

In Defence of Armchair Expertise

by

THEODORE BACH 

Bowling Green State University

Abstract: In domains like stock brokerage, clinical psychiatry, and long-term political forecasting, experts generally fail to outperform novices. Empirical researchers agree on why this is: experts must receive direct or environmental learning feedback during training to develop reliable expertise, and these domains are deficient in this type of feedback. A growing number of philosophers resource this consensus view to argue that, given the absence of direct or environmental philosophical feedback, we should not give the philosophical intuitions or theories of expert philosophers greater credence than those of novice philosophers. This article has three objectives. The first is to explore several overlooked issues concerning the strategy of generalizing from empirical studies of non-philosophical expertise to the epistemic status of philosophical expertise. The second is to explain why empirical research into a causal relationship between direct learning feedback and enhanced expert performance does not provide good grounds for abandoning a default optimism about the epistemic superiority of expert philosophical theories. The third is to sketch a positive characterization of learning feedback that addresses developmental concerns made salient by the empirical literature on expert performance for specifically theory-driven or “armchair” domains like philosophy.

Keywords: expertise, metaphilosophy, development, success, theory, theoretical

1. Introduction

A RECENT DEBATE OVER THE epistemic value of philosophical intuition has revealed ways in which empirical research into expert performance bears on questions about philosophical practice. That debate started with experimental studies that revealed how the philosophical intuitions of undergraduate students were influenced by epistemically irrelevant factors (e.g., order of problem presentation; for example, Weinberg et al., 2001; Swain et al., 2008). Several philosophers replied that these studies did not support a critical view of *expert* philosophical intuitions, which they claimed are the intuitions that matter most (e.g., Hales, 2006; Kauppinen, 2007; Ludwig, 2007; Sosa, 2007). But then came a new set of experimental studies that revealed how professional philosophical intuitions were just as influenced by the epistemically irrelevant factors (e.g., Schulz et al., 2011; Schwitzgebel and Cushman, 2012; Tobia et al., 2013). Some defenders of traditional philosophical methods then responded that these interference effects were

ultimately compatible with the claim that expert philosophers are *generally* better at intuiting and theorizing than novices.¹

Enter the empirical literature on expert performance. The critics appealed to widely accepted empirical studies of expert performance that indicated the importance of *direct* or *environmental* learning feedback for the development of high-quality expert performance (e.g., Weinberg et al., 2010; Clarke, 2013; Ryberg, 2013). According to this empirical literature, there are several domains in which experts struggle to outperform novices (clinical psychiatry, political forecasting, financial advising, etc.), and these domains have in common that they do not sufficiently make available to experts and experts-in-training direct or environmental learning feedback. According to the same empirical literature, the domains in which experts do consistently outperform novices (livestock judging, chess, meteorology, etc.) make available the direct or environmental type of learning feedback.

Having established this developmental condition for enhanced expert performance, the critics then argued that there is no empirical evidence that direct or environmental learning feedback is available for improving the intuitive judgments of training philosophers (Weinberg et al., 2010; Ryberg, 2013). Some critics sharpened the point, arguing that there is positive empirical evidence that the epistemically good type of feedback is unavailable in the domain of philosophy (e.g., Clarke, 2013). Either way, the upshot was that, putting the issue of interference effects to the side, the empirical literature on expertise provided a direct challenge to the presumption that professional philosophers *qua* experts are generally better intuiters than novice philosophers.

But at this point, the debate was no longer just about philosophical intuition. The debate had now also implicated the epistemic status of expert philosophical *theories* and their development. This is for at least two reasons. First, according to research on expert performance, the empirical basis for the importance of direct feedback concerns what is developmentally critical for *experience-based learning generally*.² Thus if this empirical literature has developmental

1 In Devitt's terms: "The [Expertise] Defense requires only that the philosophers' intuitions be better, in general ... even if just as influenced by non-truth-tracking factors as the folk's" (Devitt, 2012, p. 22). See Williamson (2011, pp. 218–219) for a similar claim.

2 As Dawes puts it, "such [immediate and correct] feedback is important if people are to learn anything at all from experience – whether it is a concept, a general idea of how to deal effectively with people, or a motor skill" (Dawes, 1994, pp. 119–120). Similarly, Hogarth (2001, pp. 87, 98–99) makes clear that the availability of direct feedback informs learning through experience generally. I will note here that one reason that intuition figures centrally in feedback-based discussions of expert performance in the empirical literature is that many institutional experts who lack knowledge for feedback-based reasons nonetheless put forward beliefs or policies – beliefs or policies derived from psychological processes that might be described as intuition-based – as knowledge. This dynamic is a recurring theme in Dawes's critique of expert psychiatry, for example (see Dawes, 1994).

implications for philosophical expertise, then these implications cannot be restricted to intuitive judgement. The implications should apply to whichever epistemic task, for example theoretical modelling, the enhanced performance of which we think is dependent on experience and training.

Second, many philosophers and psychologists model intuitive judgements as the cognitive products of internalized theories (see, e.g., Kornblith, 1998; Gendler, 2007; Kahneman and Klein, 2009; see also Weinberg et al., 2010, p. 344). On this view, one can defend the (enhanced) epistemic status of expert philosophical intuition indirectly – by defending the (enhanced) epistemic status of the (internalized) expert philosophical theories from which the intuitive judgements originate.³ A central reason that critics of philosophical intuition resource the empirical literature on non-philosophical expert performance is to address this theory-based strategy for defending expert philosophical intuition. The critics report how, according to the empirical literature on expertise, the development of “epistemically virtuous” (or truth-tracking) concepts, rules or theories depends on adequate training exposure to direct feedback (see, e.g., Weinberg et al., 2010, pp. 340–341).⁴ They then argue that we do not currently have good reasons for claiming that philosophers are sufficiently exposed to this type of feedback. If these claims are correct, then it would follow that the strategy of appealing to expert philosophers’ philosophical conceptual schemas or their psychologically internalized philosophical theories in order to ground the truth-tracking credentials of expert philosophical intuitions is unsuccessful.

3 Weinberg et al. (2010, p. 344) express this view as follows: “perhaps what makes philosophers expert intuiters is that they, unlike most undergraduates, have mastered philosophy’s well-established theories or principles, materials which they bring to bear when evaluating their thought experiments”. Weinberg et al. also allow that such theory-issued judgements can count as intuition: “We want to remain maximally neutral on questions about the necessary conditions for intuitionhood, and so it could be that some such judgments are among what philosophers have in mind when they talk about intuitions or that some satisfy the role traditionally assigned to intuitions” (Weinberg et al., 2010, p. 344). They further describe how a “slight tweak” to this view, whereby the theory is internalized in the philosopher’s cognitive or “adaptive unconscious”, would allow such theory-issued judgements to have the unconscious or unreflective ingredient that many philosophers deem necessary for intuition (Weinberg et al., 2010, p. 344).

4 Clarke (2013), who endorses and extends Weinberg et al.’s feedback-based challenge to expert philosophers, helpfully summarizes their use of the empirical literature as follows: “Weinberg et al. (2010) have delved in depth into the empirical literature on expertise and on the basis of their investigations they argue against the view that professional philosophers have reliable expert intuitions. They consider three hypotheses that might underwrite the claim that professional philosophers have superior philosophical intuitions to the folk. These are: (i) that philosophers have superior conceptual schemata than the folk, (ii) that philosophers employ more sophisticated theories than the folk, and (iii) that philosophers possess superior cognitive skills than the folk. They examine the empirical literature on expertise that is relevant to all three hypotheses in detail and point out significant difficulties that this body of work raises for all of them (Weinberg et al. 2010, pp. 336–49)” (Clarke, 2013, p. 188).

In this article, I will not be concerned with this last step of the critics' argument. To be clear: I will neither discuss nor defend the role of philosophical intuition here (even if it may be possible to deploy the arguments that follow, along with several additional assumptions not discussed here, toward such a defence).⁵ Instead, I will focus on the prior steps of the critical challenge sketched above – the steps that concern expert theory development, mechanisms for learning feedback, and the strategy of generalizing from empirical research on non-philosophical expert performance.⁶ The two claims I focus on are the following. The first is the claim that the development of epistemically virtuous or truth-tracking expert philosophical theories (concepts, rules, etc.) may require training in a specific kind of direct or environmental feedback environment. The second is the claim that there are not compelling reasons to think that philosophers train or work in the type of feedback environment that could causally explain the development of epistemically virtuous philosophical theories (concepts, rules, etc.). These claims and their sceptical implication are of considerable importance for the epistemology of philosophy (and armchair domains generally), and they have this importance regardless of their implications for the alleged practice by philosophers of appealing to intuitions as sources of evidence.⁷

The empirical researcher Robin Hogarth has described feedback structures that promote inaccurate learning as “wicked” learning environments and feedback structures that promote accurate learning as “kind” learning environments (e.g., Hogarth, 2001, pp. 98–99; see also Kahneman and Klein, 2009, p. 520). One way to think about my objective here is as an investigation into whether we should have confidence that professional philosophers who have the goal of elucidating philosophical targets (material composition, moral properties, reference, etc.) through theoretical models train and work in a “kind” rather than “wicked” feedback learning structure. The difficulty, as Ryberg (2013) has pointed out, is that it is unclear what suitable quality parameter – a *prima facie* requirement of a kind learning environment – can provide adequate feedback for the successful

5 I flag this possibility in the conclusion (section 5). Also, none of the arguments to follow – arguments that address the development of expert philosophical theorizing – will rely on the claim that philosophers employ intuitive judgement.

6 Thus rather than attempting “the next move” in “the dialectic” over philosophers’ use of philosophical intuition, I am taking up a particular strand that developed within that dialectic, looking more closely at its roots in the empirical literature, and pulling it in different directions to explore its implications for the epistemology of philosophy.

