

Are Philosophers Good Intuition Predictors?*

Shen-yi Liao

Philosophical Psychology (forthcoming)[†]

Some philosophers have criticized experimental philosophy for being superfluous. Jackson (1998) implies that experimental philosophy studies are unnecessary. More recently, Dunaway and colleagues (2013) empirically demonstrates that experimental studies do not deliver surprising results, which is a pro tanto reason for foregoing conducting such studies.

This paper gives theoretical and empirical considerations against the superfluity criticism. The questions concerning the surprisingness of experimental philosophy studies have not been properly disambiguated, and their metaphilosophical significance have not been properly assessed. Once the most relevant question is identified, a re-analysis of Dunaway and colleagues' data actually undermines the superfluity criticism.

Most existing criticisms of experimental philosophy argue that its findings are *irrelevant* to philosophical practice.¹ However, there exists another, comparatively ignored criticism of experimental philosophy, which says that its findings are *superfluous*. Traces of the superfluity criticism can be found in the early crossfires about the value of experimental philosophy. For example, Frank Jackson (1998: 36) wryly remarks:

^{*} I am very grateful to Billy Dunaway, Anna Edmonds, and David Manley for openly sharing their datasets and granting me permission to post them on the Open Science Foundation repository, and for giving feedback during this paper's formative stage. For their valuable comments, I thank Joshua Knobe, Edouard Machery, Aaron Meskin, Jonathan Phillips, Sara Protasi, Nina Strohminger, Jonathan Weinberg, two anonymous referees, and audience members at the 2015 Experimental Philosophy UK Workshop, at the 2015 Society for Philosophy and Psychology meeting, at University of Leeds, and at University of Stirling. The work on this paper was supported by a European Community FP7 Marie Curie International Incoming Fellowship, grant PIIF-GA-2012-328977.

[†] PENULTIMATE VERSION. For citation and reference, please see the definitive and final version forthcoming in *Philosophical Psychology*.

¹ Different critics give different reasons for experimental philosophy's irrelevance. First, some have argued that intuitions—philosophers' and non-philosophers' alike—are simply irrelevant to philosophical arguments and philosophical practice (Weatherston 2003; Cappelen 2012). Second, some have argued that the questionnaire-based studies employed can fail to capture non-philosophers' intuitions due to the limitations of the method (Kauppinen 2007; Cullen 2010; Huebner 2015). Third, and most prominently, some have argued that philosophers are *expert intuiters* with respect to philosophically relevant concepts, and non-philosophers are not. That is, philosophers are better able to ignore irrelevant or biasing factors in thought experiments, fill in the underdescribed but theoretically relevant details, and so on; so that their intuitions better track the contours of philosophically relevant concepts. Buckwalter (in press) and Nado (2014b) provide comprehensive overviews of this debate about philosophers' alleged expertise in intuiting.

I am sometimes asked [...] why, if conceptual analysis is concerned to elucidate what governs our classificatory practice, don't I advocate doing serious public opinion polls on people's responses to various cases? My answer is that I do—when it is necessary.

Most recently, this criticism resurfaced in Billy Dunaway, Anna Edmonds, and David Manley's aptly titled paper "The Folk Probably Do Think What You Think They Think" (2013).

Dunaway and colleagues aim to empirically demonstrate that philosophers can reliably predict results of experimental philosophy studies that they have not heard of or about. If they succeed in their aim, an upshot of their study is that, given practical considerations such as cost and time, it is often unnecessary to conduct experimental philosophy studies. Like the irrelevance criticism, the superfluity criticism of experimental philosophy threatens the viability of the research program as a whole, since cost and time are highly salient factors in researchers' considerations of which projects are worth pursuing.

In this paper, I will set aside the irrelevance criticism of experimental philosophy, and address the superfluity criticism of experimental philosophy. My approach is partly theoretical and partly empirical. I first disambiguate two senses of the question "are philosophers good intuition predictors?". I then report an alternative analysis of Dunaway and colleagues' dataset to show that, although they are correct that there is one sense in which philosophers are good at predicting non-philosophers' intuitions, there is another sense in which their data gives more equivocal results. I then argue that, considering actual philosophical practice and professional philosophy institutions, it is this other sense in which philosophers are not clearly good intuition predictors that is more relevant to the superfluity criticism. Consequently, my re-analysis of Dunaway and colleagues' dataset undermines the superfluity criticism of experimental philosophy.

Of broader significance, I will show that thinking more deeply about the superfluity criticism and Dunaway and colleagues' study reveals an overlooked reason for experimental philosophy's relevance to philosophical practice: it allows for the operation of two epistemological institutions with distinctive epistemic virtues. I conclude with some methodological reflections that I hope can spur further progress in experimental philosophy and in metaphilosophy.

