise that shields philosophers from a specific type of bias.[19]

Before proceeding, we should note that we assume that when a philosopher uses OLA in attempting to solve a philosophical puzzle, the use of such techniques will be informed by previous philosophical training. Of course, philosophical training is not a *necessary* condition for deploying OLA—after all, this methodology presumably relies on general capacities we employ to make sense of discourse in everyday life. However, given that appeal to OLA in philosophical contexts requires careful attention to philosophical texts, together with the ability to effectively reconstruct argument and situate them among other philosophical theories, it is natural to expect that this skill will typically result from extensive training in philosophy.

One of the high-water marks of ordinary language analysis is Austin's *Sense and Sensibilia*. One of the primary targets here is Ayer's (1963) version of the so-called 'Argument from Illusion'—a historically influential argument that has shaped many debates in the philosophy of perception. This argument proceeds from descriptions

---

[18] For further discussion, see Hansen (2014).

[19] Although OLA nowadays has something of a bad name, it is important to note that one can adopt the practices of OLA without subscribing to the more hefty methodological commitments that might have been shared by its initial proponents. And indeed, the frequency with which considerations about linguistic data are brought to bear on assessments of philosophical theories shows that OLA is alive and well today. For a discussion, see Hansen (2014); for a contemporary (qualified) defence of OLA, see Hansen and Chemla (2015) and Hansen (2017).

of cases of non-veridical perception—i.e., cases where objects look or appear to have properties they do not actually possess. For instance:

(A) A stick that normally looks straight may appear bent when it is seen in water.

Ayer (1963, p. 4) contends it is obviously implausible that a straight stick *becomes* bent once put in water, or that the stick can be simultaneously both crooked and straight. Therefore, if a stick can both appear to be straight in normal circumstances *and* look bent when put in water, then this suggests that in *at least* one of these cases we are not *directly aware* of the stick. Nevertheless, Ayer contends that even if the viewer is not directly aware of the stick, they must be aware of *something*. This 'something' is called a sense-datum, which is credited with the properties the stick appears to have.[20]

One of Austin's central contentions is that proponents of the Argument from Illusion get us into a muddle because they neglect the subtle implications of the verbs 'looks', 'seems' and 'appears' when describing cases of non-veridical perception (Austin 1962, pp. 33–43). As Austin aptly points out, proponents of the Argument from Illusion take themselves to use these *appearance-verbs* in a very distinctive way: specifically, they employ them to indicate *merely* how things look/appear/seem to a viewer 'there and then', without implying that viewers would be inclined to judge that things *really are* that way. Thus, they assume (quite plausibly) that most people who read descriptions like 'a straight stick half-submerged in water appears bent to S' will *not* infer that S would be inclined to *judge* that the stick really *is* bent. After all, most of us would know better than to judge that the stick *is* bent just because it appears to so be 'there and then'—we naturally assume that any other viewer would also know better.

However, as Austin skilfully teases out by way of a number of *linguistic intuitions*, clauses of the form "X seems F to S" and "X appears F to S" imply just that. That is, Austin proposes that these constructions *generally* convey that S would, lacking any countervailing information, judge that "X *is* F" (Austin 1962, pp. 36–38). To illustrate, imagine I claim that 'the new bike lanes *appear* to be quite narrow', or 'the church *seems* to be gothic'. In these cases, we *ordinarily* take the speaker to endorse the content that comes after the appearance-verb. Of course, this *doxastic* implication can be *cancelled* by explicit contradiction: E.g., if we say 'the church seems gothic, but was actually constructed only 10 years ago', we take the relevant appearance-verb to indicate *only* how things seem to the viewer 'there and then', without implying that she judges things to be that way. However, absent such cancellations, we typically interpret these verbs in their doxastic sense, and infer that subjects would judge accordingly.