7 Suppose that Cappelan (2012) and Deutsch (2015) are correct that philosophers do not use intuitions evidentially. If that were the case, it would still be of considerable metaphilosophical importance whether the empirical literature on non-philosophical expert performance supports a sceptical attitude toward the development of expert philosophical theories.

modelling of philosophical targets.⁸ And while it may be comfortable (and consistent with guild interests) for professional philosophers simply to assume that they work in a “kind” feedback learning structure relative to the goal described above, the empirical literature on expert performance warns us that such assumptions are not to be trusted.⁹ What is required is a non-question-begging and naturalistically plausible positive account of the type of feedback structures that can underwrite the development of improved expert philosophical theorizing. As I will explore, providing such an account is particularly challenging for theory-driven “armchair” domains in which important investigative questions are empirically underdetermined. Nonetheless, and as I argue in the second part of the article, such an account can be given.

I proceed as follows. In section 2, I explicate the type of sceptical challenge to philosophical theorizing (and armchair domain theorizing generally) that I argue is best supported by the empirical study of non-philosophical expert performance. That explication will reveal the importance of experts’ capacity to recognize what counts as a success. I also discuss in section 2 an important difference between expert failure in domains studied by empirical investigators, on the one hand, and the type of expert failure (if there is one) of experts in the philosophical domain, on the other.

In section 3, I distinguish between two types of expertise that I term “procedural expertise” and “content expertise”. In reference to this distinction, I explain why the sceptical argument explicated in section 2 does not provide adequate grounds for abandoning a default optimism about the enhanced accuracy of expert philosophers’ theoretical models.

Section 4 provides a positive characterization of learning feedback for specifically theory-driven or “armchair” domains like philosophy. I consider the claim that theoretical virtues – for example, simplicity and unificatory power – serve as indicators of theoretical success. I then explain how these indicators can function as sources of learning feedback and in what respect they are accessible from the armchair. Section 5 concludes.

8 Clarke puts his finger on the *prima facie* wickedness of the philosophical learning environment this way: “When a nurse has an intuition that a patient is in danger of an imminent heart attack she can ask other nurses if they share her intuition or not, and get feedback from them; but more importantly she can (and will) receive direct feedback from the environment. Either the patient does or does not go on to have an imminent heart attack. Learning whether or not her intuitions track reality enables her to train her future intuitions and improve these. In the case of philosophical intuitions, however, direct feedback from the environment is typically unavailable. We cannot directly discover what knowledge really is or what morality really demands” (Clarke, 2013, p. 193). I further discuss Clarke’s view and Ryberg’s notion of a quality parameter in section 2.

9 As Weinberg et al. (2010, p. 342) point out, “one would expect pretty much all fields to have some degree of feedback, and for the authorities in those fields to generally think it sufficient for developing highly improved performance. Yet the science here indicates that many of them, even in highly respected fields, are simply wrong about that. The overall pattern of findings in this part of the psychological literature provide us reason to think that philosophers are in this latter category of the mildly self-deceived”.

Before proceeding, I offer three clarifications regarding terminology. First: who qualifies as an “expert?” James Shanteau (1992, p. 255) explains that “experts are operationally defined as those who have been recognized within their profession as having the necessary skills and abilities to perform at the highest level”. Camerer and Johnson (1991, p. 196) – another source the critics reference often – define an expert as “a person who is experienced at making predictions in a domain and has some professional or social credentials”. To be an “expert”, then, is to have the appropriate institutional history and be regarded by peers as an expert. I follow this convention here, using the term “expert” and “expertise” in this descriptive, institutional sense. Applied to philosophy, I will assume that professional philosophers with graduate degrees from research institutions count as experts (but the arguments to follow do not require drawing that line precisely here). Occasionally, I will use the term “professional” instead of “expert”.

Second, what counts as a “theory” for our purposes? I will understand theories in terms of explanatory descriptions of philosophical explananda. While there are various ways to add specificity to this idea, I will adopt the suggestion from Paul (2012) that a philosophical theory is a description of a model or family of models intended to be isomorphic in some relevant sense to real world targets (but the arguments to follow do not require this construal). I will mostly use the expressions “philosophical theory” and “theoretical model” to refer to theories understood in this way, e.g., “Millikan’s theoretical model of proper functions”. There is also a psychological sense of the term “theory”. If the arguments to follow are successful, then what I claim about the development of successful theories (in the sense of theoretical models) should have implications for questions concerning the development of successful theories in the psychological sense.¹⁰ However, I limit my discussion here to the development of theories understood in the former, theoretical modelling, sense.¹¹

10 I thank an anonymous reviewer for pressing the importance of disambiguating different senses of the term “theory”. It is also worth mentioning in this context that Weinberg et al. (2010) treat a defence of intuition based on philosophers’ “specialized conceptual schemata” and a defence based on philosophical “theories” separately. However, these two types of informational structures are both causally and developmentally bound up together. For example, Weinberg et al. (2010, p. 338) describe higher-order concepts, which they offer as an example of specialized conceptual schemata, as concepts that encode which factors are most relevant to a question or problem. But presumably, a philosopher’s acceptance of particular philosophical theories will lead to corresponding higher-order concepts (conceptual schemata) that organize the features of a problem or question on the basis of that theory’s commitments and principles. Also relevant here are considerations raised in n. 3 concerning how descriptive theories can be psychologically internalized.

11 Though in section 4.3 I advance some remarks about a developmental relationship between the two forms or realizations of theoretical modelling.

Third, what counts as an “epistemically better” or “more successful” theory on my view? I will generally discuss epistemic betterness or epistemic success in terms of accuracy or correspondence. However, I wish to remain neutral on this point, and alternative understandings (e.g., assertability, warrant, coherence) would work just as well for the arguments that follow.

2. Untangling a Sceptical Argument against Expert Philosophical Theorizing

Are there good reasons to think that professional philosophers, in virtue of their expert training, develop more accurate theoretical models of philosophical explananda than novice philosophers? Call this claim “Expert Superiority”, or ES:

Expert Superiority (ES): Expert philosophers in general produce more accurate theoretical models of philosophical explananda than philosophical novices.

Few people believe that professional philosophical theorizing is without its fair share of blunders, false claims, wrong-headed research programmes and so forth. However, it would be a significant blow to professional philosophy (as well as a violation of prevalent assumptions about the value of philosophical expertise) if its core aim of generating accurate theoretical models of philosophical explananda (e.g., moral properties, consciousness, material composition, social categories, etc.) had a success rate that, similar to various expert projects in clinical psychiatry, long-term political forecasting and stock brokerage, was generally no better than the success rate of novices. This is what is at stake in the claim of *Expert Superiority*.

There are several challenges for those who wish to defend ES. First, one cannot empirically measure or directly observe the accuracy of philosophical theories in the way that researchers perform analogous measurements and observations in other domains. For example, empirical researchers conduct follow-up studies about the criminal behaviour of released prison inmates to measure the success of expert predictions about criminal recidivism. This and related empirical techniques have no obvious application in most philosophical contexts.¹² Second, it is not obvious from an examination of the history of philosophical theorizing that philosophers’ activity of theoretically modelling philosophical explananda has itself produced progress toward this end.¹³ A third challenge – one that emerges

¹² There are arguably exceptions to this point in specific philosophical sub-domains if select philosophical models in these domains have clear and testable implications. See, for example, Martí (2012, section 5) or Devitt (2015) on how we might test philosophical theories of reference against data on actual linguistic usage. But I will proceed on the assumption here that, from the perspective of current empirical investigative methods, most competing philosophical models are predictively equivalent.

¹³ But see Stoljar (2017) for a recent discussion and cautiously optimistic response to progress-based challenges to philosophical practice.

from the debate over philosophical intuition sketched in section 1 – is that there are compelling lines of empirical support, derived from the empirical investigation of non-philosophical expert performance, *against* the claim of *Expert Superiority* (ES). My concern here is to explore, and then outline a response to, this type of challenge to ES.

A large empirical literature on expert performance shows how experts outperform novices only in domains that make available a clear and direct or environmental type of learning feedback. This empirical finding holds across domains with starkly different subject matter, and thus there is strong empirical support for a *domain-general* principle that makes exposure to direct or environmental learning feedback a requirement for developing enhanced expert performance (see, e.g., Clarke, 2013, p. 191). As we saw in section 1, Weinberg et al. (2010) employ this empirical finding to challenge the assumption that expert philosophers have developed specialized conceptual schemata, rules and/or theories that are epistemically virtuous in the sense of being “fine-tuned” to truth-tracking factors.

Applying this domain-general principle to the philosophical aim of producing accurate theoretical models of philosophical explananda generates the following feedback requirement in reference to *Expert Superiority*:

Feedback Requirement (ES \supset FB): *Expert Superiority* (ES) is true only if expert philosophers are sufficiently exposed to direct or environmental learning feedback. (FB)

To assess the implications of *Feedback Requirement* (ES \supset FB), I need to say more about what direct or environmental learning feedback is. The essential feature of this type of feedback is that it allows the practitioner to know whether his or her performance was successful or unsuccessful. A chess player whose queen-sacrifice strategy always leads immediately to being checkmated receives direct feedback that the strategy is a poor one. A meteorologist who compares today’s weather against yesterday’s predictions can receive direct feedback about the methodology underlying yesterday’s predictions. And so on. It is feedback that communicates “clear verdicts” (Weinberg et al., 2010, p. 341), allows “apprehension of repeated or salient successes and failures” (Weinberg et al., 2010, p. 340), and where the expert-in-training can “correctly diagnose (perhaps tacitly) the causes of success or failure” (Weinberg et al., 2010, p. 340).