1. What Philosophers Think The Folk Think

The term "experimental philosophy" has no uncontroversial definition. On the broad conception, one that I personally sympathize with, experimental philosophy includes works that contribute to philosophical inquiry with empirical methods, done by both philosophers and researchers in cognate fields (Rose and Danks 2013). However, there is also a narrow conception of experimental philosophy that figures more prominently in metaphilosophical debates. On this narrow conception, experimental philosophy only consists in questionnaire-based studies that involve

soliciting non-philosophers' intuitions regarding philosophically relevant concepts. This paper will primarily discuss experimental philosophy on the narrow conception.

One way that experimental philosophy can turn out to be superfluous is that philosophers are *expert intuition predictors* with respect to philosophically relevant concepts.² Suppose that non-philosophers' intuitions are—contrary to the irrelevance criticism—relevant to philosophical practice, one might nevertheless argue that questionnaire-based studies of those intuitions are unnecessary because the studies' results can easily be predicted from the philosopher's armchair. Although traces of the superfluity criticism can be found in the metaphilosophical literature (for example, as noted earlier, in Jackson 1998), I will focus on what I regard as the best exemplar of it: Dunaway and colleagues' "The Folk Probably Do Think What You Think They Think" (2013).

Dunaway and colleagues start their paper by asking "How good are philosophers at predicting ordinary intuitions involving philosophically relevant concepts?" (Dunaway *et al.* 2013: 421). As they rightly note, their study is the only attempt to empirically answer this question thus far. I will explain their study's metaphilosophical significance in section 1.1, and their study's procedures and results in section 1.2.

1.1. *The Practical Upshot*

To be clear, Dunaway and colleagues themselves do not think that experimental philosophers should simply stop doing questionnaire-based studies that involve soliciting non-philosophers' intuitions regarding philosophically relevant concepts. Instead, they offer a more nuanced criticism: *if* the results of such studies can be reliably predicted by philosophers, *then* given limited resources (for example, with respect to cost and time), philosophers often do not need to conduct such questionnaire-based studies.³ As Dunaway and colleagues put the point,

A philosophical book might contain dozens of appeals to ordinary intuition, and the difficulty that would be involved in conducting careful experiments for each of these is enormous. Presumably we must make a trade-off between a higher risk of error on the one hand, and a drain of resources on the other. (437-438; cf. Jackson 1998; contra Stich and Weinberg 2001).

² Note that the philosophers-as-expert-intuition-predictors hypothesis makes no assumption about the relationship between philosophers' and non-philosophers' intuitions. For example, it is compatible with the possibility that philosophers actually systematically have intuitions different from non-philosophers, but nevertheless they are awesome at predicting non-philosophers' intuitions. For another example, it is also compatible with the possibility that philosophers are actually expert intuiters, but nevertheless they are awful at predicting non-philosophers' intuitions.

³ In the undercurrent of this argument is the thought that there is a negative correlation between a finding's obviousness and its importance. While many academic researchers consciously or unconsciously endorse this thought, F.D. Richard *et al.* (2001) found that laypeople do not. Specifically, they found that the social psychology findings that laypeople think to be most obvious—in the sense of being the most predictable—are also the ones that they found to be most important for researchers to investigate. In other words, laypeople in fact judge there to be a positive correlation between a finding's obviousness and its importance.

Let me state Dunaway and colleagues' position even more carefully.⁴ On their view, there are many factors that determine whether a study should be conducted. Among them, a study's predictability is one *pro tanto* practical reason against conducting a study. That is the sense in which many prominent studies in experimental philosophy, on the narrow conception, turn out to be superfluous.

1.2. *The Empirical Pudding*

So, how did Dunaway and colleagues show that philosophers are good predictors of non-philosophers' intuitions with respect to philosophically relevant concepts? To start, they selected four experimental philosophy studies on different topics that have been explicitly claimed as giving "surprising" results. (The details of these studies are given in Table 1). Then, they sent out an online questionnaire to faculty members and graduate students in top "Leiter-ranked" English speaking philosophy departments.⁵ These questionnaire respondents were clearly and repeatedly instructed and asked to not make a prediction if they've heard of or about the relevant experimental philosophy study.⁶ Finally, respondents were given the relevant vignettes of the four allegedly "surprising" experimental philosophy studies and asked to predict the results.

Here is an example that involves the vignettes and questions from one study reported in Knobe 2003. Respondents were presented with the two vignettes and questions used in the study in counterbalanced order:

Harm

The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.'

The chairman of the board answered, 'I don't care at all about harming the environment. I just want to make as much profit as I can. Let's start the new program.'

They started the new program. Sure enough, the environment was harmed.

⁴ I thank Billy Dunaway, Anna Edmonds, and David Manley for helping me clarify their position.

⁵ There were two runs with only a small difference in the wording of the prompt. Since the two runs produced the same pattern of results, I will talk about the two runs collectively. The merged dataset contains responses from a total of 331 respondents. See Dunaway *et al.* 2013, notes 11 and 27, for further details on respondent recruitment. In addition to their paper, further details of the studies can be found at <http://web.archive.org/web/20150316124156/http://www-personal.umich.edu/~dunaway/MetaSurvey.html>.