Why is any of this relevant for the Argument from Illusion? The reason is that, when describing cases of non-veridical perception, proponents of this argument intend to use appearance-verbs in a purely *phenomenal* sense: that is, they take these verbs to indicate only how things seem to a subject from that perspective, without implying anything about their inclination to judge that things really are as they appear, look, or seem to them 'there and then'. However, as Austin points out, uncancelled uses of appearance-verbs ordinarily have just this implication; moreover, this suggests that

[20] Both Russell (1912) and Broad (1923) present the Argument from Illusion in ways very similar to Ayer. However, it is noteworthy that contemporary reconstructions of this argument differ considerably from these formulations (see, e.g., Smith 2002).

uncancelled use of these verbs when describing non-veridical perception cases for the Argument from Illusion might not suffice to indicate that these verbs are meant in their purely phenomenal sense (Austin 1962, pp. 36–41). Therefore, as Austin is at pains to do throughout, it is very important to make explicit exactly what is intended by the use of certain appearance-verbs, and what is already known by the viewer. For instance, one needs to be reminded of the situation one is in when it is claimed that a stick half-submerged in water *appears* bent, and how this bears on the reading of the relevant appearance-verb; after all, it is not implausible to think that in some cases one could in fact suspect that the stick *really is* bent—say, when one did not previously know that the stick was actually straight. Thus, Austin's use of OLA calls attention to the fact that, without due caution in our philosophical discussions, certain uses of appearance-verbs can exert potentially confusing influences on our judgements.

So much for Austin's application of Ordinary Language Analysis to the Argument from Illusion; why think that this has any relevance to a successful expertise defence? Here's why: in a number of recent papers, Fischer et al. (2015, 2016, 2017b) provide empirical evidence that supports Austin's suspicions that misuse of appearance-verbs in descriptions of cases of non-veridical perception gives rise to significant philosophical difficulties. In what follows, we first present details of this empirical research, and then build on their findings to argue that Austin's use of OLA in analysing the Argument from Illusion is a clear display of significant philosophical expertise that does lend support to the expertise defence.

The empirical research by Fischer et al. (2015, 2016, 2017b) is a prime example of a project in cognitive epistemology: it aims to provide a psychological explanation of the intuitions driving the Argument from Illusion, and then builds on this explanation to assess their evidentiary value. The psychological explanation Fischer et al. provide builds on a body of work in cognitive psychology on the process called *stereotype enrichment*. This is a cognitive process that is key to our competence as speaker/hearers of a language, and which plays a central role in both language production and comprehension (Hare et al. 2009; Ferretti et al. 2001; McRae et al. 1997; Atlas and Levinson 1981). To illustrate the operations of this process, consider the following claim:

   (Tomato) "She looked at the vibrant colour of the tomato"

Most people who read the claim above infer that I was looking at the colour red—rather than green, orange, or yellow. Why do people immediately leap to this conclusion, since the tomato could have been of any one of these other colours? One well-accepted explanation is that, for most of us, tomatoes have a strong *stereotypical* association with the colour red; this strong association is what facilitates the inference to the conclusion that the viewer faces a particular colour (Atlas and Levinson 1981). In this sense, we "enrich" the utterance with our stored information about the colours that are stereotypically associated with tomatoes.

Stereotype enrichment is a generally reliable process for arriving at accurate conclusions about the world.[21] For instance, most of the tomatoes we encounter in our

---

[21] The explanation for this is that stereotypical associations are encoded in our semantic memory, which evolves from frequent activation by repeated encounters with certain stimuli in the world (Neely 1991).

daily lives are in fact red. However, there are obviously cases where such stereotype enrichments can lead us astray. The following riddle is helpful to illustrate (Giora 2003, p. 13):

> **(S)** A boy and his father get into a severe car accident. The father dies immediately, and the boy is taken to hospital. In the emergency room, the surgeon looks at the boy and says 'I can't operate on him. He's my son.' How is this possible?

Many people, when first exposed to this question, are stumped by it. A plausible explanation for why people tend to be puzzled is the fact that the term 'surgeon' has a strong gender stereotype. This strong stereotypical association immediately prompts the leap to the conclusion that the surgeon must be the father—which, in turn, gives rise to the confusion ('But the father has died!'). This example illustrates how, even under situations where we would expect the contextual information to trigger the non-stereotypical meaning (viz., the use of 'surgeon' to pick out the boy's mother), we may still draw inferences that are 'enriched' by stereotypical associations.

In recent work, Fischer et al. (2015, 2016) argue that a similar kind of bias drives the intuitions underlying the Argument from Illusion. One of the central lines of empirical support for this claim is the evidence that *all* appearance-verbs (seems, appears, and looks) have very strong stereotypical association with certain doxastic properties. In more detail, Fischer et al. provide substantial empirical evidence—by way of a variety of experimental methods—for the following proposal[22]:

> **(H)** The assertion 'x looks/appears/seems F to S' stereotypically implies that *S is inclined to judge that x is F*.