Can we reject *Expert Superiority* by invoking the *Feedback Requirement* and then denying that expert philosophers are sufficiently exposed to direct learning feedback? When discussing the development of epistemically virtuous philosophical intuitions, conceptual schemata and/or theories, Clarke (2013), Weinberg et al. (2010) and Ryberg (2013) consider various possible sources of feedback – other philosophers’ intuitions, thought-experiments, prior episodes of moral

reasoning – and they conclude that these are insufficiently reliable or clear or direct to count as sources of the right kind of feedback (see, e.g., Ryberg, 2013, p. 8; Weinberg et al., 2001, pp. 341–342; Clarke, 2013). We can view these authors as motivating the idea that the feedback requirement is not satisfied (i.e., \sim FB). At the very least, we should view them as highlighting an unnoticed explanatory burden that applies to defenders of philosophical expertise concerning how the feedback requirement is satisfied (e.g., Clarke, 2013, p. 191). Applying this explanatory burden in the context of the claim *Expert Superiority*, the specific challenge is one of providing a positive and empirically plausible account of reliable feedback mechanisms for the development of accurate philosophical theoretical models. This challenge, which I think shares important similarities to Mackie’s “argument from queerness” challenge to moral realists (see, e.g., the quotation from Clarke in n. 8), is one that I take up in section 4. But before getting there, we need to understand more clearly the constraints that the empirical literature on expert performance imposes on such a positive account. And to understand that, we first need to analyse what conditions apply to the type of informational feedback transactions that are epistemically important for the developing expert.

2.1 The role of success-identification in the sceptical argument against expert philosophical theorizing

The reason for the epistemic importance of direct feedback is straightforward: direct feedback facilitates error correction and improvement over time. Less straightforward is what is required for this feedback transaction to take place. First, there must be a source or producer of direct feedback within the domain. Second, and perhaps less obvious, for the expert-in-training to *receive* feedback from this source *as* a feedback signal the expert-in-training must be able to recognize what *counts* in the first place as success versus error. Or if the expert-in-training does not recognize what counts as success versus error, then the expert-in-training’s source of instruction or source of conditioning (a teacher, training programme, conditioning regime, etc.) would need to possess this recognitional capacity on some level. But the general lesson is this: if the practitioner (or instructor, etc.) does not know the difference between successful and unsuccessful performance, then the practitioner cannot correct errors and improve performance over time. For example, if a chess player continually identified getting checkmated with chess success, then regardless of the amount of chess practice and feedback-producers, the chess player will not improve at chess.

We can call this requirement of accurate success and failure identification for receiving direct feedback the “success-identification requirement”:

Success-Identification Requirement (FB \supset SI): a philosopher can receive direct philosophical feedback about philosophical theories (FB) only if that philosopher can know or reliably recognize what would count in the first place as an accurate philosophical theory (SI).¹⁴

This *Success-Identification Requirement* is implicit in both Weinberg et al. (2010) and Clarke (2013). It is also relevant to related discussions about “calibrating” intuitions from Cummins (1998). But among the critics, Ryberg (2013) arguably comes closest to an explicit acknowledgement of the requirement with his “quality parameter”. Here is Ryberg explaining why the requirement (or parameter) is met in chess and math but not moral philosophy:

[in chess and math] there exists a clear quality parameter – the validity of proofs and the victory in games ... In contrast, while the philosopher may have engaged in many cases of intuition-based reasoning, it seems much less plausible to hold that she has prior experiences of having made intuitive judgements which led to correct moral answers. (Ryberg, 2013, p. 8)

The last line is the most important. The *reason* that the moral philosopher cannot string together the developmentally critical prior experiences of *having made correct moral judgements* is that she (much like a hypothetical researcher monitoring the accuracy of her judgements) has no way of knowing which moral judgements were correct in the first place.¹⁵

Meeting the *Success-Identification Requirement* is a tall order for the defender of philosophical expertise, whether one is defending the epistemic virtue of philosophical intuitive judgements or philosophical concepts and theories (e.g., ES). For example, it is not enough for the expert-in-training to perform a task or generate a theoretical model and then judge, perhaps in reference to post-task information from the environment, “yes, *this* indicates success”. Nor is it enough if the entire community of professionals concurred in that judgement. For the requirement to be met, the judgement must be correct. More precisely, the *Success-Identification Requirement* requires knowledge rather than mere belief about success.¹⁶

I think that this insight regarding the importance of success-identification for the receipt of reliable feedback gets to the root of the dilemma for defenders of

14 As mentioned above, the practitioner might receive accurate information about success from an instructor or conditioning regime, in which case the *Success-Identification Requirement* is pushed back one level to the instructor or the design of the conditioning mechanism.

15 Or consider Clarke’s contrast between nurses and epistemologists (n. 8).

16 It is worth pointing out that an expert’s knowledge of what counts as institutional success (awards, degrees, prestige, etc.) is not sufficient to satisfy the success-identification requirement. We would need independent reasons to think that institutional success corresponds to success at developing and judging accurate philosophical theories. But to know that, we would need to know which theories were accurate, so we arrive back at the initial problem. And of course, there are institutions that confer institutional success in the absence of reliable performance. Success in the institutions of astrology and dianetics are notable examples.

Expert Superiority. There are various expert domains for which investigative questions are: (a) significantly underdetermined by available empirical methods and evidence, and (b) not explicable according to the rules of a formal system (compare to chess). I will call investigative domains that meet these two conditions “theoretical” or “armchair” domains.¹⁷ For these domains, which include theoretical physics, several research programmes within the cognitive sciences, and most parts of philosophy, we can always plausibly question whether the classification “this is correct” is itself correct. Given the *Success-Identification Requirement* and *Feedback Requirement* as codified above, it follows that we can always plausibly challenge whether training experts can receive feedback *as* feedback in these domains, and thus we can always plausibly challenge whether quality expertise can be developed through training in these domains.

Let’s take stock. We began by questioning whether we are warranted in claiming that expert philosophers generate more accurate philosophical theories than novices. I called this claim *Expert Superiority* (ES). We next determined that we are warranted in asserting *Expert Superiority* only if we are warranted in asserting that expert philosophers have been sufficiently exposed to direct learning feedback, and I called this the *Feedback Requirement* (ES \supset FB). We then determined that an important source of scepticism about philosophers’ exposure to direct learning feedback derives from the (often implicit) claim that to receive philosophical feedback about a theoretical schema or model, one must know or have a reliable recognitional capacity about what counts as a successful or accurate philosophical schema or theory in the first place. I termed this the *Success-Identification Requirement* (FB \supset SI).

Now, if we add to these claims the rather appealing premise that *philosophers do not know which philosophical schema and theories are accurate* (\sim SI), then we arrive at the following valid argument against *Expert Superiority*:

A sceptical argument against expert philosophical theorizing:

- (1) ES \supset FB *Feedback requirement*
- (2) FB \supset SI *Success-identification requirement*

¹⁷ I do not assume that being an investigative armchair domain in this sense is an all-or-nothing affair. It is not unusual for domains traditionally construed as armchair domains to pursue investigations that have empirical or formal deciders (e.g., certain logical investigations within philosophy), and domains traditionally construed as empirical or formal to have some investigations that lack such deciders (e.g., abstract ecological theorizing). A further point of clarification is that I am not using the expression “armchair domain” to reference investigations that merely have the aim of mapping logical space (for example, projects in what Strawson termed “descriptive metaphysics”). Instead, I am using that expression to refer to investigations that aim to tell us something about the objective world and where the accuracy or outcomes of these investigations is at present substantially underdetermined by both formal and empirical methods.

- (3) *~SI Insufficient knowledge about what counts as an accurate philosophical theoretical model*
- (4) *Therefore ~ES Expert theories are not superior*

I need to explain further the role of premise {3} in the above argument. To do so, I first need to discuss an important contrast between what accounts for experts' failure to receive adequate feedback in the domains studied by empirical investigators (stock brokerage, etc.), on the one hand, and what would account for expert philosophers' failure (if there is one) to receive adequate feedback about their theoretical models, on the other.¹⁸

In the professional domains empirically shown to nurture unreliable expertise, what explains the absence of direct feedback? There are two task-dynamics in these "expert-dubious" domains that typically explain the deficiency of feedback. First, experts in these domains are generally tasked to make predictions. Second, there is typically a significant time lag or other epistemic disconnect between prediction and assessment. Consider parole experts, who, as revealed by follow-up studies, make less reliable predictions about criminal recidivism than simple regression models (see, e.g., Carroll et al., 1982). The explanation for this poor performance is that, even when institutional mechanisms are in place to deliver outcome data back to the expert predictors (often they are not), such data will arrive many years after the initial decision and when the experts have "moved on". A similar dynamic applies to experts making judgements about who deserves graduate admission, or which stocks have long-term value. The empirical literature on poorly performing experts is rife with these types of examples.¹⁹

I submit that this pattern of expert failure reveals an important but overlooked insight regarding the developmental role of feedback. While there *is* a feedback problem in these expert-dubious domains, that feedback problem does *not* generally stem from the expert's inability to recognize or know what would count in

18 Given their strategy of employing the empirical literature on expert performance to argue that expert philosophers require sufficient exposure to direct feedback to develop epistemically virtuous (i.e., truth-tracking) concepts, rules and theories (informational structures that are candidates for underwriting truth-tracking intuitive judgements), there are reasons to think that critics of philosophical intuition like Weinberg et al. (2010), Clarke (2013) and/or Ryberg (2013) are implicitly committed to an argument like the above. But I will not press this interpretation here. The more important point, as far as I am concerned, is that if one's goal is to harness the empirical literature on non-philosophical expertise as a sceptical challenge to the epistemic status of expert philosophical theories, and if one takes into account the role that success-identification plays in the receipt of reliable feedback, one is led directly to an argument like the above.

19 See, e.g., Meehl (1954) on clinical prediction, Dawes (1971) on predicting the future success of graduate students, Armstrong (1978) on economic forecasting, Carroll et al. (1982) on predicting criminal recidivism, Tetlock (2005) on long-term political prediction, and Camerer and Johnson (1991) on prediction generally.

the first place as a successful versus erroneous performance. For example, political forecasters understand that a third world war would mean that a forecaster's prediction of unending peace was mistaken (even if they disagree on whether, in fact, a third world war is likely). An emergency care doctor who diagnoses a patient with heartburn but who finds out 10 years on that the patient was in cardiac arrest understands that this indicates an incorrect diagnosis. And so on. The feedback problem for these domains instead generally stems from challenges specific to the diachronic features of the domain's relevant tasks.

