

Philosophical Expertise and the Burden of Proof*

(penultimate version of paper forthcoming in 40th anniversary issue of *Metaphilosophy*)

Timothy Williamson

ABSTRACT: Some proponents of ‘experimental philosophy’ criticize philosophers’ use of thought experiments on the basis of evidence that the verdicts vary with truth-independent factors. However, their data concern the verdicts of philosophically untrained subjects. According to the expertise defence, what matters are the verdicts of trained philosophers, who are more likely to pay careful attention to the details of the scenario and track their relevance. In a recent paper, Jonathan Weinberg and others reply to the expertise defence that there is no evidence for such expertise. I reply to them in this paper, arguing that they have misconstrued the dialectical situation. Since they have produced no evidence that philosophical training is less efficacious for thought experimentation than for other cognitive tasks for which they acknowledge that it produces genuine expertise, such as informal argumentation, they have produced no evidence for treating the former more sceptically than the latter.

KEYWORDS: Experimental philosophy; expertise; thought experiments.

1. An eye-catching feature of contemporary analytic philosophy is the argumentative weight it lays on thought experiments. This feature has been the target of an extended critique by self-described ‘experimental philosophers’, from Weinberg, Nichols and Stich 2001 on. They have conducted extensive trials of some well-known philosophical thought experiments on a variety of subjects under a variety of circumstances. Their results suggest that the answers given to key questions in the thought experiments are sensitive to the ethnicity of the subjects, the order in which the questions are asked and other factors presumably irrelevant to the truth of the answers. On this basis, experimental philosophers have argued that the use of thought experiments in philosophy should be substantially restricted, because on our current evidence they do not deserve our trust.

In *The Philosophy of Philosophy* (2007), I developed an account of thought experiments in philosophy as employing deductively valid arguments with counterfactual premises which we evaluate as we evaluate other counterfactuals, using a mixture of imaginative simulation, background information and logic. In response to the experimental philosophers’ critique, I noted that their trials have not been conducted on professional philosophers but on lay subjects, typically undergraduates, with little or no philosophical training: ‘Yet philosophy students have to learn how to apply general concepts to specific examples with careful attention to the relevant subtleties, just as law students have to learn how to analyze hypothetical cases. Levels of disagreement over thought experiments seem to be significantly lower among fully trained philosophers than among novices. [...] We should not regard philosophical training as an illegitimate contamination of the data, any more than training natural scientists how to perform experiments properly is a contamination of their data. Although the philosophically

innocent may be free of various forms of theoretical bias, just as the scientifically innocent are, that is not enough to confer special authority on innocent judgment, given its characteristic sloppiness' (2007, 191). Call this way of defending the use of thought experiments in contemporary philosophy *the expertise defence*.

As the quotation makes clear, the expertise defence does not imply that a good philosophical education involves the cultivation of a mysterious *sui generis* faculty of rational intuition, or anything of the kind. Rather, it is supposed to improve far more mundane skills, such as careful attention to details in the description of the scenario and their potential relevance to the questions at issue.

In 'Are Philosophers Expert Intuiters?' (2010), four experimental philosophers — Jonathan Weinberg, Chad Gonnerman, Cameron Buckner and Joshua Alexander (WGBA) — provide the best developed response to the expertise defence currently available. In brief, WGBA argue that whether philosophical training confers genuine expertise (significantly greater reliability) in conducting thought experiments is a squarely empirical question, to be answered by detailed empirical investigation in the light of the extensive scientific literature on expertise, and that the burden of proof is on proponents of the expertise defence to carry out such investigations and show that they deliver the requisite results. Since no such detailed investigations have in fact been carried out, the four authors treat the experimental critique as still holding the field: in their view, philosophers are not currently justified in laying argumentative weight on thought experiments as they do.

