Consciousness and Cognition 20 (2011) 1722–1731

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Short Communication

Persistent bias in expert judgments about free will and moral responsibility: A test of the expertise defense

Eric Schulz [a,1], Edward T. Cokely [a,b,*,1], Adam Feltz [c,1]

[a] Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Germany
[b] Department of Cognitive and Learning Sciences, Michigan Technological University, United States
[c] Departments of Philosophy and Interdisciplinary Studies, Schreiner University, United States

ARTICLE INFO

Article history:
Received 2 August 2010
Available online 18 May 2011

Keywords:
Experimental philosophy
Personality
Free will
Compatibilism
Incompatibilism
Expertise
Intuition
Judgment bias
Philosophical skill

ABSTRACT

Many philosophers appeal to intuitions to support some philosophical views. However, there is reason to be concerned about this practice as scientific evidence has documented systematic bias in philosophically relevant intuitions as a function of seemingly irrelevant features (e.g., personality). One popular defense used to insulate philosophers from these concerns holds that philosophical expertise eliminates the influence of these extraneous factors. Here, we test this assumption. We present data suggesting that verifiable philosophical expertise in the free will debate—as measured by a reliable and validated test of expert knowledge—does not eliminate the influence of one important extraneous feature (i.e., the heritable personality trait extraversion) on judgments concerning freedom and moral responsibility. These results suggest that, in at least some important cases, the expertise defense fails. Implications for the practice of philosophy, experimental philosophy, and applied ethics are discussed.

1. Introduction

Philosophy endeavors to provide a rational investigation of things such as knowledge, free action, and ethics. In developing arguments and theories, many philosophers appeal to their own intuitions or to those intuitions that are widely held (e.g., Jackson, 1998; McCann, 2005; Smith, 1994; Sosa, 2007b; Stanley, 2005). However, there is reason to be concerned about this practice as scientific evidence has documented systematic diversity in philosophically relevant intuitions as a function of seemingly extraneous factors including culture (Haidt, Koller, & Dias, 1993; Machery, Mallon, Nichols, & Stich, 2004; Weinberg, Nichols, & Stich, 2001), development (Nichols, 2004; Nichols & Folds-Bennett, 2003), and heritable personality traits (Cokely & Feltz, 2009a, 2009b; Feltz & Cokely, 2007, 2009).

One popular defense used to insulate philosophers from these concerns is the *expertise defense*. This defense holds that if one has expert knowledge about some philosophical issue (e.g., the free will debate) the influences of extraneous features (e.g., personality) would not matter or would be dramatically reduced (Kauppinen, 2007; Ludwig, 2007; Sosa, 2007a; Williamson, 2007, in press). In what follows, we contribute to the emerging body of empirical evidence investigating the truth of these claims. We present data from the first study that assesses objectively verifiable expert knowledge (e.g., philosophical

* Corresponding author at: Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.
E-mail address: cokely@mpib-berlin.mpg.de (E.T. Cokely).
[1] All authors contributed equally.

1053-8100/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.concog.2011.04.007

knowledge about the free will debate), a key personality trait (i.e., extraversion), and their relation to paradigmatic judgments about free will and moral responsibility.[2]

1.1. Personality, intuitions, and the free will debate

Many theorists claim that belief in free will and moral responsibility forms a cornerstone of what it is to be human. Beliefs in free will and moral responsibility have been argued to underwrite our notions of punishment, blame, and a host of other attitudes about how we relate to ourselves and others (Cover & O'Leary-Hawthorne, 1996; Kane, 1996, 1999). Not surprisingly, many theorists hold that most people believe in free will and moral responsibility. Some even argue that if we find that people are not free and morally responsible, we should leave people to their mistaken beliefs (Smilansky, 2002).

In the contemporary free will debate, one hotly contested issue is whether free will and moral responsibility are compatible with the truth of determinism (compatibilism) or not (incompatibilism). Adherents of both camps have complex, nuanced arguments. In addition to these arguments, one strategy theorists have used to help augment their position is to appeal to "folk"[3] intuitions about the relation of determinism with free will and moral responsibility (e.g., Dennett, 1984; Ekstrom, 2002; Kane, 1999; Lycan, 2004; Pink, 2004; Strawson, 1986; Wolf, 1990). This strategy is important because those views that are consistent with folk intuitions are thought to have prima facie plausibility that the views that are inconsistent with folk intuitions do not. For example, those who have views that are consistent with folk intuitions about free will and moral responsibility do not have to explain away folk intuitions. However, those views that are contrary to folk intuitions shoulder some burden to explain why folk intuitions about freedom and moral responsibility are wrong. For these reasons, views supported by folk intuitions have been claimed to have "squatter's rights" that views clashing with folk intuitions do not (Dennett, 1984; Nahmias, Morris, Nadelhoffer, & Turner, 2006).