⁶ When I have presented this paper, experimental philosophers sometimes raised worries about technical details of Dunaway and colleagues' study, such as worries about whether they successfully ensured that respondents honestly reported that they were unable to provide an unbiased answer. I am simply not interested in such worries in this paper. That is, I am assuming that Dunaway and colleagues have conducted their study perfectly, and I am only interested in providing an alternative analysis of their data and drawing a different metaphilosophical conclusion on the basis of that alternative analysis.

(1) Did the chairman of the board *intentionally* harm the environment?

and

Help

The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, and it will also help the environment.'

The chairman of the board answered, 'I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the new program.'

They started the new program. Sure enough, the environment was helped.

(2) Did the chairman of the board *intentionally* help the environment?

Then, the respondents were asked to select one of the four following options, presented in counterbalanced order:

On average, subjects asked question (1) would respond "yes"

- significantly more often than subjects asked question 2 would respond 'yes'
- not significantly more or less often than subjects asked question 2 would respond 'yes'.
- significantly less often than subjects asked question 2 would respond 'yes'
- (Unable to provide an unbiased answer.)

As Dunaway and colleagues expected, an overwhelming majority of respondents (79.1% in both runs) correctly predicted the result of this study. The same pattern holds for the other studies that Dunaway and colleagues tested, as Table 1 shows:

	Study 1	Study 2	Study 3	Study 4
Study Article	Knobe and Fraser (2008)	Knobe (2003)	Livengood and Machery (2007)	Nichols and Knobe (2007)
Study Topic	Moral Judgment & Causation	Intentional Action	Causation by Omission	Determinism & Moral Responsibility
% of Correct Prediction	93.4%	79.1%	84.1%	85.5%

Table 1. Summary of test cases and results from Dunaway *et al.* (2013)

In *four out of four* studies tested, philosophers correctly predicted the results of these studies. Contrary to what experimental philosophers claim, these studies are not “surprising”—at least not in the sense that their results cannot be predicted from the armchair. On Dunaway and colleagues’ analysis, philosophers are rather impressive at predicting non-philosophers’ intuitions with respect to philosophically relevant concepts, such as INTENTION and CAUSATION. In turn, for reasons discussed in section 1.1, these results suggest that experimental philosophy, on the narrow conception, is superfluous.

2. What Philosophers Really Think The Folk Think

2.1. Question Disambiguation

Let us return to the question that Dunaway and colleagues ask at the start of their paper. That question, “how good are philosophers at predicting ordinary intuitions involving philosophically relevant concepts?”, turns out to be ambiguous between two questions:

[Collectivist] How good are philosophers, *as a collective*, at predicting ordinary intuitions involving philosophically relevant concepts?

versus

[Individualist] How good is *each individual* philosopher at predicting ordinary intuitions involving philosophically relevant concepts?

Dunaway and colleagues’ own analysis only addresses the collectivist question. So, in sections 2.2 and 2.3, I re-analyzed their data to address the individualist question. And, I shall argue in section 2.4, it is the answer to the individualist question that is more relevant for metaphilosophical debates about the value of experimental philosophy.

2.2. Individualist Re-Analysis

I re-examined Dunaway and colleagues’ data in order to answer the question, “How good is each individual philosopher at predicting ordinary intuitions involving philosophically relevant concepts?”⁷ First, I recoded the predictions as either correct (numerical value 1) or incorrect (0).⁸ Second, I calculated the percentage of correct

⁷ I am grateful to Billy Dunaway, Anna Edmonds, and David Manley for graciously sharing their data and discussing the re-analysis. The recoded data is available on the Open Science Foundation repository at <https://osf.io/n5zsw/>.

⁸ The coding for Study 1 and Study 2 are straightforward. For Study 3, only the Unsafe Rope case responses were used because, as Dunaway and colleagues acknowledge, the Safe Rope case is only a control (Dunaway *et al.* 2013:

predictions for each individual respondent. Third, I excluded respondents who only made two or fewer predictions; 237 qualified respondents remain in the dataset for subsequent analyses.⁹