This paper is a response to WGBA.¹ I will argue that they have misconstrued the dialectical situation, that it is currently the experimental critique of professional

philosophers' use of thought experiments that lacks adequate evidential support, and that philosophers *are* currently justified in laying argumentative weight on thought experiments. Of course, it is never completely satisfying just to return the burden of proof to one's opponents. It would be more fun to lay out a vast array of specific experimental evidence for the value of philosophical training in improving performance with thought experiments. However, it will be a long time before we have strong evidence of that kind one way or the other, and in the meantime philosophers must get on with their job. They should not be expected to abandon their use of thought experiments when there is no good evidence that doing so would improve their philosophizing. Nor should philosophers be expected to suspend their current projects in order to carry out psychological investigations of their capacity as thought experimentalists, on the basis of evidence that undergraduates untrained in philosophy are bad at conducting thought experiments. After all, we do not expect physicists to suspend their current projects in order to carry out psychological investigations of their capacity as laboratory experimentalists, on the basis of evidence that undergraduates untrained in physics are bad at conducting laboratory experiments. Standards of laboratory experimentation in physics are doubtless higher than standards of thought experimentation in philosophy; nevertheless, in both cases the point remains that it would be foolish to change a well-established methodology without serious evidence that doing so would make the discipline better rather than worse.

2. WGBA describe the target of the experimentalist critique as 'analytic philosophy's longstanding practice of deploying armchair intuitive judgments about

cases' (331).² This description is a little misleading. Critics of 'armchair philosophy' tend to forget that there are real life analogues of some philosophical thought experiments; stopped clocks really do show the right time twice a day (Williamson 2007, 192-5). One can argue against the justified true belief account of knowledge just as easily with such real life Gettier cases as with the original fictions. But experimental philosophers never suggest that actualizing the scenarios of thought experiments would help solve the methodological problem. Rather, in discussion they have typically been quick to insist that their critique should be applied equally to the analogues for the real life cases of the judgments at issue in philosophical thought experiments. Since the real life cases can be encountered far from the armchair, the word 'armchair' should be deleted from WGBA's description of the target of the experimentalist critique. That leaves 'analytic philosophy's longstanding practice of deploying intuitive judgments about cases'. Presumably, WGBA have nothing against deploying judgments about cases; one does not make philosophy more scientific by compelling philosophers to speak only in generalities. Thus the weight falls on 'intuitive'. Unfortunately, WGBA do not explain what they mean by the word. When is a judgment about a case intuitive? If I judge 'You do not know how many coins I have in my pocket', is that an intuitive judgment about a case? If experimental philosophers judge 'There is currently insufficient evidence to deny that there is knowledge in this Gettier case', is that an intuitive judgment about a case? WGBA give no help in answering such questions. If such examples do count as judgments we are not currently justified in trusting, then the experimentalist critique is self-destructively general. If we are currently justified in trusting such judgments about

cases, what is supposed to differentiate them from those judgments in which, according to the experimentalist critique, we should not trust?

The extreme unclarity about the target of the experimentalist critique does not render the critique completely vacuous. It is clear at least that full-dress philosophical thought experiments are supposed to lie in the centre of the target area; what is unclear is how far out the area is supposed to extend. That is not simply a matter to be left for further experimental investigation. For, according to the experimentalists, on *present* evidence we should already be withdrawing our trust from some ‘judgments about cases’; they should tell us, at least roughly, which ones. Presumably, they feel justified in assuming that their own judgments in the paper fall outside the present target area. Again, when they discuss ‘the areas of philosophy in which appeals to intuition about cases are still central, such as epistemology and action theory’ (345), they treat themselves as already having some capacity to discriminate between what is an appeal to intuition about a case and what is not.³

Having signalled this major problem with the experimentalist critique, I will not elaborate on it in what follows. Nor will I discuss objections that have been raised to details of the experimental designs, such as the wording of the questions. Moreover, I am quite willing to grant that the apparent disagreements between the answers of different subjects were genuine, so that if one answer was true another was false.⁴ My concern is with the experimentalist response to the expertise defence.

I will not be questioning the expertise literature, or WGBA’s interpretation of it. In one respect they sometimes misrepresent the expertise defence itself, when they speak of their opponents as claiming that a philosophical education ‘immunizes’ one against the

influence of whatever psychological factors distort the judgments of untrained subjects in their trials. It is not plausible that philosophical training will totally eradicate such influence, just as it is not plausible that historical training will totally eradicate the influence of whatever psychological factors distort the judgements of untrained subjects about historical matters. But the expertise defence requires no such extreme claim. The defence is vindicated if philosophical training substantially reduces the influence of the distorting factors, even short of total eradication. WGBA's more circumspect formulations acknowledge this obvious point: '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*' (333, their italics).

