Philosophical Psychology
Vol. 23, No. 3, June 2010, 331–355

Are philosophers expert intuiters?

Jonathan M. Weinberg, Chad Gonnerman, Cameron Buckner and Joshua Alexander

Recent experimental philosophy arguments have raised trouble for philosophers' reliance on armchair intuitions. One popular line of response has been the expertise defense: philosophers are highly-trained experts, whereas the subjects in the experimental philosophy studies have generally been ordinary undergraduates, and so there's no reason to think philosophers will make the same mistakes. But this deploys a substantive empirical claim, that philosophers' training indeed inculcates sufficient protection from such mistakes. We canvass the psychological literature on expertise, which indicates that people are not generally very good at reckoning who will develop expertise under what circumstances. We consider three promising hypotheses concerning what philosophical expertise might consist in: (i) better conceptual schemata; (ii) mastery of entrenched theories; and (iii) general practical know-how with the entertaining of hypotheticals. On inspection, none seem to provide us with good reason to endorse this key empirical premise of the expertise defense.

Keywords: Armchair Philosophy; Conceptual Schemata; Configural Rules; Experimental Philosophy; Expertise; Intuitions; Restrictionist Challenge

1. Setting the Stage: The Restrictionist Challenge and the Expertise Defense

1.1. The Restrictionist Challenge

A number of experimental philosophers in recent years (e.g., Machery, Mallon, Nichols, & Stich, 2004; Swain, Alexander, & Weinberg, 2008; Weinberg, Nichols, & Stich, 2001; see also Sinnott-Armstrong, 2008) have begun to challenge analytic philosophy's longstanding practice of deploying armchair intuitive judgments about cases. This has sometimes been called the *restrictionist challenge* (Alexander & Weinberg, 2007), as these philosophers contend that some substantial degree of

Jonathan M. Weinberg is Associate Professor of Philosophy at Indiana University, Bloomington, USA.
Chad Gonnerman is a Ph.D. candidate in Philosophy at Indiana University, Bloomington, USA.
Cameron Buckner is a Ph.D. candidate in Philosophy at Indiana University, Bloomington, USA.
Joshua Alexander is Assistant Professor of Philosophy at Siena College, Loudonville, New York, USA.
Correspondence to: Jonathan M. Weinberg, Indiana University – Philosophy, Sycamore Hall 026, Bloomington, IN 47405, United States. Email: jmweinbe@indiana.edu

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/10/030331-25 © 2010 Taylor & Francis
DOI: 10.1080/09515089.2010.490944

restriction on that practice may be warranted; we will follow that usage here. Our goal in this paper is to set out one understanding of the dialectical structure of the restrictionist challenge, and critically evaluate the prospects for one popular pro-armchair response to it: that philosophers are *experts* in a way that should render their intuitive judgments unimpeached by the experimental philosophers' findings.

Here is one way of construing the restrictionist challenge, in terms of three main moving pieces. First, there are the experimental results themselves, which at this point in time overwhelmingly concern 'ordinary' subjects (typically, but not entirely, university undergraduates). In particular, there is the claim that the studies in question reveal a worrisome pattern of responses in those subjects, such as ethnic variation or sensitivity to the order of presentation of the cases. Second, there is a metaphilosophical claim that the relevant philosophical facts do not pattern in the same way; there is a misalignment between the contours of the philosophical facts and the contours of philosophical judgments (or, at least, of those judgments as studied by the experimental philosophers). Third, there is an ampliative inference from the patterns disclosed concerning the ordinary subjects to the predicted occurrence of those patterns in professional philosophers. Putting all three pieces together, the armchair practice is thus challenged: we have reason to think that it involves deploying a source of putative evidence that is sensitive to non-truth-tracking factors.

We will use the term 'cathedrist' for those looking to defend the trustworthiness of the traditional armchair method in the face of the restrictionist challenge. There are at least three different main avenues of cathedrist response, corresponding to each of the three parts of the restrictionist challenge.[1] They can attack the first, experimental piece, by critiquing the studies themselves (e.g., by proposing confounds; see Cullen, forthcoming; Sytsma & Livengood, forthcoming). They can attack the second, metaphilosophical piece, by suggesting philosophical appropriateness of variation (e.g., by going relativistic in the face of diversity; see Goldman, 2007; Jackson, 1998; Lycan, 2006; Sosa, 2009) or contextualist in the face of instability (see discussion at the end of Swain et al., 2008). And they can attack the third piece, by demonstrating some relevant difference between the psychology of philosophers when engaged in their professional activities, and the psychology of undergraduates in the context of doing surveys, which would thereby defeat the restrictionist's ampliative inference. For example, it might be that philosophers are appealing to a different kind of psychological state than that deployed by the folk, in answering the survey questions; alternatively, it might be that philosophers' social epistemic practices would protect them from the unwanted variation observed in the folk. The remainder of this paper will be concerned with one particular version of this third sort of defense: the claim that philosophers' special training will sufficiently shield them from problems that may afflict folk intuitions, which we will call the *expertise defense*.

1.2. The Expertise Defense

According to the expertise defense, ''we should acknowledge that not all intuitions are created equal .... For example, the physical intuitions of professional scientists

are much more trustworthy than those of undergraduates or random persons in a bus station'' (Hales, 2006, p. 171). The mathematical intuitions of professional mathematicians are similarly more trustworthy than those of the folk. So, too, the philosophical intuitions of professional philosophers are more trustworthy than the intuitions of the subjects studied by experimental philosophers. In light of this, the practice of appealing to *philosophical* intuitions about hypothetical cases, properly construed, should be the practice of appealing to *philosophers'* intuitions about hypothetical cases. And so studies conducted on the intuitions of untutored folk can provide no evidence against the practice of appealing to philosophical intuitions as evidence.

It borders on the trivial to claim that philosophers' training makes them at least somewhat better than the folk, at least at some philosophically-relevant tasks. But while such a generic claim to expertise likely merits a strong default standing in these discussions, it is important to see that the expertise defense turns on a much more substantial empirical hypothesis. *Some improvement or other* is one thing, but what the purveyors of the expertise defense require is that philosophers' intuitions are *sufficiently less susceptible to the kinds of unreliability that seem to afflict the folk intuitions studied by experimental philosophers*. Only if this much more specific—and hence much stronger—empirical hypothesis is true will the expertise defense provide an adequate response to the restrictionist challenge. If philosophers' intuitions on the whole aren't sufficiently immune from the sorts of distortions that folk intuitions seem to be prey to, then—even if philosopher's intuitions are in *other* ways epistemically superior to the folk's—the cathedrist will not in fact have the makings of a successful reply to the challenge.

The specificity of the expertise claim that is needed has not, we think, been generally noted by cathedrists so far, as evidenced at least in part by the generality of the sorts of defenses they have offered. While cathedrists have on the whole recognized that some sort of argument is needed on their behalf, they have tended to offer only analogies on behalf of such claims to expertise. Hales (2006), for example, notes that philosophers and physicists alike have lots of education and training. Williamson (2005, 2007) compares the training of philosophers to that of lawyers, and Ludwig (2007) does likewise with mathematicians. The common thought behind these appeals is that we typically give much greater credence to the intuitions of these experts in their respective domains than we do the *hoi polloi*, so we should do the same for philosophers in philosophical matters. But such analogies at best secure the generic, albeit highly likely, claim to expertise that we have just noted is insufficient to the cathedrist's argumentative needs.

Why have so many very smart philosophers not accurately observed what sort of premise they really need here? Our hypothesis is that they all share a sort of *folk theory of expertise*. The rough aphorisms of this folk theory include: (i) sufficient background, training, or experience is by and large all it takes to get better at any given activity; and (ii) expertise at one aspect of an activity is closely correlated with expertise in other aspects of that activity. Since philosophers do indeed spend a lot of time philosophizing, and since many aspects of philosophy do seem to show some

sort of return on effortful investment (philosophical training does typically bring a mastery of relevant literatures both contemporary and historical, and even specific technical skills such as argument evaluation and construction), this folk theory of expertise can make it seem as though *of course* we should expect philosophers to be better (and better in the right sorts of ways) at conducting thought experiments than the 'uninitiated'.