The next point to emphasize is that, *if* philosophy were to have a feedback problem, then that problem could not be the same type of feedback problem. Philosophical theories aim to be empirically adequate, but they do not generally have predictive content or predictive aims. In other words, philosophy is not a predictive discipline. Thus, if philosophers are insufficiently exposed to direct feedback, we could not blame this fact on the types of time-lags and related institutional entanglements that are typically to blame in expert-dubious domains like stock brokerage and political forecasting (see Bach, 2019, for a more detailed discussion and argument). If philosophers are insufficiently exposed to direct feedback, it would have to be caused by something else.

But what? Unlike the expert-dubious domains studied by empirical researchers, it is not obvious that experts in armchair domains and philosophy in particular have knowledge of what counts in the first place as an accurate theoretical model. If that were the case, then the feedback artery would be blocked at the *Success-Identification Requirement* (i.e., \sim SI).

I can further illustrate this contrast by employing a distinction used by the logical positivists. The positivists distinguished between propositions that are verifiable in principle and propositions that are verifiable in practice, both of which they claimed were meaningful.²⁰ The typical feedback problem for expert-dubious domains is that, even though experts' claims are verifiable in principle, and even though experts share knowledge about what these verifications conditions are (thus satisfying the *Success-Identification Requirement*), these conditions are often difficult to verify in practice. In the domain of philosophy, on the other hand, it is not obvious that experts have knowledge (beyond tautological transformations of their own theoretical commitments) of which states of affairs could *in principle* indicate the success of one rival philosophical theory over another. This is a fundamental problem with respect to satisfying the *Success-Identification*

20 The proposition "Jupiter has a rocky core" is verifiable in principle but not in practice. According to Ayer (1936, p. 36), the proposition "the Absolute enters into, but is itself incapable of, evolution and progress" is verifiable in neither practice nor principle.

Requirement, and it is quite different from the diachronic feedback problems of other domains.²¹

If this analysis is correct, then feedback-based sceptical challenges to philosophical expertise like the one that we have been considering place a great deal of argumentative weight on the *Success-Identification Requirement* not being satisfied (\sim SI). It follows that feedback-based arguments against expert philosophical (and armchair) theorizing diverge in an interesting respect from feedback-based challenges to expert performance in other domains (e.g., the domains investigated by empirical researchers). Given this argumentative weight placed on \sim SI, we need to consider whether there are good grounds for asserting \sim SI or if defenders of *Expert Superiority* can plausibly defend SI.

2.2 *The role of expert agreement in the sceptical argument against expert philosophical theorizing*

The difficulty for defenders of *Expert Superiority*, as I see it, is to provide a non-question-begging argument in support of the *Success-Identification Requirement* (SI). In fact, it would seem that the more plausible case can be made for \sim SI. For example, if expert philosophers knew which theories for a philosophical topic *P* were accurate in the first place – or if they knew which factors counted as verification conditions for a theory of *P* – then philosophical discussion about *P* would be sharply limited. But of course, there continues to be vigorous and ongoing debate over which philosophical theories are accurate. Indeed, the idea that philosophers do not generally know which philosophical theories are accurate, or which set of conditions count as favouring one rival theory or another, seems obviously true and hardly worthy of argument.

This sense of obviousness appears to play a role in recent critiques of philosophical expertise. For example, to prevent their argument against philosophical reliability from generalizing into a global and less plausible scepticism about all expert philosophical theorizing, Weinberg et al. (2010) allow for local instances in which they say that philosophers have developed epistemically virtuous theoretical distinctions. In particular, Weinberg et al. flag expert philosophers' theories pertaining to the use–mention distinction, the semantic/pragmatic distinction,

21 Consider two philosophers who disagree about the mind–body problem. A natural interpretation is that they are not disagreeing about what the future will be like. Instead they are disagreeing about which factors are relevant to and support the truth of a general claim about what types of properties or substances currently exist. The physicalist might appeal to facts about historical scientific reductions or the nature of causal processes. The dualist might appeal to conceptual facts concerning the imagining of zombies or people trapped in black and white rooms. Both philosophers tend to think that conditions currently on hand are sufficient to tip the scale one way or another, but they disagree on what those conditions are and in which direction they tip the scale.

and the epistemological/metaphysical distinction as “successful” and as “key distinctions that philosophers have wrung from their theorizing” (Weinberg et al. 2010, p. 342). However, they say that these are “exceptions that prove the rule” (Weinberg et al. 2010, p. 342).

In assessing the epistemic privileging of these or any other theoretical distinctions, care is needed to avoid an equivocation on the notion of epistemic “virtue” or “success”. It is one thing to have success at following a rule, and it is another thing for the rule itself to be successful at achieving its purpose. For example, a quack doctor can provide pupils clear and immediate feedback for rules about the application of snake oil, and such training will cause the pupils to develop great success at following the rules of snake-oil medicine. Of course, neither the doctor nor the pupils will have success at healing patients. In this case, the *Success-Identification Requirement* is not met because the practitioners clearly lack knowledge about what is successful, and this prevents them from receiving adequate feedback.

Applying this point to Weinberg et al.’s exempted distinctions, it would not be enough that philosophers consistently follow a rule for marking, e.g., the epistemological/metaphysical distinction. It would also need to be the case that the rules themselves are epistemically virtuous rules and not the philosophical equivalent of snake-oil rules. Only then would the *Success-Identification Requirement* be met. Thus, for Weinberg et al. to be correct that reliable philosophical expertise can be developed in reference to the privileged distinctions, it must be the case that the rules and theories that give rise to these distinctions accurately describe the world (if this is what it means to be philosophically successful) and that experts reliably detect that they do so.

But what allows Weinberg et al. to make these latter assumptions? Here they simply stipulate that the rules are “key”, “important” and “epistemically successful”, and have “proven useful”. This does not strike me as a significant defence. After all, any philosopher who is challenged to defend the accuracy of his or her favoured philosophical theory (about moral properties, content externalism, free will, etc.) can make similar boasts and support them with argument and empirically adequate theoretical models.

There is, however, a crucial difference between Weinberg et al.’s boast and these other boasts. That difference concerns the extent of agreement within the community of expert philosophers. While there is robust agreement in the expert community about the epistemic goodness or success of Weinberg et al.’s exempted distinctions, there is little agreement in the expert community concerning the epistemic goodness or usefulness of most other theories and theoretical distinctions.

I submit that this difference in expert agreement is what implicitly drives a powerful but ultimately misleading intuition, or sense of obviousness, concerning the rejection of the *Success-Identification Requirement* (~SI). Here is how I see that intuition or sense of obviousness functioning in the sceptical argument

against expert philosophical theorizing. Begin with the fact that, beyond Weinberg et al.'s (2010) exempted theoretical distinctions, the community of expert philosophers does not collectively agree on much of anything (and when agreement does happen, it tends to be fleeting). Infer from this that, at least for the bulk of philosophical theories, the expert community as a whole does not know which theories are accurate and which are inaccurate (if the community cannot agree, it cannot know). Because success-identification requires knowing or reliably recognizing what counts as success (i.e., accuracy) and what counts as failure (i.e., inaccuracy), infer that for the bulk of philosophical theories expert philosophers cannot identify which count as successful. Because success-identification is a necessary condition for the receipt of reliable feedback, infer that training philosophers do not adequately receive feedback with respect to the bulk of philosophical theories. Finally, conclude that training philosophers do not develop reliable expertise with respect to the bulk of philosophical theories.

3. Deflating the Sceptical Argument against Expert Philosophical Theorizing

The line of reasoning exposed in the previous section makes explicit a premise that does important work in the sceptical argument against expert philosophical theorizing:

Expert Community Lacks Knowledge (\sim CK): The community of expert philosophers does not know which philosophical theoretical models count as accurate and which count as inaccurate.

\sim CK follows straightforwardly from substantial disagreement, both contemporary and historical, within the community of expert philosophers. On most accounts of what it could mean for a group to have knowledge, we should not doubt the truth of \sim CK with respect to most philosophical theories. But as I will explain here, \sim CK cannot ground or even provide reasonable evidence for a plausible scepticism regarding *Expert Superiority*.

There are two ways that \sim CK could be used to reject *Expert Superiority*, and each way fails. First, one could reject *Expert Superiority* after conflating \sim SI and \sim CK. The conflation is tempting, but \sim CK and \sim SI mean different things (section 3.2) and rejecting *Expert Superiority* follows validly only from \sim SI. Second, one can add a premise to the sceptical argument, premise {3} below, that claims \sim CK entails or provides strong evidence for \sim SI:

- (1) $ES \supset FB$ *Feedback requirement*
- (2) $FB \supset SI$ *Success-identification requirement*
- (3) \sim CK $\supset \sim$ SI *If the community of expert philosophers does not know which philosophical theories are accurate, then it is not the case that expert*

philosophers know or reliably recognize which philosophical theories are accurate

- (4) \sim CK *The expert community lacks knowledge of which philosophical theoretical models are accurate*
- (5) Therefore \sim ES *Expert philosophical theoretical models are not superior*

While this version of the sceptical argument is valid, premise {3}, even if considered as a claim about evidential grounding, is false. I explain why below.

3.1 Two types of philosophical expertise

Distinguish between two kinds of philosophical expertise. The first we can call “procedural expertise”. It refers to the ability to use logic and argument analysis, to understand and employ philosophical distinctions effectively, to understand and reference effectively the history of philosophy, and so forth. This type of expertise is a competency with the “philosopher’s toolkit”. Expert philosophers possess a more trustworthy version of this competency than novices, and the reason is straightforward: success here is stipulated as a type of rule-following, and to be trained in philosophy is to receive ample feedback from teachers and others regarding successful versus erroneous uses of the toolkit. Using the toolkit successfully in this procedural sense does not presuppose that the rules themselves are successful at achieving a more distal purpose – for example, carving out natural kinds or constructing accurate theoretical models. If it turned out that the toolkit’s rules were the philosophical equivalent of snake-oil rules, then it would still be the case that trained philosophers develop a more reliable form of philosophical procedural expertise than novices. Of course, in this scenario we should question the value of reliable procedural expertise.