The central upshot of **(H)** is that people will *ordinarily* interpret *all* appearance-verbs in their *doxastic sense*—i.e., to imply specific doxastic properties about the viewer. This in itself may sound rather unsurprising; after all, the work of Austin already suggests as much: we tend to interpret clauses of the form "x seems F to S" and "x appears F to S" in this dominant doxastic sense, unless this meaning is explicitly ruled out.[23] However, the findings from Fischer et al. (2015, 2016, 2017b) build on

---

Footnote21 continued

Given that the strength of a link in semantic memory is directly correlated to the observed co-occurrence of frequencies in the subject's environment, semantic memory will then tend to encode accurate stereotypes (e.g., most tomatoes are red). Of course, stereotypes are not always correct, and may be encoded by misleading information in the world—e.g., biased media and uncooperative testimony.

[22] These findings were obtained from two distinct experimental methods. The first is a distributional-semantic analysis of these verbs within a vast linguistic corpus, which suggests there to be strong links between these verbs and the relevant doxastic properties. The second was the use of a forced-choice task where participants were asked to rate the plausibility of items where the complement of the appearance-verb was inconsistent with doxastic inferences. Results showed 'look' and 'appear' to be strongly associated with doxastic features, and 'seem' to be *even more* strongly associated with such features. Notably, this last method can also be combined with experimental techniques of pupillometry, which tracks increase in levels of cognitive effort by measuring the dilation of participants' pupils. This method has been fruitful in investigation of perception-verbs ('see' and 'aware'), where results obtained from previous forced-choice plausibility-ranking tasks with such verbs (Fischer and Engelhardt 2017a) were replicated in studies employing techniques of pupillometry (Fischer and Engelhardt 2017b).

[23] Interestingly, Fischer et al's work show that *all* appearance-verbs (including clauses of the form "x looks F to S", whose doxastic uses Austin (1962, p. 36) dismisses) have strong stereotypical associations with certain doxastic properties.

Austin's suggestion in an important way: their findings show that these stereotypical associations are so strong that people will tend to interpret these verbs in their doxastic sense *even in situations which invite the non-doxastic (phenomenal) interpretation of these verbs*.

This finding is surprising; because, as mentioned above, proponents of the Argument from Illusion take it that people will interpret these verbs in their purely *phenomenal* sense when they read descriptions of non-veridical cases of perception. That is, they assume (quite plausibly) that in reading descriptions of cases such as 'a large boat viewed from afar *appears* small', people will resist inferring that the viewer in these cases is inclined to judge that the object viewed *is actually* small. Again, this assumption is well-motivated: after all, most of us know very well that a large boat does not *actually* possess the property of smallness because it is viewed from afar—and we take it that other people would also not be tempted to so judge. However, the findings from Fischer et al. show just that: people will indeed tend to interpret these verbs in their doxastic sense when reading these kinds of cases—thus implying that in this case the viewer would be inclined to judge that the object does have the properties it looks/appears/seems to have. This in itself calls for an explanation: given the plausible assumption that people would interpret these verbs in their phenomenal sense when reading the kinds of cases that motivate the Argument from Illusion, why is it that they still take them in their doxastic sense?

Relying on a large body of work from research in cognitive psychology, Fischer and colleagues propose there is a rather natural explanation for this tendency. A clue to the answer is found in explanation of the small riddle of the surgeon (**S**) presented above. Recall, the explanation for why it can be so difficult to solve this riddle is that the term 'surgeon' has a very strong gender stereotype with 'male'. Indeed, as the riddle helpfully illustrates, mere mention of the term will immediately prompt the stereotype-driven inference that the surgeon is the father—even when the context indicates that this conclusion is not possible. A similar kind of explanation can account for people's tendency to interpret appearance-verbs in their doxastic sense when reading descriptions of cases of non-veridical perception. More precisely, the strong stereotypical associations between appearance-verbs and the relevant doxastic properties immediately drives the stereotype-driven inference that viewers are inclined to judge that things are as they look/appear/seem to them. Moreover, like in the riddle of the surgeon, this inference is activated even when contextual information indicates otherwise—e.g., even in the light of the explicit information that the stick that looks bent *is* straight.