However, the real psychological facts about expertise turn out to be much more complicated, and more interesting, than this folk theory would suggest. First of all, not just any experience or training will result in expertise, no matter how prolonged or effortful. For example, Shanteau (1992) surveys a vast swath of the literature on the development of expertise, and finds tremendous diversity in the development of expertise according to the characteristics of the task and the learning environment. Some areas, such as meteorology and chess, have proved conducive to acquiring expertise; others, such as psychiatry, stock brockerage, and polygraph testing, have tended not to produce real expertise, even for those with years of experience and training (for general reviews on the insufficiency of training and experience alone in producing genuine expertise, see Camerer & Johnson, 1991; Ericsson, 2006, pp. 686, 691; Ericsson & Lehman, 1996, pp. 276–77; Feltovich, Prietula, & Ericsson, 2006, p. 60; Garb, 1989; Shanteau & Stewart, 1991). (See section 2.2 below for further philosophy-specific discussion.) Moreover, even when some expertise *does* develop, it does not follow that all of the problems that we discussed earlier will go away. For example, Damisch, Mussweiler, and Plessner (2006) found that even highly trained Olympic gymnastics judges were remarkably susceptible to order effects, depending on the perceived similarity between the performance that they were judging and the immediately preceding performance. (See also, e.g., Brown, 2009, for concerns about order effects in professional auditors.) Neither is expertise a cure-all for cultural biases, which can stubbornly persist in the face of the highest levels of expertise even in well-established scientific disciplines. For example, de Waal (2003) reviews a century-long rift between prominent Western and Japanese primatologists concerning the threat of anthropomorphism when interpreting certain behaviors as evidence of chimpanzee kinship groups and culture. While this particular rift, in de Waal's opinion, is now closed, it closed not because the intuitive assessments of expert primatologists were immune to cultural biases, but rather because the Western primatologists were confronted by large amounts of data they were unable to explain without conceding the major points of the dispute. Furthermore, de Waal cautions that such cultural biases are still alive and well in other areas of biology; Dutch and Japanese studies on African cichlids agreed to "all the data" but disagreed as to whether a marked increase in speciation should be explained in terms of "competition and exploitation" or "complementary roles within the ecosystem" (de Waal, 2003, p. 293). And de Waal claims that such cultural variation in scientific judgments calls for more active cross-cultural exchange, as "each culture is too wrapped up in its own relation with nature to step back and see it as it is" (2003, p. 298). Regardless of what specific measures are most appropriate in biology, the key morals for our purposes here are that, first, even the highly trained expertise of

biologists does not shield them from cultural variation in epistemically important judgments, and second, when such variation is found, it is something that requires an active response, and not just a shrugging of the shoulders while assuming that the existence of some degree of expert training in one's community of inquiry renders otiose any such worries. (See also Page, 2007, and Zamzow & Nichols, forthcoming, on the positive epistemic values of having a diverse community of inquirers, under the right conditions.)

Neither can cathedrists take any comfort from the ways in which philosophers clearly manifest expertise, such as our mastery of formal techniques or the corpora of our philosophical forebears. It is "one of the most enduring findings in the study of expertise" that it typically develops highly narrowly and task-specifically, and that there is "little transfer from high-level proficiency in one domain to proficiency in other domains—even when the domains seem, intuitively, very similar" (Feltovich et al., 2006, p. 47). Experts in one board game will not automatically have any particular expertise in another, similar game; surgical expertise turns out to be surprisingly specific to individual surgical tasks. (See also Norman, Eva, Brooks, & Hamstra, 2006.) So philosophers' possession of such demonstrable skills as, say, the close analysis of texts, or the critical assessment of arguments, or the deployment of the tools of formal logic, does little to nothing to raise the probability that they possess any correspondingly improved level of performance at conducting thought experiments.

Even in linguistics—perhaps as closely related to analytic philosophy as any scientific field could be—the status of "expert" intuitions is not just contested, but furthermore considered a matter for sustained empirical investigation (Labov, 1975, 1996; Schuetze, 1996; Wasow & Arnold, 2005). Even those who look to defend the viability of "armchair linguists" do so from a thoroughly empirically-oriented perspective, treating it without argument as a question to be addressed with substantial reliance on scientific investigation (Philips, forthcoming).[2] Surely if the status of expertise claims in linguistics requires empirical backing, it would be surprising if parallel claims in philosophy somehow did not.

We conclude that philosophers at this time *cannot take it for granted* that they are experts, in the relevant sense here of being sufficiently shielded from the sorts of effects that provide the raw materials for the restrictionist challenge. But that is of course not to say that we can assume that they are *not* experts in such a sense, either! Ideally, what would follow at this point would take the form of an empirical study, like the ones that experimental philosophers have conducted on folk intuitions. We don't have any such study to offer at this time (for that matter, neither do cathedrists). Although we hope that such studies may someday be undertaken, in advance of—and as relevant guidance for—any such studies, we offer a set of hypotheses as to what a psychologically respectable basis for philosophical expertise *might* consist in. The psychological literature on expertise and expert performance offers a sizeable array of hypotheses as to the basis of the improved performance of different expert populations, and we have here extracted the three that strike us as the most plausible, and which moreover jibe at least somewhat with some of the

hand-waving appeals to expertise already extant in the cathedrist literature. To be clear, we do not consider this an exhaustive account of all possible hypotheses of philosophical expertise, and encourage others to explore this rich literature for other ideas. We make no claims to be experts on expertise, but we do think we have learned a few things of clear relevance to current metaphilosophical disputations.

The three hypotheses that we will consider are that philosophers have superior *conceptual schemata* to the folk; that they deploy more sophisticated *theories* than the folk; and that they possess a more finely-tuned set of *cognitive skills* than the folk. Each of these hypotheses squares well with at least one substantial corner of the literature on expertise, and can be seen to offer at least a *prima facie* sketch of a way that the cathedrist could resist the restrictionist's arguments. On closer inspection, however, each hypothesis will be seen to be more or less problematic given the current dialectical situation. If we understand each hypothesis as taking the form 'Philosophers are expert intuiters in philosophical matters because they have psychological characteristic C', then there are at least four general problems that an appeal to C might face in the current dialectic. First, it might be that the appeal is to something that philosophers (probably) lack. Second, it might be at best an open, empirical question as to whether C would sufficiently inoculate against sensitivity to the factors that drive the restrictionist challenge. Third, C might be something that would be expected to introduce new forms of systematic error, leading the cathedrist at best to move out of the frying pan and into a different frying pan. Fourth, appealing to C might fail to address the challenge in question, by presupposing resources that would require that the challenge has already been met, and thus could do no work of its own in meeting that challenge. In what follows, we will see that appeals to conceptual schemata and sophisticated theories suffer from a mix of these four problems while the appeal to cognitive skills is as of yet too underdeveloped for us to assess its success.

2. Expertise as Better Conceptual Schemata?

2.1. The Hypothesis

One possibility here is that philosophical experts possess specialized conceptual schemata which they bring to bear in the evaluation of thought experiments. Because 'concept' is a fraught term both in philosophy and in psychology,[3] we will use 'conceptual schemata' to refer to whatever informational structures there are in the head that ultimately underwrite judgments of the application of terms, over and above one's explicit theories.

Such a model provides one way to work out the not-uncommon suggestion that philosophical expertise consists in a kind of special competence with concepts of philosophical interest. Singer (1982, p. 9), for example, argued that the professional ethicist's coursework endows them with a clearer understanding of moral concepts which helps them avoid settling moral issues in a "muddled way" (see also his 1974). More recently, Ludwig (2007, pp. 159–160) has endorsed the idea that professional

philosophers acquire a ''greater sensitivity to the structure'' of philosophical concepts, which allows them to ''focus on the right features'' of thought experiments and thereby avoid restrictionist manipulations.