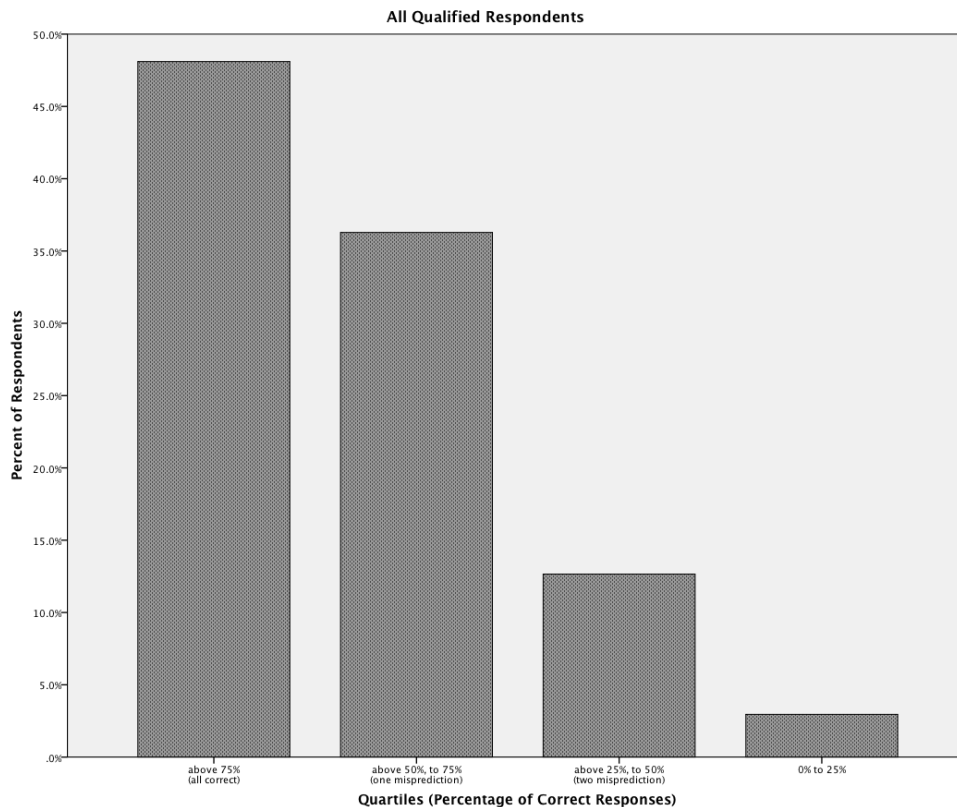


Figure 2. Quartiles of respondents who made correct intuition predictions.

Figure 2 provides a visualization of the results for all qualified respondents. One headline from these results is that the majority of individual philosophers made at least one incorrect intuition prediction. Only 48.1% of the respondents got all of their intuition predictions correct. Of the rest, 36.3% of respondents got one intuition prediction wrong, 12.7% got two intuition predictions wrong, and 3.0% got all intuition predictions wrong. Another headline from these results is that the median philosopher gets roughly one out of five intuition predictions wrong.

Given these headlines, on the individualist re-analysis, it is not clear that philosophers have an expertise in intuition prediction. An appropriate comparison

433). For Study 4, which involved asymmetric free will intuitions regarding abstract and concrete universes, a response is coded as correct if and only if it is a correct prediction of the asymmetry—in other words, if and only if it predicts free will intuitions regarding both universes.

⁹ This exclusion criterion was determined prior to any analysis. The rationale is that including the percentage of correct predictions for respondents who only answered two or fewer questions would inappropriately skew the results.

class is needed to truly assess whether philosophers are particularly good at intuition prediction. While I will not try to delineate such a comparison class in this paper, I want to note one clearly inappropriate comparison class: random guesses.

Contrary to stereotypes, professional philosophers still possess theory of mind to some degree, and they should be antecedently expected to make decent guesses about non-philosophers' intuitions. In research on the predictability of psychological research results, studies typically find that naïve participants do better than chance. (See Richard *et al.* 2001: 498–499 for an overview of extant research on this topic.) Better-than-chance performance is thus unlikely to be indicative of any specific capacity, such as a capacity for predicting non-philosophers' philosophical intuitions. Instead, such performance is more likely to be only indicative of a general capacity for predicting others' thoughts—that is, a general capacity for mindreading. Showing that philosophers are good at intuition prediction requires more than showing that they do better than random guesses.

2.3. Further Exploratory Analyses for Contextualization

I then conducted further exploratory analyses to contextualize the individualist re-analysis results.

In assessing the superfluity of experimental philosophy, one comparison we can make with this dataset is between respondents who know some experimental philosophy and respondents who do not. Remember that Dunaway and colleagues made it an option for respondents to decline to make a prediction on the ground that they have heard of or about the relevant experimental philosophy study. Using this information, we can explore the following conjecture: if one has heard of or about the results of one experimental philosophy study, which investigates the outcome of some cognitive processes, one can do better at predicting the results of another experimental philosophy study that investigates partially overlapping cognitive processes.¹⁰ To explore this conjecture, I divided respondents into two groups—philosophers who are wholly innocent of experimental philosophy (*innocents*), and philosophers who are not (*non-innocents*)—and compared their intuition prediction performances.¹¹

Figure 3 provides a visualization of the results for just the innocents. One headline from these results is that an even greater majority of individual

¹⁰ Whether experimental philosophy studies on diverse topics involve overlapping cognitive processes is controversial. On the one hand, there are attempts at unifying disparate findings. For example, Phillips *et al.* (2015) found that relevance of alternative possibilities is a factor that moderates moral judgments' influence on intuitions about freedom, causation, the doing / allowing distinction, and intentional action. On the other hand, there are good reasons to think that "intuition" refers to a highly heterogeneous category that involves many unrelated cognitive processes (Nado 2014a). This conjecture only makes a weak assumption, that there is some overlap in the studies that Dunaway and colleagues examined.