We can call the second type of expertise “content expertise”. It refers to the development, possession, judgement and application of accurate (true, assertable, etc.) philosophical theories, whether these are about material constitution, moral properties, the mind–body problem, scientific explanation, the analysis of knowledge, and so forth. If we assume that the philosopher’s toolkit is generally stocked with good rules rather than snake-oil rules, then reliable content expertise would presuppose reliable procedural expertise. However, procedural expertise underdetermines content expertise. For example, philosophers who are roughly equal in their procedural abilities can arrive at contrary content positions. More precisely: current expert knowledge of what counts as procedurally successful is generally not sufficient to adjudicate conflicts over what counts as content-successful.²²

²² There are of course cases in which the distinction between procedural and content expertise blurs. This is especially likely if the best way to apply a philosophical device depends on the content to which it is applied.

The comparative claim at stake in this article – *Expert Superiority* – refers to content expertise. Understood as a claim about content-reliability, *Expert Superiority* cannot be defended in the simple way that reliable procedural expertise can be defended, i.e., by checking whether a rule was followed. However, I submit that there are distinct grounds for a reasonable defence of (content-based) *Expert Superiority*. To develop that defence, I first acknowledge an important difference between the distribution of reliable procedural expertise, on the one hand, and the distribution of reliable content expertise, on the other.

3.2 *An uneven distribution of philosophical content expertise and its implications for the defence of expert superiority*

Expert philosophers are collectively superior to novice philosophers with respect to procedural expertise. There is also a more or less even distribution of reliable procedural expertise among professional expert philosophers. The point here is not to deny individual differences – surely some philosophers are more skilled with the toolkit. The point is that generally all expert philosophers are procedurally reliable in the same way that generally all trained meteorologists are reliable meteorologists.

We should not have the same conviction about an even distribution of reliable philosophical content expertise. Reliable philosophical content expertise is likely more difficult to achieve and thus less evenly distributed. This should not be news. Philosophers have always believed that “so-and-so” has the right take on this or that topic, that “so-and-so” is wrong about this or that commitment, that philosophers from “such and such school of thought” are wrongheaded in their allegiance to this or that principle, and so on. Regardless of what philosophers believe about the accuracy of their own theoretical commitments, they tend to think that, with respect to the theoretical commitments of the broader professional community, only some philosophers are content-reliable some of the time. This sociological point is as elitist as it is platitudinous.

Assume for the moment that this “common view” is apt and that there is an even distribution of reliable procedural expertise but an uneven distribution of reliable content expertise. In this scenario, and assuming that the philosophical toolkit has good rules rather than snake-oil rules, we would be able to trust specifically procedural philosophical expertise in roughly the same way that we can trust expertise in chess and the applied sciences. However, we would not be able to trust philosophical content expertise in quite the same way that we trust expertise in these other domains. Early articulations of the expertise defence, which were built on an analogy to expertise in chess and science, miss this disanalogy. As a result, those defences were open to cogent criticisms based on differences between philosophy and other domains (see, e.g., Ryberg, 2013).

But crucially, if the common view is apt, then it would *not* follow that the contentful theories of philosophical experts do not deserve greater credence than those of novices (the denial of *Expert Superiority*, or \sim ES). If there is an uneven distribution of content-reliability, then some or many expert philosophers tend to engage in content-success or accuracy while some or many tend to engage in content error. The former group, whether it is 10% or 70% of the expert philosophical community, will have more accurate philosophical theories than novices. It will also be the case that on average, expert philosophers have more accurate philosophical theories than novices.²³

If true, these assumptions about the distribution of content-reliability have several important implications for the sceptical challenge explicated in section 2. First, we would have a straightforward explanation of *Expert Community Lacks Knowledge* (\sim CK). Given fracturing within the expert philosophical community with respect to content-reliability, the expert community as a whole could not know which philosophical theories are accurate and which are inaccurate. We would also expect the sorts of expert disagreements that serve as evidence for \sim CK.

Second, we would have a straightforward explanation of why \sim CK neither entails nor means the same thing as claiming that the *Success-Identification Requirement* is not met (\sim SI). The *Success-Identification Requirement* refers to the ability to recognize reliably, or know what counts as, a philosophical content-success. It is possible for an individual or a group of individuals to possess this epistemic ability even if the wider community of which these experts are members does not collectively possess this ability. Compare: expert jazz flutists have reliable jazz-flute expertise even if the wider expert community of which they are members – the community of expert jazz musicians – does not collectively possess jazz-flute expertise.

Third, we will have reverse-engineered a misleading intuition or sense of obviousness concerning untrustworthy expert philosophical theorizing. At the conclusion of section 2.2, I explained how it is easy to go from an obvious fact about expert disagreement to a claim about a missing quality parameter, then to a claim about deficient feedback, and then finally to a claim about unreliable expertise.

23 In sections 4.2 and 4.3 I defend the assumption that there is a higher percentage of philosophical content-reliability (however unevenly distributed) in the expert philosophical community than in the novice community. I do so by offering a model that explains how expert philosophers can in general develop greater content-reliability than novices. While that model cannot *guarantee* that expert philosophers in general possess greater content-reliability, it does provide a naturalistically plausible account, consistent with the empirical study of expert performance and the developmental importance of feedback, of how expert philosophers in general develop greater content-reliability.

That way of arguing fails. As indicated by the previous implication, the first inferential step relies on either a conflation or a scope confusion.

But is the assumption that there is some non-negligible (perhaps substantial) percentage of content-reliable expert philosophers justified? I argue in the next section that it is.

4. A Positive Account of Feedback for the Development of Philosophical Theoretical Models

Returning to the jazz analogy, it is easy enough to recognize that some but not all members of the expert jazz community possess jazz-flute expertise. Even a novice like me can correctly identify Hubert Laws but not Ella Fitzgerald as possessing reliable jazz-flute expertise. It is also easy to offer a developmental explanation of this difference: jazz-flute training confers reliable jazz-flute expertise, and only some jazz musicians are trained in jazz-flute. In contrast, and assuming the uneven distribution of content-reliability discussed in section 3.2, the task of recognizing which expert philosophers are content-reliable is fraught with difficulty and inevitably controversial. Central to that difficulty is that, as noted in section 2, we cannot prove or empirically confirm who is content-reliable or even *that* there is a non-negligible percentage of content-reliable expert philosophers. To prove or confirm as such would presuppose that we were already in possession of either formal or scientific devices for determining the accuracy of the first-order philosophical theories over which expert philosophers disagree. We do not currently have these devices. Such are the epistemic conditions that define what I earlier termed “armchair domains”. In these domains, it is presumably the purpose of theoretical modelling to throw light into the epistemic voids created by the absence of such devices.

Going without formal and empirical deciders makes for risky epistemic exploration. Philosophers and other armchair professionals generally accept this epistemic condition as an occupational hazard. But there is an important difference between risky conditions and futile conditions. It is risky to set sail without a GPS or a compass, but one can still achieve navigational success if one knows how to use the position of celestial bodies as indicators of one’s position. Similarly, while it may be epistemically risky to pursue armchair exploration, there can still be reliable indicators about the relative success or accuracy of competing, empirically adequate theories. If there were such indicators, then they would be sources of feedback. And if a philosopher could know or reliably recognize such signals for what they

were, then this would satisfy the success-identification requirement for the receipt of feedback.

Are there such indicators? I think that there are. But before spelling out the details, I should introduce two additional desiderata for an account of philosophical feedback that follow from my earlier discussion. First, given that philosophers are generally without formal and empirical deciders for competing theories, we should expect the relevant indicators, were they to obtain, to be different in kind from those found in domains like chess and the applied sciences. Second, and given the conjecture from section 3.2 about an uneven distribution of content-reliability, we should expect such indicators, were they to obtain, to be knowable but difficult to know.

4.1 Theoretical virtues as a source of philosophical feedback

Researchers within and outside of philosophy have long recognized the importance of theoretical or “super-empirical” virtues, and they have done so independent of the sceptical challenge discussed in section 2. In this section, I build on these previous insights and make clear how theoretical virtues can serve as a source of feedback that improves the reliability of philosophical content expertise.

It is widely accepted that theoretical properties like parsimony, explanatory power, elegance and fertility are reliable indicators of a scientific model’s success. Many philosophers claim that these properties play a similar indicating role in the evaluation of philosophical theoretical models. On this view, when a philosophical model exemplifies these virtues, this is evidence that the world is the way that the model describes – at least more so than what is described by empirically equivalent models that lack these virtues. This line of thinking has become common, especially in naturalistic defences of metaphysics.²⁴

Of course, not all philosophers accept that theoretical virtues have epistemic rather than mere pragmatic significance (see, e.g., Ladyman, 2012; Saatsi, 2017). I will not enter this debate here. I will assume but not argue for the claim that theoretical virtues have epistemic significance (but see the arguments by the authors listed in n. 24). The following proposal is thus a conditional one. If it turns out that theoretical virtues lack epistemic significance, then a different case must be made for *Expert Superiority*.

²⁴ Its defenders include Swoyer (1999), Ellis (2009), Sober (2009), Sider (2009), Paul (2012), Nolan (2015) and Andow (2015).

I can sketch an application of the proposal to the challenge discussed in section 2 as follows. A philosophical theory's exemplification of theoretical virtues and particularly its explanatory and unificatory power constitutes a feedback signal indicating that theory's accuracy.²⁵ A theoretical model's failure to exemplify these virtues is a feedback signal indicating inaccuracy. The *Success-Identification Requirement* is then satisfied to the extent that philosophers can know whether and to what extent philosophical models exemplify these features.

There are several challenges for making this account more concrete. For example, can we defend the legitimacy of generalizing the role of theoretical virtues from scientific domains like evolutionary biology and astronomy to armchair domains like philosophy? Here is what Paul says in defence of this idea:

.... if simplicity and other theoretical desiderata are truth conducive in scientific theorizing, they are truth conducive in metaphysical theorizing. This is a central part of my thesis: if we accept inference to the best explanation in ordinary reasoning and in scientific theorizing, we should accept it in metaphysical theorizing. (Paul, 2012, p. 22).