Importantly, this kind of model is one of the most prominent in the psychological literature on expertise. Studies on experts in many domains have shown that they do in fact tend to develop highly-specialized conceptual schemata. To name just a few, studies of experienced medical diagnosticians, chess grandmasters, physics professors, and auditors (see Anzai, 1991; Camerer & Johnson, 1991; Charness, 1991; Norman et al., 2006; for a general view, see Ericsson & Lehman, 1996) show that they do develop specialized representations for evaluating particular cases (e.g., ''best next move'' for chess players [de Groot, 1946/1978] or ''acceleration'' for physicists [Reif & Allen, 1992]). A cathedrist might survey such studies and suggest that expert philosophers have also developed appropriately specialized conceptual schemata. To explore this idea, we review two different (but related) ways to flesh it out in psychological detail.[4]

Camerer and Johnson (1991) propose one kind of specialized decision procedure often acquired by experts, which substantially drives their intuitive judgments: the ''configural rule.'' As models of expert decision-making, configural rules are to be contrasted with actuarial or regression models. Actuarial or regression models predict the likelihood of a target attribute based on a weighted summation of all possibly-relevant cues. Configural rules, on the other hand, combine small, contextually-determined subsets of features to infer the presence or absence of the target attribute, and the impact of any one feature is a product of the presence of the others. For example, a regression model might predict the likelihood of pre-diabetes using a complex summation of scores (with context-invariant weights) for family history, age, ethnicity, socioeconomic status, blood pressure, pre- and post-fasting blood glucose levels, patient lifestyle and symptom reports, and so on. Expert diagnosticians, on the other hand, will tend to focus on small subsets of cues presumed highly diagnostic for a given context; for example, the (possibly tacit) rule might be ''if a patient is obese and over 45, fasting blood glucose over 110 mg/dL means pre-diabetic''—whereas ''if a patient is normal weight and under 30'', a different combination of cues and their scores might be used (Camerer & Johnson, 1991, pp. 204–205). The authors suggest (pp. 207–208) that experts end up encoding such structures because they are computationally less expensive and easier to learn, they are amenable to causal interpretation, and they emerge naturally from the attempt to explain past cases, yet can sometimes match or occasionally even exceed the performance of more laborious forms of reasoning. Two common classes of configural rules that ''experts are quick to develop'' can be seen to closely resemble philosophers' own preferred structures of necessary and/or sufficient conditions: conjunctive (''hire Hope for the faculty if she has glowing letters of recommendation, good grades, and an interesting thesis'') and disjunctive (''draft Michael for the basketball team if he can play guard or forward or center extremely well''). If philosophers' intuitions are driven by configural rules, and if their configurations—unlike those of the folk—are sensitive only to philosophically-relevant

features, then philosophers' judgments might be resistant to restrictionist instabilities.

Another route to conceptual expertise might lie in the acquisition of competence with specialized representations, which Boshuizen and Schmidt (1992) call "higher-order concepts,"[5] that summarize and organize cues deemed relevant to the solution of common kinds of task in that domain. For example, they argue that clinical physicians organize their diagnostic experience through a process of "knowledge encapsulation."[6] Upon reviewing think-aloud protocols (in which subjects are asked to verbalize their trains of thought during decision-making) administered to physicians at varying stages of their career, they found that the number of "lower-level" biomedical propositions mentioned (operationalized as indicating an anatomical location or bodily process) decreased proportionally and monotonically with increase in expertise. At the highest levels of expertise, explicit biomedical propositions had dropped below 10% of those mentioned (down from above 55% for beginning medical students). Instead, their experts seemed to traffic in higher-order clinical concepts difficult to operationalize in biomedical terms, such as "excessive drinker" and "atypical inflammation reaction." Knowledge encapsulation was also evidenced, they argue, by experts' tendencies to verbalize direct links between higher-order concepts without appealing to intermediary biomedical steps, such as making the jump from "drug use" to "endiocarditis" without stopping at "contaminated needles" and "bacteria-induced sepsis" (Boshuizen & Schmidt, 1992, p. 169). Because their experts were still capable of producing biomedical knowledge to justify their diagnoses in post-hoc probes, Boishuizen and Schmidt conclude that biomedical knowledge was not lost, but has rather been encapsulated, through years of experience and practice, under the higher-order concepts appearing in the protocols.[7] To return to the philosophical domain, we might thus suppose that philosophers, after years of experience dealing with problems of a certain type, acquire higher-order concepts like "deviant causal chain" or "Gettiered justification" which encapsulate lower-level logical and conceptual features of cases. Furthermore, we might suppose that it is on the basis of competence with such higher-level concepts that, for example, epistemologists are able to efficiently pick out just the epistemologically-relevant features of hypothetical cases.

Both hypotheses concerning the specialized conceptual schemata of philosophers thus suppose them resistant to restrictionist manipulations like framing effects because they will not be distracted by the philosophically-irrelevant features which vary across the experimental conditions. The problem with the folk judgments, then, would be that they are driven by conceptual schemata that include factors that do not track the philosophical truths, and these factors provide the hooks that can be manipulated in experimental contexts (and elsewhere) to lead their judgments astray.

2.2. The Critique

We have extracted from the empirical literature two suggestions as to how philosophers might possess better conceptual schemata: their conceptual schemata

might consist of configural rules or "higher-order concepts" that encode just the philosophically relevant features of cases. Our first reaction to these hypotheses is something along the lines of: "boy, they sound like awfully interesting empirical hypotheses, and it sure would be nice if someone would test them." Putting that aside for now, though, we are concerned that even if philosophers showed hallmarks of configural rules or knowledge encapsulation, this still might not provide a sufficient response to the restrictionist challenge at this time.

For let us thus grant, for the sake of argument, that experts have some form of specialized conceptual schemata. Although we've no evidence for it, we will also grant, for the sake of argument, that the philosophical intuitions of experts will stabilize on some set of verdicts in such a way that they are able to give the same answer across conditions in the standard restrictionist experiments. We'll grant this because we don't think it makes much difference to the dialectic. The problem is that the cathedrist's work has only just begun once they show that philosophers' judgments are consistent across conditions of restrictionist experiments. One easy route to stabilization, for example, would be to "lock in" a sensitivity to some philosophically-irrelevant factors, such that one's intuitions are internally coherent but have no more (and perhaps even *less*) claim to the truth than the intuitions of novices. For example, philosophers may "lock in" to their initial assessment of a case as primed in one context and render later judgments consistent with the earlier one. This sort of outcome would not be psychologically surprising: for example, Ariely, Lowensterin, and Prelec (2003) report a number of experiments demonstrating that subjects spontaneously impose such "coherent arbitrariness" on their judgments in a variety of tasks; and persons more inclined towards reflection have been demonstrated primacy effects in tasks involving the evaluation of evidence in highly structured argumentative settings (Petty, Tormala, Hawkins, & Wegener, 2001), not too dissimilar from the way that philosophers typically encounter arguments. While such judgments would appear to be stable in restrictionist thought experiments, there would be no more—and perhaps even *less*—reason to trust these judgments as opposed to those of the novices. To return to the scheme outlined at the end of the previous section, it may be that philosophers' intuitive assessments are driven by configural rules or higher-order concepts which render their judgments stable across conditions, but without actually rendering their judgments sufficiently stable against the influence of philosophically-irrelevant factors. (Note that even settling very stably on one judgment over time is consistent with still suffering from unwanted sensitivity in judgment, if the matter of *which* end-point one has settled upon is too much influenced by inappropriate factors. For example, the rule "regarding any philosophical question, believe the first thing you are told, and then do not change your mind no matter what" might result in highly stable beliefs across time, but it clearly suffers from a radical sensitivity to who your initial teachers may be.)