¹¹ I want to stress that these analyses are only exploratory, especially given that there exists a potential confound in grouping the respondents thusly. The potential confound exists because innocents typically made more intuition predictions than the non-innocents, and so they typically have more opportunities to get an intuition prediction wrong. However, given the way that Dunaway and colleagues recruited their respondents, not all respondents—and so, not all innocents—answered all the questions. The heterogeneity within each group makes any comparison between the two inexact. I thank Edouard Machery and David Manley for discussion about this potential confound.

philosophers made at least one incorrect intuition prediction. Only 41.7% of the respondents got all of their intuition predictions correct. Of the rest, 39.3% of respondents got one intuition prediction wrong, 14.9% got two intuition predictions wrong, and 4.2% got all intuition predictions wrong. Another headline from these results is that the median philosopher gets roughly one out of four intuition predictions wrong.

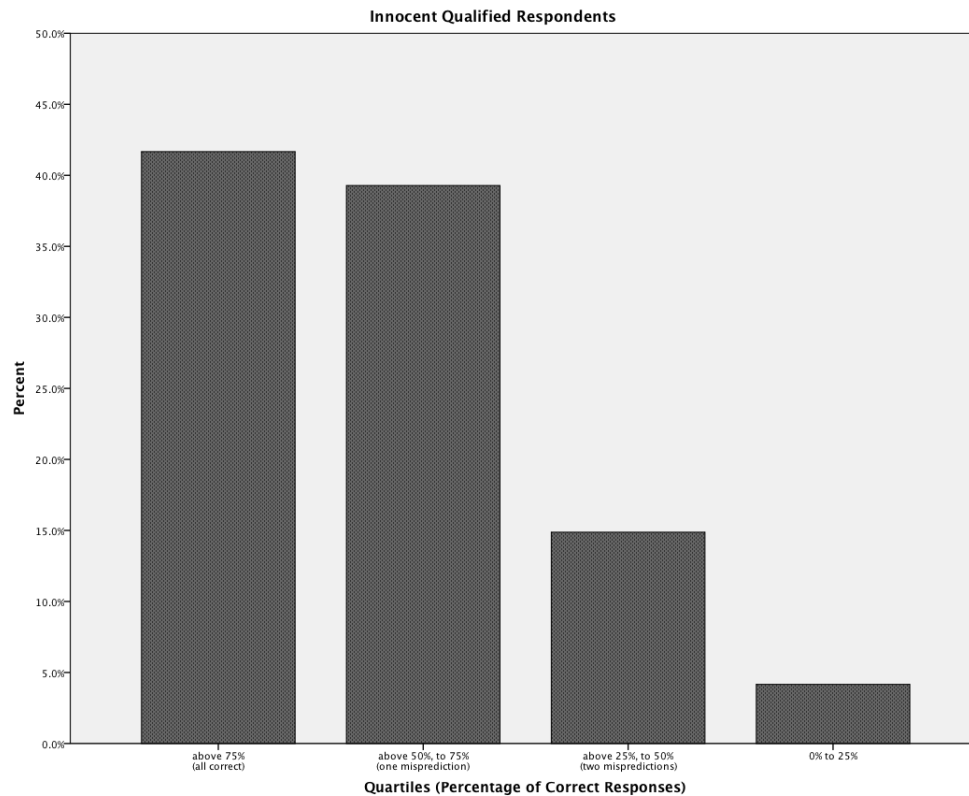


Figure 3. Quartiles of innocent respondents who made correct intuition predictions.

There exist differences between the intuition prediction performances of the innocents and the non-innocents. The non-innocents ($M = 0.855$, $SD = 0.210$) have a higher percentage of correct intuition predictions than the innocents ($M = 0.788$, $SD = 0.214$), $t(128.926) = 2.216$, $p = 0.028$, effect size $r = 0.192$. In particular, the non-innocents (63.8%) are more likely than the innocents (41.7%) to get all intuition predictions correct, $\chi^2(1) = 9.571$, $p = 0.002$, effect size *Cramer's V* = 0.201.

These differences suggest that a philosopher who has heard of or about at least one of the experimental philosophy studies that Dunaway and colleagues investigated can better predict results of the other experimental philosophy studies investigated than a philosopher who has not heard of or about any of the experimental philosophy studies investigated. Prior knowledge of experimental philosophy results seems to make a small but noticeable contribution to

philosophers' intuition prediction performances. Insofar as philosophers possess some expertise in intuition prediction, then, part of that expertise seems to be the product of (perhaps unconsciously) assimilating existing experimental philosophy findings.¹² These results therefore point to at least one utility of experimental philosophy studies: giving philosophers a better sense of folk responses to other experimental philosophy studies that exploit partially overlapping cognitive processes.