While I agree with Paul, it is important to see why this would not suffice as a response to the critical challenge explicated in section 2. Critics of philosophical expertise can reply to Paul's suggestion by claiming that empirical studies of expert performance provide grounds for challenging whether the right type of continuity – the type that could justify generalizing the indicator role of theoretical virtues – indeed holds between reliable scientific modelling and philosophical modelling. This challenge about continuity would have both a metaphysical and an epistemological component. Concerning metaphysical continuity, critics could rightly request an empirically plausible account that explains how the targets of philosophical models literally serve as sources or providers of feedback.²⁶ (And if defenders of *Expert Superiority* were to claim that theoretical virtues are the sources of philosophical feedback, then the critics could rightly request an empirically plausible account of how the targets of philosophical models literally give rise to such virtues.) Regarding the epistemological continuity, critics could rightly request an empirically plausible account that explains how philosophers reliably identify such feedback signals for what they are.

In response to both hypothetical requests, the next two sections advance programmatic considerations about the metaphysics and epistemology of theoretical

25 It is worth noting here that information-carrying relations are not restricted to causal relations (see, e.g., Dretske, 1988).

26 See, e.g., the concern expressed by Clarke in n. 8.

virtues as these relate to feedback mechanisms and the development of philosophical content-reliability.

4.2 Some remarks on the metaphysics of theoretical virtues

In the sciences, what grounds the connection between accurate theoretical models and theoretical virtues? A standard answer is that scientific models exhibit explanatory power just in case they identify, track and/or describe a natural kind. At the root of this answer is the conviction that the world is not, as Locke once speculated, one in which all things “differ but in almost insensible degrees” (Locke, 1975, III, iv, 12). Instead there are “gaps” and “chasms” (which Locke denied), or what Millikan (2010) calls “clumps and clots”, or what Boyd calls “homeostatic property clusters”, or what we can simply call “natural kinds”. The common idea here is that while much of logical space is empty, in select locations of that space the laws of nature cause the clustering of particular property syndromes. Given the way that the world is, some property syndromes are generally stable – H₂O, *Canis familiaris*, geodes – and some are not – HO₂, talking cucumbers, lamppost-supernova hybrids, etc.

The epistemic side of this coin is that, if you get on the track of one of these stable property syndromes – perhaps by description, perhaps by mere ostension – then you can reap explanatory rewards.²⁷ And if you have a model that is deficient in explanatory pay-off, then probably it has failed to identify or describe such a natural kind. Aristotle’s *superlunary* category, which grouped together all things outside the moon’s orbit, offered no explanatory pay-off. Following Griffiths (1997, p. 171), this was a clear sign that the *superlunary* category failed to identify a natural kind.

We are building to the idea that a successful philosophical model will place investigators on track of a natural kind. Such a model would corral investigative efforts toward one of these occupied sections of logical space and as a result exhibit theoretical virtue and particularly explanatory power. But returning to the critics’ perspective, one can object here that scientific natural kinds are accessible in a way that the kinds targeted by philosophical models are not. This difference in accessibility (or perhaps underlying reality), the objection continues, undercuts generalizing the explanatory role of scientific kinds to purported philosophical kinds. This type of objection generalizes Mackie’s well-known “argument from queerness” against moral realism. It balks at the plausibility of there being a type of information about the

²⁷ This connection between kinds and explanation is a theme that runs through the work of Quine, Boyd, Millikan, Kornblith, Griffiths and many others. Kornblith (1995) is a particularly clear statement.

epistemic goodness of a philosophical theory that is literally produced (the metaphysical queerness) and literally received (the epistemological queerness).²⁸

The Mackie-style objection is less forceful than it seems. It would be more forceful if defenders of explanatory philosophical kinds were restricted to thinking of such kinds as metaphysically exotic or separate – as Platonic forms or abstracta or some such. But there is no reason to think of explanatory philosophical kinds in that way. Instead, we should view the type of natural kind implicated by successful philosophical models as continuous with the types of kinds implicated by successful scientific models. This continuum does not advance from the scientifically acceptable to the metaphysically bizarre. Rather, it advances from kinds that accommodate explanatory models with limited scope but considerable force to kinds that accommodate models with more limited force but considerable scope.²⁹ As the continuum advances in this direction, higher-order relational properties take on an increasingly central explanatory role.³⁰

28 Both Pust (2017) and Schechter (2010) describe related objections that they term “the argument from inexplicable reliability” and “the reliability challenge” respectively. As Schechter notes, the general objection is an old one, and it has various analogs (the Benacerraf–Field problem for mathematical Platonists, Peacocke’s integration challenge, etc.). The specific objection I am considering here can be viewed as a narrow version of the general objection. This more specific objection works by resourcing the empirical literature on expert performance to vindicate and explain, in terms of a feedback deficiency, the general reliability objection. Viewed this way, I am addressing whether such a vindication and explanation indeed holds in armchair domains generally and the philosophical domain specifically.

29 On the relationship between natural kinds and these two features of induction (scope and force), see Griffiths (1999). For a closely related discussion in terms of explanatory breadth versus guidance, see Blanchard et al. (2018).

30 A few qualifications are in order. The claim that higher-order relational kinds have important epistemic status enjoys strong support. For example, in developmental psychology, Dedre Gentner and colleagues have convincingly demonstrated the importance of higher-order relational kinds for cognitive development and analogical reasoning (see, e.g., Gentner, 2003; Markman and Stillwell, 2001; Bach, 2012). In the philosophy of reference, Richard Boyd has advanced the influential view that referential practices, including those conducted in what I am calling armchair domains, provide access to natural kinds “whose real essences consist of complex relational properties, rather than features of internal constitution” (Boyd, 1979, p. 358). Boyd’s view makes sense of how economic models about *money*, for example, are explanatorily fruitful if they coordinate reference to a complex of relational facts about social convention and human rationality (and unfruitful if they were instead to coordinate reference to the microstructural features of currencies). Pushing this idea even further, Boyd explains how metaphors like “the mind is a computer” can function as a type of ostension that place investigators on the track of better understanding an abstract relational kind, even if that kind’s “fundamental properties have yet to be discovered” (Boyd, 1979, p. 364). The claim that there are philosophical natural kinds that have relational essences is a natural extension of this idea. If philosophical concepts like free will, counterfactual causation, reference and moral goodness refer at all, then they would likely refer to natural kinds with abstract relational essences in this sense. And if that were the case, then philosophical models that coordinated reference to these kinds would offer more explanatory power than empirically equivalent rival models that directed investigators to less populated regions of logical space. Boyd himself develops the view that “moral goodness” refers to such a natural kind. Of course, there are plenty of philosophers

An example will help explain the idea. Ruth Millikan's theory of proper functions claims (very roughly) that we can determine an item's function in reference to the activity performed by its ancestors and that helps account for the proliferation of the item's lineage. It is the proper function of the human liver to detoxify blood, of beaver slaps to warn of danger, of car headlights to illuminate driving terrain, and (more controversially) of beliefs to provide maps that bring about the satisfaction of desires. We do not currently possess empirical tests or formal methods for determining whether Millikan's model is more accurate – more reflective of nature's stable and explanatorily rich clusters of relations – than empirically adequate but rival models (e.g., Cummins's 1975 dispositional analysis of functions). Nonetheless, Millikan's etiological account is widely embraced, and for good reason: the account offers considerable explanatory and unificatory power. For instance, it provides a common model for the function of artefacts, biological items, language and cognitive items. Millikan's account is also simple and elegant, explaining misrepresentation in terms of malfunction – a phenomenon already accepted in our ontology – whereas rival theories tend to make post-hoc adjustments or appeal to Meinongian fictional objects. And perhaps most important for our purposes, there is nothing metaphysically exotic or offensive about the underlying kind to which Millikan's theory coordinates reference. The term "proper function" describes a complex of relations that includes a type of ontogenetic relation (i.e., a replicative or reproductive relation between an item and a historical lineage) as well as a type of phylogenetic relation (i.e., between a historical lineage and its historical environment with reference to past activities that conferred fitness and help explain the reproduction of the lineage). Given the way that the world is, *this* set of relations happens to cluster and repeat. Conceptual tools like Millikan's theory of proper functions that help organize our epistemic practices around this induction-supporting kind non-accidentally yield explanatory success.

who think that Boyd's model fails to provide epistemic access to an explanatory kind in the way that he contends. Such critics claim that there is no such natural kind, or that moral discourse is better explained by an alternative model. Generalizing this concern, it is possible that there are no philosophical natural kinds in the sense described above. If that were the case, then the best philosophical work would generally be negative: some philosophers posit relational kinds and flag their explanatory power, and other philosophers correctly maintain that the explanatory relation and the grounding kinds are absent. But note that, even in this scenario, there would be a difference between philosophical content-reliability and content-unreliability. (Sider, 2011, p. vii, makes a similar point: "Those who say that questions of ontology are 'merely verbal', for example, are best regarded as holding that reality lacks ontological structure. Such deflationary metametaphysical stances are thus themselves metaphysical stances".) And while I think that the general account on offer above has important implications for philosophical investigations into descriptive metaphysics (in Strawson's sense) as well as modal reality, I leave the exploration of those implications to another occasion.

Of course, philosophers can and have challenged whether Millikan's theory best exemplifies these theoretical virtues. I take up concerns about the epistemology of theoretical virtues in the next section.

4.3 Some remarks on the epistemology of theoretical virtues

Even if we were to accept that there are such philosophical natural kinds and that models that coordinate reference to them exhibit theoretical virtue, do we have reason to think that expert philosophers can reliably recognize which competing philosophical models possess theoretical virtue? For example, who is to say whether Millikan's theory of functions (or Boyd's theory of moral goodness, etc.) exemplifies greater theoretical virtue than rival models? This is the challenge of success-identification but applied to the positive claim that theoretical virtues are proximal indicators of success. Recalling that success-identification is a necessary condition for the receipt of feedback (section 2), defenders of *Expert Superiority* who recruit theoretical virtues as proximal success-indicators owe an answer to such questions. In addition, that answer needs to indicate why some but not all philosophers reliably recognize which philosophical models have greater theoretical virtue. Otherwise, the claim that there is an uneven distribution of reliable content expertise – an important claim for the defence of *Expert Superiority* – is unmotivated.