A fascinating incongruity in expertise research, remarked upon by both Camerer and Johnson (1991) and Shanteau (1992), is found between the studies of expertise conducted in cognitive psychology and those conducted in the psychology of judgment and decision-making. While the cognitive studies have generally found

experts to be well-tuned to their tasks, the literature on decision-making has surprisingly often found that the judgments of experts, while highly-specialized, are no more reliable than those of novices. A primary focus of both Camerer and Johnson and Shanteau was to explain how both of these conclusions can be simultaneously true. Camerer and Johnson argue that while experts do exhibit some epistemic virtues not typically found in novices—such as thinking in terms of subtle distinctions and novel categories, and possessing an increased domain-specific memory and better skill at generating new hypotheses and measuring variables—their judgments *also* still often fall prey to many of the same errors as novices. For example, the configural rules of experts often invoke the same illusory correlations as the decision procedures of novices.

Indeed, another sort of problem to be considered with the "better conceptual schemata" approach is that such schemata run a very real risk of introducing *new* systematic sources of error. For example, the expert's extensive experience will make them even more likely to induce a configural rule from a small sample and then overgeneralize (Camerer & Johnson, 1991). For example, Lewandowsky and Thomas (forthcoming) discuss research of how master level (though, interestingly, not grand master) chess players were susceptible to an order effect on a set of chess puzzles involving finding the best route to checkmate. It seems that such players can be primed with a set of cases involving smothered mate positions, leading them to over-notice such positions in later puzzles. Even though the *prima facie* promise of "better conceptual schemata" is that such schemata might exclude factors to which the folk are inappropriately sensitive, nonetheless to have configural rules and higher-order concepts is at the same time to have another piece of cognitive machinery upon which priming effects and the like can get a hold. In short, there is a threat that any "stability" that does get afforded by switching to configural rules or higher-order concepts may end up being bought only at the cost of introducing yet a new source of error—a new source of grist for the restrictionist's mill.

This is not to say, of course, that experts are always doomed; the right kind of training, on the right kinds of tasks, can very frequently enable experts to overcome their cognitive limitations (at least, on those tasks). One of the most robust consensus findings of the study of expertise is that expert judgments can only become more reliable where experts are readily confronted with clear, reliable feedback on which to train.[8] Configural rules are acquired through experience, and are accepted, shaped, and rejected through the apprehension of repeated or salient successes and failures. The tuning here, however, depends upon the ability of the expert-in-training to focus intently on practice, break the task down into components, and correctly diagnose (perhaps tacitly) the causes of success or failure. Where feedback is unclear, infrequent, or absent, inaccurate configural rules are likely to become progressively more entrenched; and when experts are left to their own devices to find negative feedback, confirmation biases will often allow inaccurate rules to persist even when ample disconfirming instances are available (Camerer & Johnson, 1991, p. 210).

So: one might expect configural rules and higher-order concepts *of some sort or other* to develop in the cognition of anyone seeking expertise in any domain; what the

cathedrist needs in the minds of philosophers, however, are *epistemically virtuous* concepts or rules. Given that the experts' specialized conceptual schemata often introduce new forms of error in many domains, and especially that experts in other domains often seem sensitive to the same sorts of order effects as novices (albeit in a different way), the cathedrist offering the expertise defense needs to provide substantive reasons to believe that the expertise defense's reliance on specialized conceptual schemata would not have to pay such a self-destructive bill. The natural response to these worries would be to argue that the training of philosophers is rich with the kind of clear, reliable feedback which would allow them to hew their conceptual schemata more closely to the philosophical truth. But what, specifically, might that feedback be?

There is a live possibility here that philosophers' intuitions about cases do not receive anything like the kind of substantial feedback required for such virtuous tuning. Worse still, various of the psychological mechanisms that can plague human cognition in general—overgeneralization, overconfidence, cognitive dissonance, attribution error, belief bias, belief perseverance, and so on—might very well lead philosophers to believe that we have attuned intuitions, even if in reality in many places we have simply been systematically reaffirming early impressions and incorrectly attributing our professional successes (e.g., in debate and publication) to their validity. We should certainly expect that *some* degree of non-virtuous tuning occurs in any discipline, and the degree to which it persists will be inversely proportional to the amount of reliable feedback received and the amount of motivation to correctly tune. Philosophers would be far from alone in this regard. Shanteau (1992) lists such highly-regarded professionals as stock brokers, psychiatrists, and intelligence analysts as ones in which such false expertise has been found.

Another suggestion might be that philosophers train their intuitions against other, already-certified expert intuitions.[9] But this appears to be a non-starter, since it just invites an explanatory regress: how did the purveyors of *those* intuitions develop their expertise? And whatever it is, why is it not *that* that we are appealing to in order to meet the restrictionist challenge, instead of attempting an expertise defense? One might appeal to something like the wisdom of the philosophical ages here, the slow accretion of certified cases over generations. But analytic philosophy's history does not seem to have a large amount to offer in such terms.

It is important to keep the relevant contrast domains firmly in mind here, for they are what will provide the meterstick by which we can evaluate just what works as the right kind of feedback, and in what needed amounts, in order to produce effective training of real expertise. The fields in which competent experts routinely develop are those like meteorology, livestock judging, and chess. In such areas, experts are confronted with a truly vast array of cases, with clear verdicts swiftly realized across a wide range of degrees of complexity or difficulty. Philosophy rarely if ever (outside its formal subareas) provides the same ample degree of well-established cases to provide the requisite training regimen. Feedback on errors in thought experimenting is more limited, usually only coming when we offer our arguments publicly in some way; such feedback is much more unclear and ambiguous even when it comes, as we will

often take those who do not share our intuitions simply to be themselves in error, and we can easily, even without meaning to, surround ourselves with like-minded intuiters; and when a correction signal final does materialize, it can come with time-lags measured more in years than in the minutes or days more typical of the fields with more successful results of training. (Just ask yourself how often you change your mind about a thought experiment, and how long you have typically taken the case under consideration before finally undergoing such a change.) Most of all, the size of a typical philosophical "training set" seems to be orders of magnitude smaller than those in areas with successful inculcation of this sort of expertise. There may seem to be lots of, say, trolley cases, or fake barn cases, or teletransporter cases. But compared to, say, the number of times a chess player will have practiced a given opening en route to mastery, it seems that we are not even dealing in the same order of magnitude. This stark divergence between philosophical training regimens and those of areas in which expertise is well-documented is the sort of thing that attention to the relevant empirical literature can highlight, but which is obscured by unsubstantiated analogies to other disciplines.

In evaluating such claims, we would stress that it is important not to backslide into relying on an armchair sense of how much or what sorts of training is or is not sufficient. It would be very easy, very tempting, for philosophers to look at the rough handful of places where we do get some sort of feedback and say to themselves, self-reassuringly, "yes, surely that's enough." But evaluating whether we should expect that philosophers have enough of it and of the right sort by the lights of the expertise literature requires examining that literature, and getting some sense of what the differences really are between the endeavors that do and do not regularly inculcate expert levels of competence in their practitioners. After all, 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, and at a minimum, these considerations indicate that proponents of the expertise defense need to offer real, substantive scientific evidence that this is not so.