3. In assessing the dialectical status of the expertise defence, it is useful to start with some general points about observational evidence. Since they are near-platitudes, they are presumably points of agreement in theory between proponents and opponents of the expertise defence. The issue will be whether opponents of the defence have respected them in practice.

Experimentation and other systematic forms of observational evidence-gathering use scarce resources of time, energy and money (for brevity, I will say only 'experiments' in what follows). Even on a comparatively long timescale, the human race will only perform a tiny fraction of all the experiments it is humanly feasible to perform. Many possible experiments appear to lack any value; no outcome of them appears to provide significant evidence on any significant theoretical or practical question. Other possible

experiments have more apparent value than that, but still deserve far lower priority than more urgent ones to which the resources should go instead.

What attitude should we take to the outcome of an unperformed experiment? It may sound laudably open-minded to insist that we should not commit ourselves as to the outcome. On reflection, however, that attitude reveals itself as a damaging form of scepticism. For let T be a scientific theory so well confirmed by a mass of experimental and theoretical considerations that it is unreasonable to continue testing T, and reasonable to commit ourselves to T. Nevertheless, we cannot have separately tested *all* the experimentally testable consequences of T, since there are infinitely many. Thus T has some experimentally testable but untested consequence O. The proposed attitude to unperformed experiments requires us not to commit ourselves to O. But since T entails O, commitment to T involves commitment to O. Thus the proposed attitude requires us not to commit ourselves to T. But, by hypothesis, it is reasonable to commit ourselves to T. Thus the attitude requires us not to do something it is in fact reasonable to do. Hence the attitude is not binding. Indeed, it is worse than that. For the argument is very general: the attitude in question forbids commitment to virtually any scientific claim, however well confirmed within the limits of human feasibility. We should not take such an attitude. The case of scepticism about global warming shows just how pernicious such an ‘open-minded’ attitude to missing data can be. No one is more dogmatic than the sceptic in his scepticism. It is sometimes reasonable to commit oneself as to the outcome of an experiment that has never been performed, and perhaps never will be. More generally, it is sometimes reasonable to commit oneself to a hypothesis (such as O) that could be tested by systematic experiment but never has been, whether or not it ever will be.

Care is needed in applying the argument. Presumably, T does not entail that the experiment will not be performed incompetently or on an unluckily unrepresentative sample. It may be unwise to assume that no misfortune or mistake will occur in the performance of the experiment. But that is not the issue. As a consequence of T, O too does not rule out such performance noise. What is reasonable is to commit oneself to O itself, which could be tested by systematic experiment but never has been. Similarly, the mere fact that the expertise defence could be tested by systematic experiment but never has been is consistent with the present reasonableness of commitment to the expertise defence. Any critique of it must be based on far more specific considerations.

For purposes of comparison, consider the hypothesis that professional physicists tend to display substantially higher levels of skill in cognitive tasks distinctive of physics than laypeople do. The hypothesis could be tested by systematic experiment. But even before that has happened, one can reasonably accept it. More generally, consider how philosophers of science (in the broadest sense) proceed when working on the philosophy of mathematics, physics, chemistry, biology, psychology, economics, linguistics, history or almost any other academically well-established discipline with departments in most major universities across the world. They normally assume that professional academics in that discipline tend to display substantially higher levels of skill in its distinctive cognitive tasks than laypeople do. For example, they assume that professional judgments on its distinctive questions carry more weight than do the judgments of laypeople or philosophers. The assumption is defeasible: external criticism of the discipline is not forbidden, but must be based on a body of evidence strong enough to defeat the initial presumption that the professionals are the people best placed to distinguish between good

and bad work within their own discipline. In practice, that initial presumption is hard but not impossible to overturn.