Fischer et al. (2015, 2016) argue that this provides a psychological explanation for why proponents (and readers) of the Argument from Illusion effortlessly leap from descriptions of cases of non-veridical perception to the negative conclusion that in such cases we are not aware of objects themselves. Very roughly put, the reason why this conclusion seems to follow so naturally is that when people read, e.g., 'a straight stick half-submerged in water looks bent', they then immediately infer that the viewer is inclined to judge that the object viewed *is* bent. However, it is explicitly clear from the description of the case that the stick is straight, rather than bent. This conflict between the stereotype-driven inference (that the object viewed is judged to be bent) and the explicit description of the case (where the stick is described as being straight)

then leads to the conclusion that the viewer does not know what object she is viewing, and is not aware of that object—but of something else.

The psychological explanation above reveals how the Argument from Illusion gains its intuitive force by trading on a confusion between the *phenomenal* and *doxastic* senses of the relevant appearance-verbs. This then suggests that the Argument from Illusion is underwritten by an incisive kind of cognitive bias and that the intuitions driving this argument are epistemically worthless. The particular bias in question is traceable to an identifiable set of vitiating circumstances: specifically, this bias arises because proponents of the Argument from Illusion make use appearance-verbs in their less salient *phenomenal* sense when describing cases of non-veridical perception, without making it explicit that this use deviates from the (dominant) doxastic sense of these verbs. The unmarked use of this less salient technical sense leads competent speakers to mistakenly draw inferences in line with the ordinary (doxastic) sense of this verb, which conflict with the inferences that are natural to draw with respect to its phenomenal sense. Moreover, it is precisely the confusion between the doxastic and phenomenal senses of the relevant verbs that prompts the conflicting intuitions driving the Argument from Illusion. Given that such linguistic confusions should have no bearing whatsoever on philosophical investigations of the nature of perception, then it is plausible to conclude that these intutions are epistemically defective.

Let us now return to the discussion of the expertise defence. We propose that the empirical work of Fischer and colleagues reveals that Austin's use of OLA amounts to a form of philosophical expertise that does suffice to shield philosophers from certain biases, and therefore supports an instance of the expertise defence. As discussed above, one of Austin's central contentions was that proponent of the Argument from Illusion get us into a muddle when they use appearance-verbs without due attention to their subtle pragmatic implications. By paying attention to these implications, we can avoid the misleading influences they can exert on our judgements. To support his claim, Austin employed a **Critical** use of the vignettes motivating the Argument from Illusions. As the empirical work of Fischer et al. demonstrate, Austin's **Critical** examination of these cases is correct: the confusion Austin identifies is indeed the source of what we take to be an uncontroversial instance of a bias.[24] The bias in question stems from the operation of a generally reliable cognitive process known as stereotype enrichment which, due to the vitiating circumstances arising from misuse to the subtle implication of appearance-verbs, gives rise to intuitions that have no evidentiary value. In this way, the empirical evidence reviewed above shows that a specific form of philosophical training (in OLA) can lead to improved performance in particular practices surrounding the use of intuitions (using thought-vignettes in a **Critical** way to uncover *and suppress* a bias stemming from a specific cognitive process) for the purposes of improved philosophical inquiry into the nature of perception.

---

[24] It is worth noting that this discussion suggests a further locus for philosophical expertise: framing verbal descriptions of cases in such a way that can avoid engaging biases arising from linguistic confusions in the first place. This form of expertise can be enabled by using empirical research in psycholinguistics, which reveals the precise ways in that these linguistic confusions can lead us astray. E.g. see Fischer and Engelhardt (2017b).

Does this mean that ordinary language analysis suffices to defeat biases across the board? Almost certainly not. As we have attempted to emphasise throughout this paper, more care is required in attributions of philosophical expertise, and relating it to the expertise defence. We have not provided evidence that ordinary language philosophy can help to avoid problematic effects of other biases, or that its practices are truth-conducive in other instances. But, this is not an unwelcome upshot of our view; as we have suggested, this piecemeal approach is required to do justice to the complexity of philosophical inquiry, and to provide a realistic evaluation of the expertise defence. It remains to be seen whether the expertise defence can be upheld for other practices, in other lines of philosophical inquiry. In advance of the careful empirical work that is required to identify the processes driving cognitive biases, and considering how these intersect with our current philosophical practices, we recommend an attitude of studied agnosticism.