Now, there are some important philosophical exceptions that prove the rule here. There truly are some key distinctions that philosophers have wrung from their theorizing, and because we have ways of explicitly inculcating sensitivity to these distinctions in our students, we can train them in ways that one should expect to be epistemically successful. Use/mention, epistemological/metaphysical, semantic/pragmatic—we certainly think that philosophical training *can* inculcate expert conceptual schemata structured in terms of *these* dimensions. But the explicitness and clarity of these distinctions,[10] and of our tools for dealing with them,[11] stands in very sharp contrast to the complete inarticulateness of the . . . well, *whatever* it would be, that is supposed to help trained philosophers to categorize Gettier cases as non-knowledge, and to be insensitive to cultural biases and framing effects. This is why it is important

that the experimental findings in the first piece of the restrictionist arguments not be explicable in terms of any such well-entrenched philosophical distinctions. Where such deflationary explanations of the experimental findings are successful, then a version of the expertise defense would be in good order: philosophers' education in these distinctions would predict precisely the sort of improvement over the untrained folk that is needed to block that third piece in the restrictionist's challenge. Yet such deflationary explanations have not yet proved generally successful. They simply have not managed to fit the patterns found in the data. (We will note two possible cases of deflationary counter-explanations in a footnote.)[12]

Another suggestion that might be made here is that philosophers may train on consonance with favored philosophical theories as their source of feedback. For example, perhaps a student comes to accept epistemological reliabilism, and over time ceases to have the "standard" Truetemp intuitions because they are judged inconsistent with her reliabilism. In this case, whether the intuition has been properly attuned or led astray depends upon whether reliabilism is true. Or, suppose that a philosopher answers that a thought experiment does not present a case of knowledge because the justification offered in the example is Gettiered. This philosopher is not thus reporting about an intuition concerning their *knowledge* concept directly, but more basically a judgment deploying the concept "Gettiered justification." Whether *that* judgment provides a trustworthy assessment of the case as a case of non-knowledge thus depends on whether the right theory of knowledge prohibits Gettiered beliefs from counting as knowledge. In both cases, we would again have a regress, in this case one that percolates through one extra step of theory-formation: expert intuitions are trained upon entrenched theory, which is itself entrenched by its compliance with . . . expert intuitions, about which we are still stuck having to ask how they were trained up in the first place. If the cathedrists already have some *other* way to tell which philosophical theories were true, then they can simply appeal to this source, and everyone can go home happy!

It seems, then, that both the suggestions that expert intuitions are trained against consonance with the intuitions of other experts, and that they are trained against favored philosophical theory, face persistent problems of explanatory regress. They can only begin to answer the restrictionist challenge if they help themselves to resources that would already suffice to answer the challenge. Such resources may yet be secured, and we offer no argument here that they cannot. Our point is the limited one that, should such resources be available, then *they* would be what is doing all the argumentative heavy-lifting, and the cathedrist would thus be offering something other than the kind of expertise defense that we are trying to problematize here.

3. Expertise as Better Theories?

3.1. The Hypothesis

Another story that we might tell about the source of expertise in philosophical intuiting begins with a rather uncontroversial thought. Time spent developing at least

a moderate understanding of a discipline's rich and well-established theories or principles is often repaid with more trustworthy domain-related judgments. That is why, for example, we may reasonably predict that those who had recently completed a course in physics will in at least some domain-related tasks outperform those with no formal instruction in physics. One such task, we might suggest (especially if we note that Newton's laws are taught in most, if not all, introductory courses), is to specify the trajectory that a ball will take when, in the absence of any external forces, it exits a curved, metal tube that is lying on its side. And, in this case, the prediction appears to bear out. McCloskey, Caramazza, and Green (1980) report that subjects who had recently completed one or more courses in physics were more likely to accurately predict that the ball will move in a straight line. Those with no formal instruction in physics were more likely to predict, inaccurately, that the ball will take a curved trajectory. It is plausible that what made the difference here is that, in carrying out the task, the former subjects, unlike the latter, applied some explicit bit of theoretical knowledge, specifically of Newton's law of inertia. (See also Norman et al., 2006, and Schmidt, Norman, & Boshuizen, 1990, for suggestions that this may apply in medical domain as well.)

The suggestion for philosophy, then, is that 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.

One worry for this line of thought is that the deployment of explicit theory does not itself count as relevantly *intuitive*, as appeals to explicitly-believed theory may not meet some necessary conditions on intuitionhood (e.g., Lynch, 2006). 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. There's no need for us to fight that fight here. (And, importantly, we will not need to take a stance on this question in the critical discussion that follows.) Still, to satisfy those who are disinclined to regard such judgments as intuitions, we should mention that with a slight tweak we can make the theory-driven model applicable to unconsciously generated judgments as well. Perhaps the source of expertise in philosophical intuiting consists primarily in the mastery of philosophy's well-confirmed theories or principles, a mastery which is so thorough that the relevant material has become internalized, or incorporated within the philosopher's ''adaptive unconscious'' to borrow a term from Wilson (2002). (See also Anderson, 1982, on *proceduralization* or Karmiloff-Smith, 1992, on *modularization.*) The thought, then, is that the philosopher's mastery of the relevant material exercises its influence from this part of her cognitive apparatus. More fully, it is that when confronting a philosophical thought experiment, the adaptive unconscious quickly and quietly applies the internalized material to the targeted content in imagination. And, since the application of well-confirmed theories or principles tends to increase the reliability of domain-related judgments, intuitive or otherwise, the proponent of the expertise defense might suggest that we can expect

professional philosophers to have an advantage over, say, the typical undergraduate at the evaluation stage of the intuition-generating process.

3.2. The Critique

If the ''better theory'' hypothesis is to help meet the restrictionist challenge, then its proponent must claim both (i) that philosophers have such a stock of sufficiently entrenched theories and (ii) that their deployment of these theories would shield their intuitions against the irrelevant factors. For example, the proponent would claim that if we were to give philosophers the survey from Swain et al. (2008), their intuitions would fail to display an order effect as long as they are sure to apply the relevant theories to each case in the ordering. But we worry that, simply put, philosophy does not yet have in its hands the kinds of theories needed to do that kind of work. The problem here is that there is just no good candidate for a rich, well-established body of theory the mastery of which can constitute a relevant source of expertise for philosophers.

This is not to say that philosophers do not have *anything* that would count as such mastery. Historians of specific periods or figures surely have such, and those who work on the ''philosophy of X'' for various values of X will draw upon a large body of knowledge, such as art-critical practices, mathematics, or scientific psychology. And note that the restrictionist is *not* aiming to challenge the judgments of such philosophers about the contents of the domains they are working on.[13] Yet in the areas of philosophy in which appeals to intuition about cases are still central, such as epistemology and action theory, there is just nothing out there that can serve for ''solving a philosophy problem'' anything like the role that is played by the contents of a good physics textbook for solving a physics problem. There is, indeed, little enough even by way of settled, consensus theories in such areas to even imagine what such a textbook would look like.

Another area in which many philosophers today have an indisputable mastery of a base of established knowledge is in formal logic. But the sorts of patterns seen in the experimental results are not in any way formally inconsistent. Taking Gettier cases to be cases of knowledge is not a contradiction—it is just, at worst, false. Similarly, there is no reason to think that any individual subject in the Swain et al. (2008) study had an actually contradictory set of judgments. Part of what is disturbing is just that different subjects had different sets of internally-consistent judgments, where this difference seems to be driven by the order in which the cases were presented. So there is no reason to think that better training in logical reasoning and detection of fallacies would be of much help here.

Even if philosophers currently had *some* theories of the right sort, the claim that their application would sufficiently inoculate their intuitions against restriction-generating factors is a substantive empirical claim. So, if the proponent of the expertise defense means to answer restrictionism while staying firmly planted in her armchair, this model will not work for her. The following toy example well illustrates the claim's substantive nature. Suppose that the so-called traditional theory of

knowledge, knowledge as justified true belief, is well-confirmed. The reader might think, then, that mastery of the content and the use of this theory is the key for stabilizing responses to the case orderings found in Swain et al. (2008). The thought is reasonable. But its truth or falsity is not armchair accessible. For all we know, the order effect arises because the clear cases of knowledge and non-knowledge differentially prime the cognitive structure(s) underlying categorizations of justified belief. If so, then the regular application of the traditional theory will not stabilize responses. The reason is that the locus of the instability is internal to the theory, so to speak. In a similar fashion, knowing that $F = ma$ would not lead to identical calculations of an object's force across contexts if estimates of its mass or acceleration were to vary across contexts. The reader might reply that what is needed then is a well-confirmed theory of justified belief. But that fails to address the root cause of the problem for the armchair proponent of the theory-driven model. How philosophically irrelevant factors affect our intuitive processes, and so whether the application of a theory or two can inoculate against their effects, are in general complex contingent matters. Such matters tend to be too messy for cognizers like us to discern from armchairs. The upshot is that whether the "better theory" hypothesis can meet restrictionism's challenge is a serious empirical question, one that philosophers should feel free to take up—scientifically, that is.