Of course, professional training filters as well as educates. Professional academics in a discipline might tend to display substantially higher levels of skill in its distinctive cognitive tasks than laypeople do even if their professional training did not enhance those skills, but merely selected people who already had them to a higher degree than others did. In practice, that ‘mere selection’ hypothesis is grossly implausible for many cognitive skills in most academic disciplines. If it were true of skill in thought experimentation in philosophy, that would anyway suffice for purposes of the expertise defence, but in this paper the focus is on professional academic training as an enhancer of cognitive skills in given individuals.

To some extent, the efficacy of professional training in academic disciplines as an enhancer of relevant cognitive skills is a matter of common experience. In principle, it can be assessed in more systematic ways too, but such assessment itself involves reliance on cognitive skills distinctive of an academic discipline such as psychology. Without an initial presumption that such skills are higher amongst those with relevant professional training than amongst laypeople, the assessment would be problematic. Moreover, it is hard to devise and apply credible tests of a skill in an intellectual discipline without relying on someone’s already accredited skill in that very discipline. If every implicit claim to cognitive skill faced a burden of experimental proof, inquiry would grind to a halt. The defeasible presumption in favour of the relevant cognitive skills of those trained in a discipline plays a significant role in enabling intellectual progress.

From a sociological perspective, philosophy is a fairly normal academic discipline. Consequently, since thought experimentation is a cognitive task distinctive of contemporary analytic philosophy, the initial presumption should be that professional analytic philosophers tend to display substantially higher levels of skill in thought experimentation than laypeople do. Although that initial presumption is in principle open to experimental test, it does not follow that the onus is on proponents of the expertise defence to do the testing. Rather, the burden of proof is on experimental philosophers to demonstrate that, contrary to initial expectations, professional training in analytic philosophy fails to enhance skill in one of its central cognitive tasks, and the corresponding professional qualifications do not select for such skill. They must point to specific features of our present evidence that tell against the expertise defence. What are those features?

Thoughts naturally turn to the difference in track record between philosophy and many other academic disciplines. Although it would be myopic to deny that philosophy has made *some* progress, one must admit that in most areas it has not made as much progress as the natural sciences (formal logic is an exception). The suggestion is that the comparative lack of philosophical progress is what defeats the initial presumption in favour of genuine philosophical expertise. However, this is not what WGBA intend, for it does not distinguish between different cognitive skills in philosophy. For some cognitive skills, WGBA explicitly concede that philosophical expertise is genuine. In particular, they assert that ‘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’ (334), without providing any experimental

evidence such as they require their opponents to produce for genuine expertise in thought experimentation. Similarly, they grant ‘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’ (335), without explaining how such skills have been demonstrated in ways for which thought experimentation would have no analogue. In these cases, they treat the positive effect of philosophical training as obvious. Thus their objection to the expertise defence must turn on specific differences between thought experimentation and other cognitive skills in philosophy, not on the general phenomenon of philosophy’s poor track record.

Thought experiments in any case constitute an unpromising scapegoat for the discipline’s lack of progress, for if the category is understood narrowly enough to save the experimentalist critique from self-defeat, it has played a comparatively small role in the history of philosophy, even though one can find examples in Plato and other great philosophers. Nor was thought experimentation to blame for what experimental philosophers might regard as some of the more embarrassing episodes in the history of philosophy, such as the shift from logic to rhetoric in the Renaissance or the idealist turn in the eighteenth and nineteenth centuries (to paint with the broadest of brushes).

WGBA must therefore specify which differences between thought experiments and other cognitive tasks in philosophy are supposed to explain why the philosophical training they presume to enhance the latter cannot be presumed to enhance the former. One of the problems they face in doing so is that thought experimentation *overlaps* the skills they presume philosophical training to grant. For example, ‘the close analysis of texts’, which WGBA describe as a ‘demonstrable skill’ possessed by philosophers, is

exactly what one needs adequately to take in and digest the description of the scenario in a thought experiment. Similarly, on many accounts of thought experiments, including that in *The Philosophy of Philosophy*, thought experiments employ arguments. In effect, conducting a thought experiment is a special case of ‘argument construction and evaluation’, which WGBA describe as a ‘technical skill’ that ‘philosophical training does typically bring’. WGBA appear not to notice this problem. Although they might classify the areas of overlap as somehow untypical (they would need to say why), the tasks for which they regard philosophical expertise as presumptively bogus are strikingly close to some of those for which they regard it as obviously genuine.