These are, to repeat, only exploratory analyses aiming to contextualize the individualist re-analysis results. While these results provide some support for the conjecture that prior knowledge of experimental philosophy makes a difference to intuition prediction success, the difference is small by conventional standards.¹³ As such, we should be cautious about drawing definitive conclusions from these results and more empirical explorations of the conjecture are welcomed.

2.4. *Metaphilosophical Ramifications*

On Dunaway and colleagues' analysis, philosophers *as a collective* are rather impressive at predicting non-philosophers' intuitions: the collective correctly predicted the results of every single study tested. However, on my analysis, the median *individual* philosopher gets roughly one out of five intuition predictions wrong, and the median *individual* innocent philosopher gets roughly one out of four intuition predictions wrong.

This discrepancy between the collective's and the individuals' intuition prediction performances turns out to be metaphilosophically significant when we consider actual philosophical practice. Insofar as philosophers utilize their intuition predictions in theory construction, there is likely to be a kind of second-order opacity. First, each individual philosopher is unlikely to know whether they are a good, mediocre, or bad intuition predictor. After all, respondents in Dunaway and colleagues' study clearly varied in their intuition prediction performances. Second, each individual philosopher is unlikely to know in which case their intuition prediction is incorrect. After all, different respondents in Dunaway and colleagues' study got different intuition predictions wrong.

What explains the discrepancy between the collective's and the individuals' intuition prediction performances? My answer is that the impressiveness of the collective's performance is actually an artifact of another factor: wisdom of the crowd. Roughly, *wisdom of the crowd* refers to an effect in which an aggregation of individual judgments produces a judgment that is epistemic superior to any of the individuals'. The basic idea is that aggregation makes noises from the individual

¹² This is an instance of the wider phenomenon of *hindsight bias*, in which people are generally so good at integrating new evidence into their existing frameworks that it can feel to them as if they knew the evidence all along. For an accessible introduction to the phenomenon and its many instances, see the aptly titled book *Everything is Obvious: *Once You Know the Answer* (Watts 2011).

¹³ However, this small effect is arguably expected in this research context. As section 2.2 suggests, it is plausible that a general shared capacity, such as the capacity for mindreading, explains a large part of philosophers' intuition prediction performance. If so, then we should expect prior knowledge to make only a small difference on top of that.

variations disappear so that only the truth-tracking signals in the individual judgments remain.

While the basic idea behind wisdom of the crowd is old, going at least as far back as Galton 1907, in recent years economists and other social scientists have been able to demonstrate its power in a variety of domains (cf. Surowiecki 2004). Furthermore, social scientists have identified two key conditions that are conducive to wisdom of the crowd:

- (a) INDEPENDENCE. The responses from the crowd are independent from one another.
- (b) DIVERSITY. The crowd consists in cognitively diverse individuals.

When these conditions hold, wisdom of the crowd makes individual variations matter far less. As Dunaway and colleagues' original analysis shows (Table 1), by aggregating intuition predictions from different respondents and thereby washing out the noisy individual differences, philosophers *as a collective* correctly predicted non-philosophers' intuitions in all four studies tested, even if few *individual* philosophers did the same. And the way that Dunaway and colleagues conducted their study is relatively conducive to wisdom of the crowd: while the cognitive diversity of the respondent pool in Dunaway and colleagues' study is somewhat questionable, the responses were in fact independent from one another—especially for the innocents. On my diagnosis, wisdom of the crowd is what turns an equivocal answer to the individualist question into an affirmative answer to the collectivist question.

3. Thinking About Structures For Thinking

3.1. Professional Institutions and Practices

Real world professional philosophy discussions do not (yet!) typically advance via questionnaires. In some ways, professional institutions and practices seem to facilitate wisdom of the crowd (cf. Jackson 1998). For example, philosophers can take into account the diversity of intuitions by presenting their work at conferences or talking about cases with colleagues and students.¹⁴

However, in many other ways, professional institutions and practices seem to actually work against wisdom of the crowd by undermining INDEPENDENCE and DIVERSITY.¹⁵ Graduate school training and the journal refereeing process both involve intuition inculcation to some degree. When influential philosophers get their intuitions into print, it is often very hard to get it out of the philosophical literature,

¹⁴ I thank Joshua Knobe for reminding me to also take note of these positive aspects of professional institutions and practices.