But, even setting aside this problem, the fact that philosophers' intuitions are in tune with our best philosophical theories doesn't show that they are immune from the kinds of problematic sensitivity that form the basis of the restrictionist challenge. That is, not unless it is presupposed that our best philosophical theories are supported by intuitions that are themselves immune from that kind of problematic sensitivity. But, if we could already be assured of that, then we would already have reason to take the restrictionist challenge as met. We find ourselves in the same explanatory regress as seen in the case of conceptual schemata.

4. Expertise as Practical Know-How about Considering Hypotheticals?

The two models of expert competence we have considered thus far both apply to the evaluative component of the intuiting process—the stuff that, once you're entertaining the hypothetical, tells you whether or not you've got an instance of the target property on your hand. Yet it might also be the case that philosophical training might improve one's competence at thought experiments by inculcating a *skill* at how best to entertain the hypotheticals in the first place. Much of the literature on expertise concerns not just factual competences but procedural ones (e.g., Anderson, 1982). And while it may be obvious that expertise in domains such as sports require such "knowledge how" as much as or more than "knowledge that," the relevance of procedural expertise also applies to such domains as auditing (Bonner & Walker, 1994), electronic circuitry diagnostics (Carlson, Khoo, Yaure, & Schneider, 1990), molecular biology (Sahdra & Thagard, 2003), and scientific reasoning more generally (Kuhn, 1989; Tchirgi, 1980; Zimmerman, 2000; though she

emphasizes that you can't totally disentangle skill from knowledge here). Procedural expertise can be expected to be found not just in sensorimotor tasks, but also in cognitive ones; for example, VanLehn writes, in the course of a summary of the history of psychological research on expertise, "phenomena that were often associated with motor skills, such as the power-law of practice and the identical elements model of transfer, were found to be important with cognitive skills as well" (1996, p. 515). So one might conjecture that philosophers, due to the specific training of their profession, might well outperform novices in intuition-mongering even if they were using the same evaluative capacities as the novices.

So we wish to explore the question: can philosophers be expected to have any particular expertise in the procedural aspects of intuition-mongering? But we run up against an initial challenge, which is that it is not quite clear what *specific* hypotheses of such expertise might be good contenders here. We have some idea what it would mean to have conceptual schemata that are more closely or efficiently hewed to the borders of its target, or to deploy a rich and well-established theory to derive an answer to a question. But here there are no *obvious* candidates for what a philosopher's procedural expertise might consist in—if it is to be something that would provide an appropriate cathedrist response to the restrictionist challenge, along the lines of the dialectic set forth at the end of section 1.

For example, one might conjecture that the folk in general pay less attention to the particulars of a case than philosophers do—that philosophers have a better ability to properly get all those details out of the description of the scenario, and entertain them fully in their imaginations. The conjecture is certainly plausible in light of work done on experts in other domains. Chase and Simon (1973), for example, report that a chess master, after having observed a chessboard for no longer than five seconds, can reproduce on average the placement of 16 out of 25 pieces, as long as they are arranged in a manner that can arise in actual chess matches. If they are arranged randomly, a master can reproduce very few placements, performing as novices do in both conditions. And this kind of finding is not isolated to chess. Similar results have been found with knowledgeable baseball fans (Spilich, Vesonder, Chiesi, & Voss, 1979), computer programmers (McKeithen, Reitman, Rueter, & Hirtle, 1981), medical diagnosticians (Patel & Groen, 1991), and even juvenile burglars (Logie, Wright, & Decker, 1992). It seems that, across lots of domains, experts have an expanded ability to extract quickly and to store temporarily lots of domain-related information. It is plausible that, when entertaining thought experiments, philosophers do as well. (Though one might also plausibly conjecture that the folk will pay more attention to the particulars, as opposed to deploying more abstract schemata. As we have stressed throughout, armchair theorizing about these psychological matters is highly unreliable! But put that aside for now.)

One might thus predict a systematic variation between the folk and trained philosophers, regarding cases in which the correct verdict turns on a very subtle detail. But that's not what is needed here, dialectically.[14] This hypothesis will not help explain away a difference in intuitions found between different groups of the folk, or between different orders of consideration of cases by the folk, that would lead us to

expect philosophers not to recapitulate the same variation. This particular hypothesized difference in skill does not do that, since it does not predict the right kind of systematic differences in reported intuitions.

One kind of difference that might make a difference here is suggested by a move that Ernest Sosa offers to fend off one part of the challenge:

> When we read fiction we import a great deal that is not explicit in the text. We import a lot that is normally presupposed about the physical and social structure of the situation as we follow the author's lead in our own imaginative construction. And the same seems plausibly true about the hypothetical cases presented to [the] subjects. Given that these subjects are sufficiently different culturally . . . they may, because of this, import different assumptions as they follow in their own imaginative construction the lead of the author of the examples . . . . Perhaps, for example, subjects who differ enough culturally . . . will import . . . different background assumptions about how likely it is that an American who has long owned an American car will continue to own a car and indeed an American car. (2009, pp. 107–108)

As it stands, this proposal has two big problems, of different kinds. First, it depends very strongly on a thoroughly unsubstantiated double-"perhaps": that the East Asian and Western European subjects actually had substantially different background assumptions in the first place; and that such different assumptions would sufficiently influence their knowledge attributions to bring about the observed differences. Whether the 'perhaps' is also an 'in fact' only be settled by a systematic empirical investigation of a sort that cannot be conducted from the armchair.

The second problem is articulated in Alexander and Weinberg (2007):

> If we cannot know that two experimental subjects are really disagreeing when they have putatively divergent intuitions, it would follow that we cannot know that two philosophers are really agreeing when they have putatively convergent intuitions. A skepticism about intuitions would be the result. Sosa's argument would thus score a victory for armchair philosophy that is both Pyrrhic and Pyrrhonian. (pp. 67–68)

In our terms here, Sosa's proposal is not one that would predict that philosophers should be immune to the cultural difference in the first place, and if we have to worry about such differences in filling-in, then we will have to worry about them with philosophers as much as the laity.

But it is here that an expertise-as-know-how proposal might be able to do some work. Plausibly, philosophers' training produces a *standardization* of what does or does not get filled in, such that cultural differences would not be reproduced. (Some of the possible psychological machinery of such a standardization, in terms of what is involved in mastering a genre, is considered in Weinberg, 2008.) Now, whether or not this is so is also an empirical question that would require some systematic investigation to decide, and its ultimate relevance would still depend on the issue mentioned above, as to whether or not the cultural differences can be explained in terms of differences in filling-in in the first place. So we do not think that this hypothesis is one that can now be counted as already meeting the restrictionist challenge. We offer it as an example of one kind of direction in which future

investigations into philosophical expertise might extend, and which could, should such investigations pan out, ultimately play a substantial role in meeting that challenge.[15] Though this has not been the most popular type of appeal to expertise made by cathedrists, it nonetheless may have a better chance of success than the appeals to finely-tuned conceptual schemata or better theories, which on inspection have not seemed shown themselves to be very plausible avenues of response to the restrictionist. We would encourage cathedrists—or anyone else—interested in the idea of philosophical expertise to try to develop some further hypotheses of thought experiment know-how that might serve here. The process of doing so, and of testing such hypotheses scientifically, would at the very least begin to shed more light on this cognitive activity that still lies at the dark heart of current analytic methodology.