WGBA accuse proponents of the expertise defence of giving a merely generic argument, without the requisite specificity to skill in thought experimentation. They conjecture that we are relying on a ‘folk theory of expertise’ according to which ‘expertise at one aspect of an activity is closely correlated with expertise in other aspects of that activity’ (333). I rely on no such theory. It takes very little experience of teaching philosophy to know that expertise in solving logic problems is not closely correlated with expertise in reading historical texts. WGBA cite my comparison between the training of philosophers and the training of lawyers as an example of the generic approach, failing to notice that the comparison was specific to skills relevant to thought experimentation: ‘philosophy students have to learn how to apply general concepts to specific examples with careful attention to the relevant subtleties, just as law students have to learn how to analyze hypothetical cases’ (Williamson 2007, 191). Nothing they say undermines the analogy. They neglect it just as they neglect the overlap between the skills they explicitly treat as enhanced by philosophical training and those relevant to thought experimentation.

4. WGBA do try to identify some relevant differences between thought experiments and other cognitive tasks in philosophy in terms drawn from the scientific literature on expertise. In that literature, various characteristics of training regimes have turned out to be conducive to the production of genuine expertise. WGBA maintain that these characteristics are absent from philosophical training with respect to thought experiments (and presumably not with respect to the cognitive tasks at which they take philosophical training to confer genuine expertise). We might therefore interpret WGBA as accepting the gist of the analysis in section 3 of the dialectical situation, while attempting to discharge the burden of proof on them by providing specific evidence of the relevant differences between thought experiments and other cognitive tasks in philosophy.

From WGBA's discussion of the expertise literature, one can extract three characteristics of training regimes that have turned out to be conducive to the production of genuine expertise. They are:

- (a) repetitive practice with fast, accurate feedback;
- (b) decomposition of the task into sub-tasks;
- (c) use of external decision aids.

I accept that (a)-(c) are conducive to the production of genuine expertise, and that their absence has the opposite effect. In the published paper, WGBA concentrate on arguing that training regimes in philosophy are deficient with respect to (a). Let us take each feature in turn.

(a): By the time one has a Ph.D. in analytic philosophy, one has typically read many dozens of articles and books in which thought experiments play a key role, thought,

talked and written about them on numerous occasions, and received extensive feedback on one's reactions from one's teachers, much of it immediate (for example, in class). These uses of thought experiments often involve exploring many variations on the same theme (brains in vats, twin earths, Gettier cases, trolley cases). According to WGBA, the number of such occasions for a given individual is still orders of magnitude less than for a chess player practicing a given opening (342). But who ever claimed that the difference in skill at thought experimentation between a professional philosopher and an undergraduate is as dramatic as the difference in skill at chess between a grandmaster and a beginner? A more relevant comparison is with the number of occasions on which the trainee philosopher receives feedback with respect to philosophical skills for which WGBA acknowledge the efficacy of a standard training, such as the close analysis of texts, and the critical assessment of arguments. Another relevant comparison is between feedback in legal and philosophical training with respect to hypothetical cases. WGBA's vague remarks ignore the more appropriate comparisons. They also confuse the issue by failing to distinguish between feedback for trainee philosophers and feedback for already trained philosophers (341-2). In short, they provide no serious evidence of deficiency with respect to (a), and so fail to shift the burden of proof onto their opponents.

(b) It is not hard to decompose the task of thought experimentation into consciously discernible sub-tasks. First, one must read and digest the description of the scenario; this is the part that corresponds to WGBA's 'demonstrable skill' of 'the close analysis of texts'. Then one must judge what would be the case in the scenario described, which in turn often decomposes into answering several questions, such as 'Is it a belief?', 'Is it true?', 'Is it justified?' and 'Is it knowledge?'. One must also judge whether the

scenario is really possible, for otherwise the thought experiment may not be fit for purpose. Finally, one must determine whether the premises, if verified, do entail the proposed conclusion; this part corresponds to WGBA's 'technical skill' of 'argument construction and evaluation'.