5 Conclusion

We have been exploring the viability of the expertise defence as a response to sceptical claims regarding the use of intuitions in philosophy. By critiquing extant positions, we found that both parties in this debate operate with an inadequate conception of philosophical expertise. In responding to these worries, we have offered a novel account of philosophical expertise that is sensitive to the variety of different practices and methods that comprise philosophical inquiry. This account better reflects the complex and variable relationship between philosophical expertise and experimentally motivated scepticism about intuitions. The upshot of our account is that we cannot vindicate or reject the expertise defence all at once. Rather, we must adopt a nuanced and localised approach that attends to the details of precisely what philosophical methods are being used, and whether these methods suffice to shield philosophers from specific distorting effects that skew our intuitions. And finally, we illustrated how one might take up our novel approach by appealing to recent work in cognitive epistemology, which we propose supports a local instance of the expertise defence. Our arguments show that much work remains to be done on the extent to which appealing to philosophical expertise can mitigate the negative implications of experimental philosophy; nonetheless, we have offered a framework in which both parties to the debate can make concrete progress.

References

Alexander, J., Betz, D., Gonnerman, C., & Waterman, J. P. (2017). Framing how we think about disagreement. *Philosophical Studies*. https://doi.org/10.1007/s11098-017-0971-9.

Alexander, J., Gonnerman, C., & Waterman, J. (2015). Salience and epistemic egocentrism: An empirical study. In J. Beebe (Ed.), *Advances in experimental epistemology*. London: Continuum.

Antony, L. M. (2016). Bias: Friend or foe? Reflections on saulish skepticism. In M. Brownstein & J. Saul (Eds.), *Implicit bias and philosophy. Metaphysics and epistemology* (Vol. 1). Oxford: Oxford University Press.

Atlas, J. D., & Levinson, S. C. (1981). It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In *Radical pragmatics* (pp. 1–62). Academic Press.

Austin, J. L. (1962). *Sense and sensibilia*. Oxford: Oxford University Press.

Ayer, A. J. (1963). *The foundations of empirical knowledge*. New York: Macmillan.

Baz, A. (2016). Ordinary language philosophy. In H. Cappelen, T. S. Gendler, & J. Hawthorne (Eds.), *The Oxford handbook of philosophical methodology*. Oxford: Oxford University Press.

Beebe, J. R., & Buckwalter, W. (2010). The epistemic side-effect effect. *Mind & Language*, *25*(4), 474–498.

Boyd, K., & Nagel, J. (2014). The reliability of epistemic intuitions. In E. Machery & E. O'Neill (Eds.), *Current controversies in experimental philosophy*. Routledge.

Broad, C. D. (1923). *Scientific thought*. London: Routledge and Kegan Paul.

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

Cohen, S. (1999). Contextualism, skepticism, and the structure of reasons. *Noûs*, *33*(s13), 57–89.

DeRose, K. (1991). *Consciousness explained*. Baltimore: Penguin Books.

Dennett, D. C. (1991). *Consciousness explained*. Baltimore: Penguin Books.

DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, *52*(4), 913–929.

Deutsch, M. (2015). *The myth of the intuitive*. Cambridge: The MIT Press.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*(4), 516–547.

Fischer, E. (2014). Philosophical intuitions, heuristics, and metaphors. *Synthese*, *191*(3), 569–606.

Fischer, E., & Collins, J. (Eds.). (2015). *Experimental philosophy, rationalism, and naturalism. Rethinking philosophical method*. London: Routledge.

Fischer, E., & Engelhardt, P. E. (2016). Intuitions' linguistic sources: Stereotypes, intuitions and illusions. *Mind and Language*, *31*(1), 67–103.

Fischer, E., & Engelhardt, P. E. (2017a). Diagnostic experimental philosophy. Teorema: International. *Journal of Philosophy*, *36*(3), 117–137.

Fischer, E., & Engelhardt, P. E. (2017b). Stereotypical inferences: Philosophical relevance and psycholinguistic toolkit. *Ratio*, *30*(4), 411–442.

Fischer, E., Engelhardt, P. E., & Herbelot, A. (2015). Intuitions and illusions: From explanation and experiment to assessment. In E. Fischer & J. Collins (Eds.), *Experimental philosophy, rationalism, and naturalism. Rethinking philosophical method* (pp. 259–292). London: Routledge.

Gendler, T. (2010). *Intuition, imagination, and philosophical methodology*. Oxford: Oxford University Press.

Gerken, M., & Beebe, J. R. (2016). Knowledge in and out of contrast. *Noûs*, *50*(1), 133–164.

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford: Oxford University Press.

Greene, J. D. (2015). Beyond point-and-shoot morality: Why cognitive (neuro) science matters for ethics. *The Law & Ethics of Human Rights*, *9*(2), 141–172.