We would note, though, that at least some good hypotheses for expert skills will not have any obvious parallels for philosophers, namely, those that involve developing a skill at using an external tool. Experts are often not such not purely in virtue of their own freestanding cognitive excellences, but because of their possession of and mastery with various external decision aids (and their failure to do so can sharply limit the excellence of their performance; see Shanteau, 1992; Sieck & Arkes, 2005). Where philosophers have something like that, we should expect training to be able to produce an appropriate skill, as with the proof techniques of formal logic as an aid to judgments of validity. But we do not currently have anything of this sort for judgments about cases of knowledge, or personhood, and so on.

5. Conclusion

Cathedrists can no longer offer quick armchair analogies between philosophers' judgments about thought experiments, and those of the practitioners of other fields in their own home domains. First and foremost, they cannot do so because the nature and extent of philosophical expertise in the relevant task, especially with regard to the kind of effects that restrictionists are talking about, is not a question that is particularly amenable to armchair inquiry. Answering it will require systematic and controlled empirical investigation; in other words, science.

Second, attention to the extant scientific literature suggests—though it certainly does not *prove*—that there are key disanalogies here. In general, other disciplines develop expertise in places where there are robust sources of feedback which can serve as a teaching signal for the would-be expert. The theories of the physicist, or the predictive models of the meteorologist, or even just the clear win/loss/draw signal at the end of the game for a chess player—these all provide the requisite sort of external check for which philosophers have drastically less that is comparable for their intuitive judgments about thought experiments. A further disanalogy that may be relevant here is that other areas by and large do not rely so centrally and heavily on their expert judgments as the key source of data for their further activities. So failures to develop real expertise in these particular judgment tasks in philosophy will have a much greater negative impact than it might in other professions.

The main point of our paper is to point out that playing the ''expertise card'' is not at all the trump in this debate that some philosophers have thought. Even if philosophers do, as a matter of fact, turn out to have the right sort of cognitive differences from novices to shield them from the restrictionist's findings, this isn't any help to the cathedrists unless they are willing to do a lot more work, and the right kind of work, *to show that this really is so*. Cathedrists offering the expertise defense have generally seemed to do so in order to close down a line of argument, namely, the restrictionist's argument from various results of experimental philosophy. Yet it turns out that they have at best only opened up this matter for discussion, and for further scientific investigation. We welcome further attempts to clarify what the nature of any philosophical intuitive expertise might be, and whether and how it might be inculcated, either by our practices as they stand or, perhaps more interestingly, by new practices as they might be developed. That philosophers have no reason to expect now, and from the armchair, that we are intuitive experts of the required sort, is not in itself a reason to think that we could not become, with the help of scientific psychology, the kind of experts that some cathedrists have assumed that we must already be.

Acknowledgements

This paper is highly collaborative, and authors are listed in reverse alphabetical order. The authors would like to thank Joachim Horvath; Edouard Machery; Stephen Stich; the organizers and participants of the 2008 ''Armchair in Flames?'' workshop at the University of Cologne; the philosophy faculty of the University of Nevada – Las Vegas; our colleagues at the Experimental Epistemology Laboratory; the organizers and participants of the 2009 ''Philosophy Without Intuitions?'' Conference at the Arche Centre at St. Andrews, especially Jonathan Ichikawa and Timothy Williamson; and an anonymous referee for this journal, for valuable feedback on earlier drafts of this paper.

Notes

[1] This is not to foreclose on the possibility of other lines of response, e.g., a parity of reasoning argument suggesting that the restrictionist's arguments would have to be skeptical and/or self-undermining. We will not consider such arguments here (though see Weinberg, 2007.)

[2] See Myers (2009) for a timely review of the literature. See also Devitt (2006) for a somewhat dissenting view on the expertise of linguists; but see also Culbertson and Gross (forthcoming).

[3] For a review, see Machery (2009, chapters 1–2).

[4] See also Murphy (2002) for other views on the connections between expertise and concepts.

[5] This is their usage; again, we intend to remain agnostic here on all questions as to the nature of concepts. Note also that ''higher-order concept'' here does not have what its default meaning might be to an analytic philosopher, i.e., that it be a concept pertaining to other concepts. Rather, the idea is that it is a concept that operates at a further remove from the most immediate details of a situation.

[6] For another approach to this idea in a domain perhaps more similar in structure to philosophy, see Zeitz (1994) on the differences in the "derived representations" of literary experts from those of non-experts.

[7] Boishuizen and Schmidt do not intend their hypothesis to conflict with other more familiar explanations of such results, e.g. that experts acquire more sophisticated *scripts* or are capable of managing larger and more specialized working memory *chunks*. They only introduce "encapsulation" to emphasize the ability of experts to unpack their chunks and verbalize lower-level justifications when needed.

[8] For more on the importance of reliable feedback, deliberate practice, concentration, and effective metacognition in the acquisition of expertise, see Ericsson, Krampe, and Tesch-Roemer (1993); Ericsson and Lehmann (1996, pp. 278–279); Feltovich et al. (2006, p. 60). Interestingly, Ericsson (2006) notes many domains where expert performance actually appears to *worsen*, in the face of continued experience, as a function of years since graduation from a training program.

[9] Another possibility unrelated to expertise but worth mentioning at this point is that philosopher's intuitions may not be tuned at all, but rather those students who fail to have the "consensus" intuitions on particular thought experiments get "weeded out" of programs (e.g., "I could not do metaphysics because I simply did not have the appropriate intuitions"). While no real tuning would occur in this case, the end result would be the same.

[10] Which is not to say that these distinctions are without significant extant philosophical puzzles! Indeed, one might wonder whether these distinctions can possibly count as clearer than the knowledge/non-knowledge distinction itself, and if philosophers are unproblematically masters of those more esoteric distinctions, how can we consistently aim to make trouble for such plain vanilla ones as knowledge/non-knowledge. But this would be to compare apples to oranges: restrictionists are not at all trying to undermine the general distinction between knowledge and its lack. They are more worried, rather, about some particular contours of that distinction, and more to the point, about how we are to determine the more subtle twists and turns in that perimeter. In areas where the distinctions like pragmatics/semantics become similarly problematic, so too will they provide less help in sifting reliable from unreliable uses of intuitions. It is not a matter of which distinctions are or are not somehow intrinsically clear, for almost all distinctions will have problematic and unproblematic zones. Rather, it is a matter of which distinctions have proved useful in offering at least some calibration of our intuitive methods.

[11] Consider our wealth of notational resources that philosophers have developed for helping to keep use straight from mention, or the set of tests (like explicit cancellation) we can apply to help disentangle pragmatic from semantic implicature.

[12] Some have noted that a speaker meaning/semantic meaning distinction may offer the means to such an argument with regard to the initial Machery et al. (2004) findings; see, e.g., Deutsch (2009) but also Deutsch, Carroll, Sytsma, and Machery (unpublished manuscript) for a empirical disconfirmation of such an argument. A clearer case is in the socio-economic status results of Weinberg et al. (2001). Because (i) the researchers there did not use any comprehension checks; (ii) the cases in which they found SES-based differences were long and complicated; and (iii) the low-SES (i.e., non-college-educated) subjects differed from the high-SES subjects by being closer to 50%—it follows that the pattern of answers they observed *might* just be a matter of the low-SES subjects very likely not comprehending the cases as well as the high-SES subjects, and thus just guessing more. Note that the viability of this confound depends essentially on very specific methodological issues, the particular pattern of this one set of observations, and the independent high likelihood of the hypothesis. Any attempts to provide such confounds for other findings must satisfy similar standards.

[13] Which is not to say that such "philosophers of X" couldn't often have their judgments improved by a serious inclusion of experimental methods in their toolbox. Since these areas are often vast, it is easy for a philosopher—for anyone—to over-generalize from what may turn out to be a biased or otherwise limited sample. See, e.g., Stotz and Griffiths (2004).

[14] With some, albeit very few, exceptions; see note 13.