¹⁵ As an anonymous referee notes, another way in which DIVERSITY may be undermined by professional institutions and practices is via the lack of demographic identity diversity. I welcome this suggestion. As an example of this mechanism, Dotson (2012) argues that the pervasive culture of justification embedded in professional philosophy excludes many diverse practitioners, who must continually defend their work against the charge that it is not real philosophy.

even when many others do not share those intuitions. For example, as Robert Cummins (1998: 116) laments, “The Putnamian take on [Twin Earth] cases is widely enough shared to allow for a range of thriving intramural sports among believers. Those who do not share the intuitions are simply not invited to the games”. In the same spirit, John Turri (in press) points to several *thought experimenter-biases*, analogous to experimenter biases in psychology, which takes place in professional philosophy institutions. As an example, Turri reports that his paper questioning the original intuition to Gettier’s cases received many dismissive and hostile referee reports in the journal refereeing process. Given the existence of such attitudes in the journal refereeing process, it is easy to imagine their existence in graduate school training, perhaps to an even greater degree because the power differential in that context is greater.

When INDEPENDENCE and DIVERSITY hold to a lesser degree, the kind of second-order opacity highlighted in section 2.4 becomes increasingly worrisome. Consider one motivation behind the superfluity criticism: if philosophers were good at predicting intuitions, then even if they were not good at intuiting, there would still be no need for experimental philosophy. The earlier results and discussions undercut this motivation. Given that individual philosophers typically do not know whether they are good at intuition prediction and they typically do not know in which cases they make the incorrect predictions, even if these individual philosophers use intuition predictions instead of intuitions as data for their theorizing, it remains likely that misleading data will get into the literature and distort subsequent discussions.

3.2. Epistemological Institutions of Experimental Philosophy

Experimental philosophers can avoid these pitfalls by operating two kinds of epistemological institutions.¹⁶ First, they can work within a community of theorists who can work together to construct thought experiments that track theoretically relevant factors.¹⁷ Second, they can operate a separate community of study participants, who are more cognitively diverse than the community of professional philosophers and can give responses independently, that takes advantage of wisdom of the crowd.¹⁸

What Dunaway and colleagues have shown, in my view, is that wisdom of the crowd is useful for studying intuitions regarding philosophically relevant concepts, regardless of whether the intuitions come from philosophers or non-philosophers. However, experimental philosophers can already take advantage of wisdom of the crowd by giving questionnaires to a cognitively diverse and independent participant pool. Moreover, experimental philosophers can also rely on

¹⁶ I thank Joshua Knobe for suggesting this idea.

¹⁷ In the same spirit, Helen De Cruz (2015) argues that philosophers’ expertise consists in eliciting intuitions in the dialectical context of professional philosophy.

¹⁸ One intriguing possibility that Dunaway and colleagues’ study suggests is that experimental philosophers should consider asking non-philosophers for their intuition predictions (or, perhaps, intuition expectations) instead of their intuitions.

what is arguably the real expertise of philosophers qua theorists: their ability to incorporate a variety of evidence and construct elegant and testable theories that account for them. In turn, being in a position to take advantage of these two distinct epistemological institutions explains why experimental philosophy, especially on the broad conception, is not only not superfluous, but also not irrelevant.

3.3. Methodological Reflections for Experimental Philosophers and Metaphilosophers

Social science is messy. Depending on analysis and interpretation, the same set of data can tell different stories. It is crucial to scientific progress that researchers can have access to data that they did not collect so that a dataset can be re-examined via the lenses of different theories and auxiliary hypotheses, and that a theory can be continually re-assessed with growing evidence.

In this case, I was in a fortunate position to re-analyze and re-interpret this dataset because Dunaway and colleagues generously shared it when I asked. However, an even more ideal scenario would be one where I did not even have to ask; where the dataset would be already available in a public repository for anyone to re-examine. In addition to the first-order results, I hope my re-analysis has shown the value of opening up one's research results for facilitating collaborative and cumulative research. Going forward, I hope it will be a norm for experimental philosophy researchers to make their data and material available on repositories such as Open Science Foundation, so that further progress can be made collaboratively and cumulatively. (If every social movement needs a hashtag to propagate, I recommend "#openxphi" for this one.)

The availability of open data can also encourage progress in metaphilosophical debates concerning the viability of experimental philosophy. Too often, criticisms are levied against specific experimental philosophy studies on armchair speculations alone.¹⁹ Against this trend, it is encouraging to see recent criticisms of experimental philosophy studies via replications (for example, Adleberg *et al.* 2015; Kim and Yuan 2015; Seyedsayamdost 2015).²⁰ My suggestion here is that proponents and critics alike can add non-experimental tools like re-analysis and meta-analysis (for example, Feltz and Cova 2014) to their toolbox, in addition to direct and conceptual replications. It is my hope that open science, and the variety of resources it offers, can give us more opportunities to make progress in metaphilosophical debates, as well as in first-order debates.

¹⁹ Huebner (2015) is a notable exception. He examines statistical properties of particular studies to illustrate and substantiate his general methodological worries concerning the use of Likert scales in experimental philosophy research.