Hansen, N. (2014). Contemporary ordinary language philosophy. Philosophy. *Compass*, *9*(8), 556–569.

Hansen, N. (2017). Must we measure what we mean? Inquiry: An Interdisciplinary. *Journal of Philosophy*, *8*, 1–31.

Hansen, N., & Chemla, E. (2015). Linguistic experiments and ordinary language philosophy. *Ratio*, *28*(4), 422–445.

Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, *111*(2), 151–167.

Jackson, F. (1982). *Epiphenomenal qualia. Philosophical Quarterly*, *32*(127), 127–136.

Jackson, F. (2003). Mind and illusion. *Royal Institute of Philosophy Supplements*, *53*, 251–271.

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

Knobe, J. (2016). Experimental philosophy is cognitive science. In J. Sytsma & W. Buckwalter (Eds.), *A companion to experimental philosophy*. New York: Blackwell.

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

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.

Lowe, E. J. (1983). On the identity of artifacts. *Journal of Philosophy*, *80*(4), 220–232.

Ludwig, K. (2007). The epistemology of thought experiments : First person versus third person approaches. In P. A. French & H. K. Wettstein (Eds.), *Midwest studies in philosophy* (Vol. 31, pp. 128–159). New York: Blackwell Pub. Inc.

Machery, E. (2015). The illusion of expertise. In J. Collins & E. Fischer (Eds.), *Experimental philosophy, rationalism, and naturalism: Rethinking philosophical method*. London: Routledge.

McRae, K., Ferretti, L., & Amyote, T. R. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*(2–3), 137–176.

Mizrahi, M. (2015). Three arguments against the expertise defense. *Metaphilosophy*, *46*(1), 52–64.

Nado, J. (2014a). Philosophical expertise and scientific expertise. *Philosophical Psychology*, *28*(7), 1026–1044.

Nado, J. (2014b). Why intuition? *Philosophy and Phenomenological Research*, *89*(1), 15–41.

Nado, J. (2016). The intuition deniers. *Philosophical Studies*, *173*(3), 781–800.

Nagel, J. (2010). Knowledge ascriptions and the psychological consequences of thinking about error. *Philosophical Quarterly*, *60*(239), 286–306.

Nagel, J. (2011). The psychological basis of the harman-vogel paradox. *Philosopher's Imprint*, *11*(5), 1–28.

Nagel, J. (2012a). Intuitions and experiments: A defense of the case method in epistemology. *Philosophy and Phenomenological Research*, *85*(3), 495–527.

Nagel, J. (2012b). Mindreading in Gettier cases and skeptical pressure cases. In J. Brown & M. Gerken (Eds.), *Knowledge ascriptions*. Oxford: Oxford University Press.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition*, *11*, 264–336.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In N. Rescher (Ed.), *Essays in honor of Carl G* (pp. 114–146). Hempel: Reidel.

Parfit, D. (1984). *Reasons and persons*. Oxford: OUP.

Rini, R. A. (2015). How not to test for philosophical expertise. *Synthese*, *192*(2), 431–452.

Russell, B. (1912). *The problems of philosophy*. New York: Barnes & Noble.

Scaltsas, T. (1980). The ship of theseus. *Analysis*, *40*(3), 152–157.

Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind and Language*, *27*(2), 135–153.

Schwitzgebel, E., & Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection. *Cognition*, *141*, 127–137.

Smart, B. (1972). How to reidentify the ship of theseus. *Analysis*, *32*(5), 145–148.

Smith, A. D. (2002). *The problem of perception*. Cambridge: Harvard University Press.

Sorensen, R. A. (1998). *Thought experiments*. Oxford: Oxford University Press on Demand.

Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, *94*, 1395–1415.

Tobia, K., Buckwalter, W., & Stich, S. (2013). Moral intuitions: Are philosophers experts? *Philosophical Psychology*, *26*(5), 629–638.

Weinberg, J. M., Alexander, J., Gonnerman, C., & Reuter, S. (2012). Restrictionism and reflection. *The Monist*, *95*(2), 200–222.

Weinberg, J. M., Gonnerman, C., Buckner, C., & Alexander, J. (2010). Are philosophers expert intuiters? *Philosophical Psychology*, *23*(3), 331–355.

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

Williamson, T. (2011). Philosophical expertise and the burden of proof. *Metaphilosophy*, *42*(3), 215–229.