[15] Sosa has recently offered a counter-response to these objections, pointing out that philosophical thought experiments take place in a context of active, reflective dialogue in which all parties are aiming for mutual comprehension. His hypothesis, then, is that such a setting allows us to be sufficiently sure that we (philosophers) are all filling-in in the same way. This also strikes us as a plausible hypothesis, but, again, one that would require actual empirical investigation. (We would note that it is not clear how it applies to the philosophical community *on the whole*—most philosophers encounter the thought experiments of most other philosophers only in asymmetric third-person settings like the printed page, and even a conference Q&A allows for only very limited exchange. For our purposes here, though, it is enough to note that Sosa's suggestion instantiates a different kind of cathedrist strategy, one that appeals not to the expertise of individual armchair practitioners, but rather to the extended set of *social epistemic practices* in which the armchair appeals to intuition are embedded. This category of cathedrist reply is one that we find intriguing, and certainly worthy of further exploration and critical consideration.

References

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

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 64–92). Cambridge: Cambridge University Press.

Ariely, D., Lowenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *Quarterly Journal of Economics*, *118*, 73–106.

Bonner, S. E., & Walker, P. (1994). The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review*, *69*, 157–178.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, *16*, 153–184.

Brown, C. (2009). Order effects and the audit materiality revision choice. *The Journal of Applied Business Research*, *25*, 21–36.

Camerer, C. F., & 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 & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 195–217). Cambridge: Cambridge University Press.

Carlson, R. A., Khoo, B. H., Yaure, R. G., & Schneider, W. (1990). Acquisition of a problem-solving skill: Levels of organization and use of working memory. *Journal of Experimental Psychology*, *119*, 193–214.

Charness, N. (1991). Expertise in chess: The balance between knowledge and search. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 39–63). Cambridge: Cambridge University Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Culbertson, J. & Gross, S. (forthcoming). Are linguists better subjects? *British Journal for the Philosophy of Science*.

Cullen, S. (forthcoming). Survey-driven romanticism. *European Review of Philosophy*.

Damisch, L., Mussweiler, T., & Plessner, H. (2006). Olympic medals as fruits of comparison? Assimilation and contrast in sequential performance judgments. *Journal of Experimental Psychology: Applied*, *12*, 166–178.

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

Deutsch, M., Carroll, C., Sytsma, J., & Machery, E. (unpublished manuscript). Cross-cultural semantics and the speaker's/semantic reference distinction.

Devitt, M. (2006). Intuitions in linguistics. *British Journal for the Philosophy of Science*, *57*, 481–513.

de Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). The Hague: Mouton (Original work published 1946).

de Waal, F. (2003). Silent invasion: Imanishi's primatology and cultural bias in science. *Animal Cognition*, *6*, 293–299.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge: Cambridge University Press.

Ericsson, K. A., Krampe, R. Th., & Tesch-Roemer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305.

Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies in expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 41–67). Cambridge: Cambridge University Press.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin*, *105*, 387–396.

Goldman, A. I. (2007). Philosophical intuitions: Their target, their source, and their epistemic status. *Grazer Philosophische Studien*, *74*, 1–26.

Hales, S. D. (2006). *Relativism and the foundations of philosophy*. Cambridge, MA: MIT Press.

Jackson, F. (1998). *From metaphysics to ethics: A defense of conceptual analysis*. Oxford: Oxford University Press.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press/Bradford Books.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*, 674–689.

Labov, W. (1975). Empirical foundations of linguistic theory. In R. Austerlitz (Ed.), *The scope of American linguistics* (pp. 77–133). Lisse: Peter de Ridder.

Labov, W. (1996). When intuitions fail. *Chicago Linguistics Society*, *32*, 77–105.

Lewandowsky, S., & Thomas, J. L. (forthcoming). Expertise: Acquisition, limitations, and control. In F. T. Durso (Ed.), *Reviews of Human Factors and Ergonomics: Vol. 5*. Santa Monica: Human Factors and Ergonomics Society.

Logie, R., Wright, R., & Decker, S. (1992). Recognition memory performance and residential burglary. *Applied Cognitive Psychology*, *6*, 109–123.

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

Lycan, W. (2006). On the Gettier problem problem. In S. Hetherington (Ed.), *Epistemology Futures* (pp. 148–168). Oxford: Oxford University Press.

Lynch, M. P. (2006). Trusting intuition. In P. Greenough & M.P. Lynch (Eds.), *Truth and realism: Current debates* (pp. 227–238). Oxford: Oxford University Press.

Machery, E. (2009). *Doing without concepts*. Oxford: Oxford University Press.

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

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naïve beliefs about the motion of objects. *Science*, *210*, 1139–1141.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, *13*, 307–325.

Murphy, G. L. (2002). *The big books of concepts*. Cambridge, MA: MIT Press.

Myers, J. (2009). Syntactic judgment experiments. *Language and Linguistics Compass*, *3*, 406–423.

Norman, G., Eva, K., Brooks, L., & Hamstra, S. (2006). Expertise in medicine and surgery. In K.A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 339–353). Cambridge: Cambridge University Press.

Page, Scott (2007). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton: Princeton University Press.

Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 93–125). Cambridge: Cambridge University Press.

Petty, R., Tormala, Z., Hawkins, C., & Wegener, D. (2001). Motivation to think and order effects in persuasion: The moderating role of chunking. *Personality and Social Psychology Bulletin*, *27*, 332–344.

Phillips, C. (forthcoming). Should we impeach armchair linguists? In S. Iwasaki (Ed.), *Japanese/Korean Linguistics: Vol. 17*. Palo Alto: Center for the Study of Language and Information.

Reif, F., & Allen, S. (1992). Cognition for interpreting scientific concepts: A study of acceleration. *Cognition and Instruction*, *9*, 1–44.

Sahdra, B., & Thagard, P. (2003). Procedural knowledge in molecular biology. *Philosophical Psychology*, *16*, 477–498.

Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. A. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine*, *65*, 611–621.

Schuetze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

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

Shanteau, J., & Stewart, T. R. (1992). Why study expert decision making? Some historical perspectives and comments. *Organizational Behavior and Human Decision Processes*, *53*, 95–106.

Sieck, W., & Arkes, J. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making*, *18*, 29–53.

Singer, P. (1974). Moral experts. *Analysis*, *32*, 115–117.

Singer, P. (1982). How do we decide? *The Hastings Center Report*, *12*, 9–11.

Sinnott-Armstrong, W. (2008). Framing moral intuitions. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Volume 2: The cognitive science of morality* (pp. 47–76). Cambridge, MA: MIT Press.

Sosa, E. (2009). A defense of the use of intuitions in philosophy. In D. Murphy & M. Bishop (Eds.), *Stich and his critics* (pp. 101–112). Oxford: Wiley.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, *18*, 275–290.

Stotz, K., & Griffiths, P. (2004). Genes: Philosophical analyses put to the test. *History and Philosophy of the Life Sciences*, *25*, 5–28.

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

Sytsma, J., & Livengood (forthcoming). A new perspective concerning experiments on semantic intuitions. *Philosophical Psychology*.

Tchirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, *51*, 1–10.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology*, *47*, 513–539.

Wasow, T., & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua*, *115*, 1481–1496.
Weinberg, J. M. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy*, *31*, 318–343.
Weinberg, J. M. (2008). Configuring the cognitive imagination. In K. Stock & K. Thomson–Jones (Eds.), *New Waves in Aesthetics* (pp. 203–223). New York: Palgrave Macmillan.
Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, *29*, 429–60.
Williamson, T. (2005). Armchair philosophy, metaphysical modality and counterfactual thinking. *Proceedings of the Aristotelian Society*, *105*, 1–23.
Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell.
Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.
Zamzow, J. & Nichols, S. (forthcoming). Variation in ethical intuitions. *Philosophical Issues*.
Zeitz, C. M. (1994). Expert-novice differences in memory, abstraction, and reasoning in the domain of literature. *Cognition and Instruction*, *12*, 277–312.
Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, *20*, 99–149.