(c) Formal methods as decision aids facilitate some, although not most, thought experiments. For example, consider the proposed law of tense logic 'If P then it will be the case that it was the case that P'. One can test it by a thought experiment in which one envisages a last moment of time, using formal techniques to check that the schema has a false instance in that scenario. The exercise is no merely formal one, for it concerns the intended interpretation of the tense operators. A more commonplace example is the regular use of outcome tables and other visual aids in perspicuously displaying the structure of thought experiments in decision theory. Although aids of that kind are 'purely notational', a good notation can do much to facilitate understanding and insight, as mathematicians know.

On closer inspection, therefore, philosophical training with respect to thought experiments may have about two and a half of the three characteristics conducive to the production of genuine expertise, for all WGBA say. Their elaborate invocation of the expertise literature threatens to undermine their own argument.

WGBA make several points that could be construed as objections to the foregoing assessments of (a)-(c). These points must now be evaluated.

First, WGBA insist that we cannot determine from the armchair *how much* practice is needed for genuine expertise, and likewise for the other factors. That is obviously correct, but it is a quite generic point; it does not discriminate between thought

experimentation and the skills WGBA acknowledge to be developed by philosophical training. For example, my comments about practice and feedback on thought experimentation could equally be applied to practice and feedback on ‘argument construction and evaluation’ (WGBA’s comments on that ‘technical skill’ are not aimed at formal logic, and their own arguments are informal). After all, it is often the thought experiments that absorb classroom time because their vivid details grip the imagination, to the detriment of drier material on the structure of informal arguments. Since WGBA provide no evidence that thought experimentation fares worse in such respects than the other skills, they give no reason to expect philosophical training to be relevantly less efficacious for the former than for the latter.

Second, to the suggestion ‘that philosophers train their intuitions against other, already-certified expert intuitions’, WGBA respond ‘this appears to be a non-starter, since it just invites an explanatory regress: how did the purveyors of *those* intuitions develop their expertise?’ (341). Such an objection might be made concerning the feedback philosophy students receive from their teachers on thought experiments, mentioned above under (a). Incompetent feedback is not conducive to genuine expertise. This point too is dangerously generic for WGBA’s purposes. When students receive feedback from their teachers on ‘argument construction and evaluation’ or ‘the close reading of texts’, how did their teachers develop their expertise? The infinite regress concern would be more serious if thought experimentation did not decompose into subtasks, for then there might seem to be little for the feedback to consist of beyond bare verdicts. Even there, however, the teacher might also suggest other related thought experiments for purposes of comparison. Moreover, in most branches of philosophy there

are many sufficiently uncontentious thought experiments, such as fictional cases of *unjustified* true beliefs that do not constitute knowledge, on which beginners are often started; it is their very uncontentiousness that makes them comparatively inconspicuous. In any case, given the decomposition of the task of thought experimentation into subtasks, described under (b), feedback can be far more articulated. For example, the teacher can draw the student's attention to overlooked aspects of the description of the scenario. In any academic discipline, the capacity of teachers to provide correct and useful feedback depends to some extent on the teachers' expertise, but the regress need not be vicious. We sometimes have a high enough level of expertise to bootstrap ourselves to a higher level of expertise by mutual criticism without input from anyone already at the higher level. Pupils sometimes surpass their teachers without having more innate ability. WGBA provide no evidence that this does not happen for thought experimentation just as it happens for other cognitive skills.

Third, WGBA complain about a hypothesis on which trained philosophers do better than laypeople when 'the correct verdict turns on a very subtle detail' that it is 'not what is needed here, dialectically' because it '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' (347-8). But that is to impose an unreasonable explanatory demand. The effect of education is often to increase uniformity on some cognitive task; explaining the effectiveness of the education need not involve explaining the specific patterns of variation amongst the uneducated. For example, one can explain why very few professional historians are Holocaust deniers or very few professional

biologists are creationists without explaining why Holocaust denial or creationism is much commoner amongst relevantly uneducated people in some countries than in others. WGBA provide no reason to expect a different pattern in philosophical training on thought experiments.