²⁰ See also the invaluable resource "Replications in Experimental Philosophy", currently maintained by Joshua Knobe and Christian Mott, and available at <http://pantheon.yale.edu/~jk762/xhipage/Experimental%20Philosophy-Replications.html>.

References

- Adleberg, Toni, Morgan Thompson, & Eddy Nahmias (2015). "Do Men and Women Have Different Philosophical Intuitions? Further Data". *Philosophical Psychology* 28(5): 615–641.
- Buckwalter, Wesley (in press). "Intuition Fail: Philosophical Activity and the Limits of Expertise". *Philosophy and Phenomenological Research*.
- Cappelen, Herman (2012). *Philosophy Without Intuitions*. New York: Oxford University Press.
- Cullen, Simon (2010). "Survey-Driven Romanticism", *Review of Philosophy and Psychology* 1(2):275–296.
- Cummins, Robert C. (1998). "Reflection on Reflective Equilibrium". In Michael R. DePaul & William Ramsey (eds.), *Rethinking Intuition*. Lanham, MD: Rowman & Littlefield, 113–128.
- De Cruz, Helen (2015). "Where Philosophical Intuitions Come From". *Australasian Journal of Philosophy* 93(2): 233–249.
- Dotson, Kristie (2012). "How Is This Paper Philosophy?". *Comparative Philosophy* 3(1): 3–29.
- Dunaway, Billy, Anna Edmonds, & David Manley (2013). "The Folk Probably Do Think What you Think They Think". *Australasian Journal of Philosophy* 91(3): 421–441.
- Feltz, Adam, & Florian Cova (2014). "Moral Responsibility and Free Will: A Meta-analysis". *Consciousness and Cognition* 30: 234–246.
- Galton, Francis (1907). "Vox Populi". *Nature* 75: 450–451.
- Huebner, Bryce (2015). "What is a Philosophical Effect? Models of Data in Experimental Philosophy". *Philosophical Studies* 172(12): 3273–3292.
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. New York: Oxford University Press.
- Kauppinen, Antti (2007). "The Rise and Fall of Experimental Philosophy". *Philosophical Explorations* 10(2): 95–118.
- Kim, Minsun & Yuan Yuan (2015). "No Cross-Cultural Differences in the Gettier Car Case Intuition: A Replication Study of Weinberg *et al.* 2001". *Episteme* 12(3): 355–361.
- Knobe, Joshua (2003). "Intentional Action and Side-Effects in Ordinary Language". *Analysis* 63(3): 190–194.
- Knobe, Joshua, & Ben Fraser (2008). "Causal Judgment and Moral Judgment: Two Experiments", in Walter Sinnott-Armstrong (ed.), *Moral Psychology*. Cambridge, MA: MIT Press: 441–448.
- Livengood, Jonathan, & Edouard Machery (2007). "The Folk Probably Don't Think What You Think They Think: Experiments on Causation by Absence". *Midwest Studies in Philosophy* 31(1): 107–127.
- Nado, Jennifer (2014a). "Why Intuition?". *Philosophy and Phenomenological Research* 89(1):15-41.
- Nado, Jennifer (2014b). "Philosophical Expertise". *Philosophy Compass* 9(9): 631–641.

- Nichols, Shaun, & Joshua Knobe (2007). "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions". *Nous* 41(4): 663–685.
- Nisbett, Richard E., & Wilson, Timothy DeCamp (1977). "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review* 84(3): 231–259.
- Phillips, Jonathan, Jamie B. Luguri, & Joshua Knobe (2015). "Unifying Morality's Influence on Non-moral Judgments: The Relevance of Alternative Possibilities". *Cognition* 145: 30–42.
- Richard, F.D., Charles F. Bond Jr., & Juli J. Stokes-Zoota (2001). "'That's Completely Obvious... and Important': Lay Judgments of Social Psychological Findings". *Personality and Social Psychology Bulletin* 27: 497–505.
- Rose, David, & David Danks (2013). "In Defense of a Broad Conception of Experimental Philosophy". *Metaphilosophy* 44(4): 512–532.
- Seyedsayamdost, Hamid (2015). "On Gender and Philosophical Intuition: Failure of Replication and Other Negative Results". *Philosophical Psychology* 28(5): 642–673.
- Stich, Stephen, & Jonathan Weinberg (2001). "Jackson's Empirical Assumptions". *Philosophy and Phenomenological Research* 62(3): 637–643.
- Surowiecki, James (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York: Doubleday.
- Turri, John (in press). "Knowledge Judgments in 'Gettier' Cases". In Justin Sytsma & Wesley Buckwalter (eds.), *A Companion to Experimental Philosophy*. Malden: Blackwell.
- Watts, Duncan J. (2011). *Everything is Obvious: *Once You Know the Answer*. New York: Crown.
- Weatherson, Brian (2003). "What Good Are Counterexamples?". *Philosophical Studies* 115(1): 1–31.