While it is important to recognize this explanatory burden for defenders of *Expert Superiority*, it is equally important not to impose uncharitable demands on such an explanation. For example, it would not be surprising if philosophers (or cognitive psychologists who might study philosophers) lacked explicit knowledge of the psychological processes and mechanisms through which some philosophers – and armchair experts more generally – reliably parse subtle differences in the theoretical virtue-profiles of competing philosophical models.³¹ If that were

31 Nolan (2015, p. 225) makes a similar point, claiming that “avoiding inelegant or arbitrary solutions, or finding some explanations satisfying and others lacking, and other such activities can proceed in advance of being able to say what one is doing and why, just as we can justifiably use induction before being able to satisfactorily answer the riddle of induction”. I add to this that, just as there have been numerous attempts to solve the riddle of induction, there have been numerous attempts to codify theoretical virtues so as to make them more amenable to explicit reasoning. I further add that philosophical discourse is the way that one would expect it to be if content-reliable philosophers had implicit rather than explicit knowledge about a philosophical model's degree of theoretical virtue. Imagine how things would be if this knowledge were explicit: we should be able to operationalize the identification of philosophical theoretical virtue, program computationally powerful and bias-free machines with such reliable processes, and then resolve philosophical stand-offs. This might describe some far-off philosophical future (wonderful or dystopian, depending on your taste), but it does not describe contemporary philosophical discourse. Although, there have been attempts to generate such computational mechanisms. Thagard's (1989) ECHO program, which is designed to detect various parameters of what Thagard terms “explanatory coherence”, is a notable example.

the case, then expert philosophers' reliable evaluation of theoretical virtues would be a genuine but implicit epistemic skill, much like many other epistemic competencies that are not in doubt (including skills that separate reliable experts from unreliable novices).

Of course, it is not dialectically satisfying simply to gesture at the possibility of an implicit skill. More satisfying is to provide an account that, while stopping short of formalizing or operationalizing that skill, explains its development while addressing the types of concerns about feedback raised in section 2. In other words, defenders of *Expert Superiority* who invoke philosophers' reliable evaluation of theoretical virtue to meet the success-identification requirement need to explain how this expert skill develops and why it develops unevenly.

To sketch such an account, I start with a claim advanced by Nolan (2015). Nolan invokes the "*a posteriori* study of the phenomenon of theorizing" to defend his view that theoretical virtues in philosophy have epistemic and not mere pragmatic significance. According to Nolan, when we examine signature epistemic breakthroughs in science – Nolan flags the scientific revolution and Copernican cosmology – we observe an accompanying exemplification of theoretical virtues like simplicity and explanatory power. Nolan interprets this as a type of *a posteriori* evidence for the epistemic value of these virtues generally.

Now, if one's goal is to justify the success-indicator role of theoretical virtues in philosophy, then I agree that past scientific successes are important (but as discussed in section 4.2, I think that a metaphysical grounding of the application of those virtues to philosophical models is further required). But the type of account needed for present purposes belongs more directly to descriptive epistemology.³² That is, we need to understand what content-reliable philosophers have been doing that would explain how they developed skill at identifying and evaluating philosophical, theoretical virtue.

The basic outline of my proposal is that content-reliable expert philosophers (and content-reliable armchair professionals generally) have profited from the *a posteriori* study of the phenomenon of theorizing. The suggestion here is not that content-reliable philosophers have special or significant prior training in science or the theories of the scientific revolution. To get clear on what I am claiming, we first need to expand Nolan's notion of the "*a posteriori* study of the phenomenon of theorizing" well beyond the Copernican-like scientific breakthroughs that are of interest in Nolan (2015).

It is often overlooked that the phenomena of successful and unsuccessful theorizing are in wild abundance. The reason that they are so plentiful is simple: everybody theorizes about nearly everything. People routinely construct

32 See Goldman (1993) on descriptive versus normative epistemology.

theoretical models for how the mice are getting into the basement, for how their kids will turn out, for the direction of local politics, for who believes what, and so on. This habit of theorizing is a domain-general activity, subserved by domain-general cognitive mechanisms like association, analogy and inference to the best explanation. Developmental psychologists claim on empirical grounds that we are already skilled domain-general theoretical model builders during infancy. Before the age of two, we posit unobserved explanans (e.g., a false belief-state) to account for observed phenomena (e.g., intentional behaviour). Before the age of two, we employ probabilistic techniques like Bayesian inference to update and improve the accuracy of our theoretical models so that they better track causal and explanatory dependencies. In sum, we are all incessant builders of theoretical models, across domains with starkly different content, and we have always been so.³³

If the phenomenon of theoretical modelling is in wild abundance, then so too are opportunities for the *a posteriori* study of its successful instances. We are often exposed to the theoretical models of others, and we are especially exposed to our own theoretical models. We test the accuracy of these models through behavioural intervention and, more simply, by waiting to see what happens. This constant exposure to both accurate and inaccurate theories presents ample opportunity for grasping which theoretical properties correlate with successful theoretical models. It provides a wealth of epistemic resources from which we can acquire second-order knowledge of the often subtle differences between accurate and inaccurate theoretical models. The resulting second-order knowledge may be implicit, but it need not be developmentally mysterious.

Before applying this insight to the situation of the armchair expert, it is worth reflecting on how it differs from some of the developmental alternatives sketched by critics of philosophical expertise. Recall Ryberg's (2013, p. 8) claim that purported experts in moral philosophy lack the necessary training exposure to the "correct moral answers". Recall also how Clarke (2013, p. 193) faulted the training epistemologist for being unable to "directly discover what knowledge really is". These critics construe the development of reliable philosophical expertise as requiring some sort of causal or confirmational experience with the literal targets of first-order philosophical models. In contrast, the current proposal construes the content-reliable armchair expert as training against a far more accessible and prosaic phenomenon: domain-general theoretical model-building.

I suggest that there are two phases of the *a posteriori* study of the phenomenon of theorizing that are individually necessary and perhaps jointly sufficient for the development of reliable armchair content expertise. For the first phase, the

33 In support of these theorizing abilities of infants, see, e.g., Gopnik and Wellman (2012).

epistemic training resources are the domain-general, all-purpose and often mundane theoretical models described above. For the second phase, the epistemic resources are the specialized contentful theoretical models of the relevant armchair domain. I discuss each resource in turn.

When discussing expert performance, it is perhaps easy to overlook the developmental importance of phase-one theoretical models. That would be a mistake. If an armchair expert never developed an implicit grasp of the marks of theoretical virtue as these marks manifest in everyday theoretical models, then probably that armchair professional will not have the foundation on which to develop reliable content expertise in his or her field of specialization. This is analogous to how, prior to specialized athletic training, one needs to have superior eye–hand coordination (or raw speed, etc.) to become an elite performing professional athlete.

In the same way that having natural eye–hand coordination is not sufficient for developing elite tennis skill, having second-order knowledge of the marks of excellence for phase-one theoretical models is insufficient for conferring armchair content-reliability. For example, in the case of philosophical content expertise, one cannot reliably judge the accuracy of competing philosophical theories about mental content if one does not understand the competing theoretical models in the first place. Nor is one likely to understand what a theoretical model *ought* to explain if one does not grasp the network of problems made salient by contemporary and historical philosophical discussions. Thus phase-one knowledge of theoretical excellence puts one in position to recognize specialized content-success, and pending adequate exposure to phase-two epistemic resources, it allows for the receipt of a form of feedback that can underwrite the development of reliable content expertise.

This thesis opens up at least two explanations for why some (or many) armchair experts fail to develop content-reliability. The first is that the expert-in-training possessed neither the motivation nor the cognitive skill to cull from the ubiquitous phenomena of everyday theoretical model-building an implicit and fine-grained knowledge of the marks of theoretical excellence. This failure would occur prior to expert training. The second explanation is that the expert-in-training lacks competency with respect to phase-two resources. This might happen, for example, if one's graduate training had a myopic focus on a wrong-headed research programme.

Speaking generally, there is no guarantee that that these failures are not rampant among expert philosophers, and thus there is no guarantee that expert philosophers have even *in general* developed more content-reliability than novices. Perhaps undetected institutional forces or scholastic trends have rendered most schools of philosophical thought and training, and thus their graduating experts,

less able to recognize philosophical theoretical virtue than unschooled philosophical novices.

But we should not hope for guarantees about these matters – such are the epistemic realities and risks for investigative armchair domains. What we should hope for, and what I have offered, is a plausible and naturalistically accommodating account – one that comports with principles derived from the empirical study of expert performance and particularly the developmental importance of feedback – of how expert philosophers develop reliable content-expertise with respect to theoretical modelling.

5. Conclusion

There are neither empirical nor formal tests for whether the theoretical models of armchair experts deserve greater credence than those of armchair novices. This observation, together with empirical data indicating the importance of direct learning feedback for the development of reliable expert performance, encourages a sceptical view about the reliability of armchair expertise generally and expert philosophical theorizing specifically.

I argued that the type of sceptical challenge to philosophical theorizing best supported by the empirical literature on non-philosophical expert performance rests on implicit claims about success-identification and expert agreement. After interrogating these assumptions, I introduced a distinction between the distribution of reliable procedural expertise, on the one hand, and the distribution of reliable content expertise, on the other. From these efforts, I concluded that the empirical literature on expert performance warrants neither a general nor surprising scepticism about the quality of expert armchair investigation.

I next offered a positive characterization of learning feedback for theoretical modelling in armchair domains generally and the philosophical domain specifically. I sketched a metaphysics and epistemology for the reliable identification of theoretical virtues that, I argued, addresses an important set of developmental concerns that stems from the empirical investigation of expert performance.

My overall analysis supports the claim that, with respect to the professional goal of accurately theoretically modelling philosophical targets, expert philosophers are generally superior. If expert philosophers make intuitive judgements on the basis of philosophical theoretical models that they have psychologically internalized, then my analysis provides resources for a defence of expert philosophical intuition. I leave that topic to another occasion.