In summary, the dialectical situation is this. The experimental critique presents evidence that philosophically untrained subjects perform poorly at thought experimentation, a cognitive task characteristic of contemporary analytic philosophy. In general, given a cognitive task characteristic of a discipline, it is unwarranted to project data about the performance at the task of subjects untrained in the discipline onto subjects trained in the discipline, without specific evidence that training in the discipline makes no substantial difference to skill at that task. WGBA's attempt to provide such specific evidence consists of a few vague and casual claims about training in philosophy and thought experimentation. They provide no significant evidence that thought experimentation is worse off in the relevant respects than the cognitive skills they acknowledge to be enhanced by training in philosophy, such as informal argumentation and the close analysis of texts. Consequently, they provide no reason to rely less on trained philosophers' skill at thought experimentation than on their skill at those other cognitive tasks.

5. The fear is sometimes expressed that philosophical training merely enforces orthodoxy in thought experiments. It socializes the malleable into eventually accepting the standard judgments, whatever their initial views. Those who stubbornly resist are excluded from the profession. They fail to get into a top graduate school, or fail to get their doctoral

dissertation accepted, or fail to get a proper job in philosophy. Even if they somehow manage to sneak into the profession, referees for prestigious journals and publishers reject their papers and book manuscripts. WGBA briefly raise such a possibility (351).

Of course, one can see academic training in many disciplines in such reductively sociological terms. It surely has some tendency to filter out unpopular views in all academic fields, including the natural sciences. But a view may be unpopular for good reason. By the arguments above, the onus is on those who suspect the professional consensus in philosophical thought experiments of being a merely sociological phenomenon to provide solid evidence for their suspicion, to distinguish this professional consensus from more benign ones. Otherwise the suspicion is just one more conspiracy theory. WGBA provide the sceptic with no such evidence.

We have more to rely on than that general consideration. As WGBA note, philosophical training fosters a variety of cognitive skills, which they treat as obviously genuine (close analysis of texts, argument construction and evaluation, formal logic ...). We might expect that if thought experimentation were a rogue pseudo-skill, orthodoxy in thought experiments would be at best poorly correlated with possession of all or most of the genuine cognitive skills in philosophy. Since a significant minority even of Western students give unorthodox responses to thought experiments, according to the experimental philosophers' own results, such responses should sometimes be combined with genuine cognitive skills in philosophy, if the latter are poorly correlated with orthodoxy. Given that highly rated performance on most dimensions can compensate for poorly rated performance on one or two in academic tests, we should not expect philosophical training to exclude all or almost all of those who deviate from orthodoxy in

thought experiments, any more than it excludes all or almost all of those who are not much good at formal logic.

Furthermore, orthodoxy in thought experiments is not all or nothing. Someone who ascribes knowledge in a Gettier case may give orthodox answers in other thought experiments. If they fail in epistemology, they can try metaphysics or moral philosophy instead. If they are good enough in all other respects, they can still make it in the profession. Having achieved tenure and prestige, they are in a position to go back to their old grievance, deliver lectures in which they skilfully construct arguments to show that their unorthodox answer in the thought experiment fits a better overall theory, and use their reputation to have their arguments published in books and papers. After all, a powerful challenge to orthodoxy brings rich professional rewards in philosophy.

Once one seriously considers what it would take to enforce a given response to a particular thought experiment across the philosophical profession purely by a process of social exclusion, with no deeper cognitive basis, the scenario looks increasingly paranoid. It is, in any case, not the scenario most experimental philosophers had in mind.

6. The claims of this paper do not entail that we should be complacent about trained philosophers' skill at thought experimentation. There are too many internal tensions between common verdicts in different cases for that.⁵ But we should also not be complacent about trained philosophers' skill at the construction and evaluation of informal arguments. Given the widespread negative evaluations of the experimental philosophers' informal arguments, and the many arguments against their conclusions, experimental philosophers presumably cannot rate trained philosophers' skill in that

respect very highly either. Plainly, however, the proper response is not to give up the practice of informal argumentation in philosophy. That would only make things worse (*much* worse). Rather, we must try to refine the practice from within, as we do. Why should we not do the same with thought experimentation?