Acknowledgements

I thank two anonymous reviewers from this journal for their helpful comments and suggestions. Support for work on this article was generously provided by Bowling Green State University and the Office of the Provost in the form of Faculty Improvement Leave for Spring 2019.

References

- ANDOW, J. (2015) "Expecting Moral Philosophers to Be Reliable." *Dialectica* 69(2): 205–220.
- ARMSTRONG, J. S. (1978) *Long Range Forecasting: From Crystal Ball to Computer*. New York: Wiley.
- AYER, A. J. (1936) *Language Truth, and Logic*. London: Gollancz.
- BACH, T. (2012) "Analogical Cognition: Applications in Epistemology and the Philosophy of Mind and Language." *Philosophy Compass* 7(5): 348–360.
- BACH, T. (2019) "Why the Empirical Study of Non-philosophical Expertise Does Not Undermine the Status of Philosophical Expertise." *Erkenntnis*. <https://doi.org/10.1007/s10670-019-00141-2>.
- BLANCHARD, T., VASILYEVA, N., and LOMBROZO, T. (2018) "Stability, Breadth and Guidance." *Philosophical Studies* 175(9): 2263–2283.
- BOYD, R. (1979) "Metaphor and Theory Change: What Is 'Metaphor' a Metaphor for?" In A. Ortony (ed.), *Metaphor and Thought*, pp. 356–408. Cambridge: Cambridge University Press.
- CAMERER, C. F. and JOHNSON, E. J. (1991) "The Process-Performance Paradox in Expert Judgment: How Can Experts Know So Much and Predict So Badly?" In K. A. Ericsson and J. Smith (eds), *Toward a General Theory of Expertise: Prospects and Limits*, pp. 195–217. Cambridge: Cambridge University Press.
- CAPELEN, H. (2012) *Philosophy without Intuitions*. Oxford: Oxford University Press.
- CARROLL, J. S., WINER, R. L., COATES, D., GALEGHER, J., and ALIBRIO, J. J. (1982) "Evaluation, Diagnosis, and Prediction in Parole Decision Making." *Law and Society Review* 17: 199–228.
- CLARKE, S. (2013) "Intuitions as Evidence, Philosophical Expertise and the Developmental Challenge." *Philosophical Papers* 42(2): 175–207.
- CUMMINS, R. (1975) "Functional Analysis." *Journal of Philosophy* 72(20): 741–764.
- CUMMINS, R. (1998) "Reflection on Reflective Equilibrium." In M. DePaul and W. Ramsey (eds), *Rethinking Intuition*, pp. 113–128. Lanham, MD: Rowman and Littlefield.
- DAWES, R. M. (1971) "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making." *American Psychologist* 26: 180–188.
- DAWES, R. M. (1994) *House of Cards: Psychology and Psychotherapy Built on Myth*. New York: The Free Press.
- DEUTSCH, M. E. (2015) *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. Cambridge, MA: MIT Press.
- DEVITT, M. (2012) "Whither Experimental Semantics?" *Theoria* 72: 5–36.
- DEVITT, M. (2015) "Testing Theories of Reference." In J. Haukioja (ed.), *Advances in Experimental Philosophy of Language*, pp. 31–63. London: Bloomsbury Academic.

- DRETSKE, F. (1988) *Explaining Behavior*. Cambridge, MA: Bradford, MIT.
- ELLIS, B. (2009) *The Metaphysics of Scientific Realism*. Durham, NC: Acumen.
- GENDLER, T. S. (2007) "Philosophical Thought Experiments, Intuitions, and Cognitive Equilibrium." *Midwest Studies in Philosophy* XXXI(1): 68–89.
- GENTNER, D. (2003) "Why We're So Smart." In D. Gentner and S. Goldin-Meadow (eds), *Language in Mind: Advances in the Study of Language and Cognition*, pp. 195–236. Cambridge, MA: MIT Press.
- GOLDMAN, A. I. (1993) "Epistemic Folkways and Scientific Epistemology." *Philosophical Issues* 3: 271–285.
- GOPNIK, A. and WELLMAN, H. M. (2012) "Reconstructing Constructivism: Causal Models, Bayesian Learning Mechanisms, and the Theory Theory." *Psychological Bulletin* 138 (6): 1085–1108.
- GRIFFITHS, P. (1997) *What Emotions Really Are: The Problem of Psychological Categories*. Chicago, IL: University of Chicago Press.
- GRIFFITHS, P. (1999) "Squaring the Circle: Natural Kinds with Historical Essences." In R. A. Wilson (ed.), *Species: New Interdisciplinary Essays*, pp. 209–228. Cambridge, MA: MIT Press.
- HALES, S. D. (2006) *Relativism and the Foundations of Philosophy*. Cambridge, MA: MIT Press.
- HOGARTH, R. M. (2001) *Educating Intuition*. Chicago, IL: University of Chicago Press.
- KAHNEMAN, D. and KLEIN, G. (2009) "Conditions for Intuitive Expertise: A Failure to Disagree." *American Psychologist* 64(6): 515–526.
- KAUPPINEN, A. (2007) "The Rise and Fall of Experimental Philosophy." *Philosophical Explorations* 10: 95–118.
- KORNBLITH, H. (1995) *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*. Cambridge, MA: MIT Press.
- KORNBLITH, H. (1998) "The Role of Intuition in Philosophical Inquiry: An Account with No Unnatural Ingredients." In M. R. DePaul and W. M. Ramsey (eds), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, pp. 129–142. Lanham, MD: Rowman and Littlefield.
- LADYMAN, J. (2012) "Science, Metaphysics and Method." *Philosophical Studies* 160(1): 31–51.
- LOCKE, J. (1975) In P. H. Nidditch (ed.), *An Essay Concerning Human Understanding*. Oxford: Oxford University Press.
- LUDWIG, K. (2007) "The Epistemology of Thought Experiments: First-Person Approach vs. Third-Person Approach." *Midwest Studies in Philosophy* 31: 128–159.
- MARKMAN, A. and STILLWELL, C. (2001) "Role-Governed Categories." *Journal of Experimental and Theoretical Artificial Intelligence* 13: 329–358.
- MARTÍ, G. (2012) "Empirical Data and the Theory of Reference." In W. P. Kabasenche, M. O'Rourke and M. H. Slater (eds), *Reference and Referring. Topics in Contemporary Philosophy*, pp. 63–82. Cambridge, MA: MIT Press.
- MEEHL, P. E. (1954) *Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. Minneapolis, MN: University of Minnesota Press.
- MILLIKAN, R. G. (2010) "On Knowing the Meaning; with a Coda on Swampman." *Mind* 119 (473): 43–81.
- NOLAN, D. (2015) "The A Posteriori Armchair." *Australasian Journal of Philosophy* 93(2): 211–231.
- PAUL, L. A. (2012) "Metaphysics as Modeling: The Handmaiden's Tale." *Philosophical Studies* 160(1): 1–29.

- PUST, J. (2017) "Intuition." In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. CA: The Metaphysics Research Lab Center for the Study of Language and Information Stanford University Stanford. <https://plato.stanford.edu/archives/sum2017/entries/intuition/>.
- RYBERG, J. (2013) "Moral Intuitions and the Expertise Defence." *Analysis* 73(2): 3–9.
- SAATSI, J. (2017) "Explanation and Explanationism in Science and Metaphysics." In M. Slater and Z. Yudell (eds), *Metaphysics and the Philosophy of Science: New Essays*, pp. 165–192. Oxford: Oxford University Press.
- SCHECHESTER, J. (2010) "The Reliability Challenge and the Epistemology of Logic." *Philosophical Perspectives* 24(1): 437–464.
- SCHULZ, E., COKELY, E. T., and FELTZ, A. (2011) "Persistent Bias in Expert Judgments about Free Will and Moral Responsibility: A Test of the Expertise Defense." *Consciousness and Cognition* 20: 1722–1731.
- SCHWITZGEBEL, E. and 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.
- SHANTEAU, J. (1992) "Competence in Experts: The Role of Task Characteristics." *Organizational Behavior and Human Decision Processes* 53: 252–266.
- SIDER, T. (2009) "Ontological Realism." In D. Chalmers, D. Manley and R. Wasserman (eds), *Metametaphysics*, pp. 384–423. Oxford: Oxford University Press.
- SIDER, T. (2011) *Writing the Book of the World*. Oxford: Oxford University Press.
- SOBER, E. (2009) "Parsimony Arguments in Science and Philosophy – A Test Case for Naturalism_p." *Proceedings and Addresses of the American Philosophical Association* 83 (2): 117–155.
- SOSA, E. (2007) "Experimental Philosophy and Philosophical Intuition." *Philosophical Studies* 132(1): 99–107.
- STOLJAR, D. (2017) *"Philosophical Progress." Defence of a Reasonable Optimism*. Oxford: Oxford University Press.
- SWAIN, S., ALEXANDER, J., and WEINBERG, J. M. (2008) "The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp." *Philosophy and Phenomenological Research* 76(1): 138–155.
- SWOYER, C. (1999) "Explanation and Inference in Metaphysics." *Midwest Studies in Philosophy* 23: 100–131.
- TETLOCK, P. E. (2005) *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press: Princeton, NJ.
- THAGARD, P. (1989) "Explanatory Coherence." *Behavioral and Brain Sciences* 12(3): 435–467.
- TOBIA, K., BUCKWALTER, W., and STICH, S. (2013) "Moral Intuitions: Are Philosophers Experts?" *Philosophical Psychology* 26(5): 629–638.
- WEINBERG, J. M., NICHOLS, S., and STICH, S. (2001) "Normativity and Epistemic Intuition." *Philosophical Topics* 29: 429–460.
- WEINBERG, J. M., GONNERMAN, C., BUCKNER, C., and ALEXANDER, J. (2010) "Are Philosophers Expert Intuiters?" *Philosophical Psychology* 23(3): 331–355.
- WILLIAMSON, T. (2011) "Philosophical Expertise and the Burden of Proof." *Metaphilosophy* 42(3): 215–229.