Psychological evidence may well have a significant role to play in refining our skill at thought experimentation. It can alert us to unexpected sources of bias and distortion in our verdicts, and help us correct for them. We are likely to have most to learn from general psychological theories of judgment that are well-established on the basis of a broad range of evidence, rather than from data gathered with a specific philosophical (or anti-philosophical) agenda on complex, philosophically contested judgments. Some such work is already available.⁶ That is a far more promising way forward than a wholesale ban on thought experimentation. Indeed, given the point from section 1 that the target of the experimental critique is not just thought experimentation but the more general practice of relying on ‘intuitive judgments about cases’, whether made in or out of the armchair (since otherwise the experimental critique would not make the intended difference), it is quite unclear what philosophy without the practice at issue would be, if such a thing is even possible.

Consider, for example, a theory of confirmation. We may hope to test it by drawing out its predictions for a range of specific counterfactual cases, kept artificially simple in order to make it as clear as possible, independent of the theory, which hypotheses would really be better-confirmed than which. Those tests are thought experiments. To follow the experimentalists’ advice not to use such tests is to make philosophy less scientific, not more.

New College

Oxford OX1 3BN

U.K.

timothy.williamson@philosophy.ox.ac.uk

References

- Nagel, Jennifer. 2008. "Knowledge ascriptions and the psychological consequences of changing stakes", *Australasian Journal of Philosophy* 86, no. 2 (June): 279-94.
- Nagel, Jennifer. 2010. "Knowledge ascriptions and the psychological consequences of thinking about error", *The Philosophical Quarterly* 60, no. 239 (April): 286-306.
- Weinberg, Jonathan M. 2009. "On doing better, experimental-style", *Philosophical Studies* 145, no. 3 (September): 455-464.
- Weinberg, Jonathan M., Chad Gonnerman, Cameron Buckner, and Joshua Alexander. 2010. "Are philosophers expert intuiters?", *Philosophical Psychology* 23, no. 3 (June): 331-55.
- Weinberg, Jonathan M., Shaun Nichols, and Stephen Stich. 2001. "Normativity and epistemic intuitions", *Philosophical Topics* 29, nos. 1 and 2: 429-60.
- Williamson, Timothy. 2005. "Contextualism, subject-sensitive invariantism and knowledge of knowledge", *The Philosophical Quarterly* 55, no. 219 (April): 213-35.
- Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.
- Williamson, Timothy. 2009. "Replies to Ichikawa, Martin and Weinberg", *Philosophical Studies* 145, no. 3 (September): 465-476.

Notes

* I thank Jonathan Weinberg for detailed written comments on related material, both him and others for useful discussion at an Arché workshop on ‘Philosophy without Intuitions?’ at St Andrews, a graduate conference at Trinity College Dublin, the 26th International Philosophical School of the Institute for Philosophical Research of the Bulgarian Academy of Sciences on ‘Applied and Experimental Philosophy in Knowledge Based Society East and West’ in Sofia, and a meeting of the Oriel College Philosophy Society in Oxford, where I presented other versions of this paper, and likewise participants in the symposium marking the 40th anniversary of the founding of the journal *Metaphilosophy*, ‘The Future of Philosophy: Metaphilosophical Directions for the 21st Century’, at the Institute of Philosophy, School of Advanced Studies, University of London.

1 The present paper builds on points briefly made in Williamson 2009, 471-75, in response to Weinberg 2009.

2 All quotations are from, and pages references to, Weinberg, Gonnerman, Buckner and Alexander 2010 unless otherwise specified.

3 WGBA’s concern in the quoted passage is not only with *explicit* appeals to intuitions about cases.

- 4 WGBA provide references to several sorts of response to the experimental critique other than, although compatible with, the expertise response.
- 5 See for example Williamson 2005 in the case of knowledge ascriptions.
- 6 Nagel 2008 and 2010 constitute a promising recent example.