Even though incompatibilism is thought by many theorists to be supported by folk intuitions, "in contemporary discussions of free will, incompatibilists self-identify as the underdog" (Nichols, 2007, p. 261). That is, incompatibilism may not be the dominant position among those who are knowledgeable about free will. For example, Robert Kane (1996) writes "Compatibilism has surely become the dominant view among philosophers today" (p. 12). As Derk Pereboom notes, "the demographic profile of the free will debate reveals a majority of soft determinists, who claim that we possess the freedom for moral responsibility, that determinism is true, and these views are compatible" (1995, p. 21). In contrast, a more recent discussion of the free will landscape indicates that many professionals specializing in free will think that incompatibilism has become the dominant view.[4] However, there is very little empirical evidence concerning whether compatibilism or incompatibilism is the dominant view among those who are knowledgeable about the free will.[5]

What empirical evidence there is concerning intuitions about free will draws largely on free will intuitions of the folk (Nahmias, Morris, Nadelhoffer, & Turner, 2004, 2005; Nahmias et al., 2006; Nichols & Knobe, 2007). To date, in most studies concerning folk intuitions about free will, there has been a substantial dissenting minority. Rather than reflecting random or unpredictable variation, emerging evidence suggests that some of the diversity in intuitions about free will is systematically related to a heritable personality trait (i.e., personality is associated with specific biases or judgment tendencies). Specifically, in some paradigmatic cases the global personality trait *extraversion* predicts compatibilist intuitions (Cokely & Feltz, 2009b; Feltz & Cokely, 2009; Nadelhoffer, Kvaran, & Nahmias, 2009). Because the different sets of intuitions can be predicted by personality and are associated with reliable bias in some representative circumstances, it lends credibility to the idea that those who are in the minority are not merely expressing error (note: bias refers to a judgment tendency which may be right or wrong, beneficial or harmful depending on environmental context). For these reasons, it appears likely that there is not one but several distinct sets of folk intuitions about free will and moral responsibility. The existence of several distinct sets of intuitions about free will and moral responsibility directly challenges the notion that anybody in the free will debate has "squatter's rights." The existence of stable and systematic compatibilist and incompatibilist intuitions requires *both* compatibilists and incompatibilists to explain away contrary intuitions.

Another major worry is that the general biases associated with various personality traits may also be present in skilled philosophers. Because philosophers are also humans with personality, personality could be related to or influence philosophers' intuitions about free will and moral responsibility. Many philosophers would want to deny this striking suggestion because then extraneous features would factor in their study of free will and moral responsibility.[6] Accordingly, some philosophers rely on a strategy to defend themselves from such concerns—i.e., the expertise defense.

1.1.1. The expertise defense

The expertise defense is a common and important defense against concerns about potential influences of extraneous features on philosophers' intuitions (Kauppinen, 2007; Ludwig, 2007; Sosa, 2007a; Williamson, 2007, in press). Generally, the

[2] In this paper, we use 'intuitions' and 'judgment' interchangeably. Some theorists maintain there is a distinction between these two things, whereas others do not. For our purposes, we take it that there is no important difference. For a more detailed discussion, see Feltz and Bishop (2010).

[3] The "folk" are generally thought to be individuals who have no detailed knowledge of the theoretical debate being investigated.

[4] For a discussion of this, see http://gfp.typepad.com/the_garden_of_forking_pat/2005/02/how_many_battal.html.

[5] Nichols (2007) offers some psychological reasons for the rise of compatibilism. Namely, that it is motivationally attractive to those who think determinism is true.

[6] Whether personality is an extaneious factor depends on what project one uses intuitions for (Feltz & Cokely, in press-a; Feltz & Cokely, in press-b; Stich, 2010). See below for a more detailed discussion of the projects one could use intuitions for and when personality is likely to be an extraneous factor.

expertise defense holds that "the folk" are the wrong class of people to ask about philosophical issues. The folk do not have the required sophistication with the concepts involved and they may not be able to reason with the same depth or precision as professional philosophers or others with high levels of philosophical skill (e.g., knowledge about debates concerning free will and moral responsibility). Accordingly, the appropriate class of people to ask about the relation of determinism to free will and moral responsibility are philosophers who have sufficient expert knowledge about free will and moral responsibility.[7] Because of philosophers' richer understandings of the debate and concepts involved, it is argued that the intuitions and biases found among the folk will not hold for skilled individuals. Hence, it would be illegitimate to infer that just because heritable personality traits can predict biases in the folk that they will likewise predict biases among skilled experts. Expertise should make personality-related biases disappear. Indeed, there is considerable evidence suggesting that verifiable expert performers—i.e., individuals who show superior, reproducible performance on representative tasks (e.g., chess masters playing chess)—can show dramatic, qualitative differences in reasoning, judgment abilities, overcoming biases common to novices (Ericsson & Lehman, 1996; Ericsson, Prietula, & Cokely, 2007).

Nevertheless, the expertise defense has a critical empirical premise (Alexander & Weinberg, 2007; Weinberg, 2007)—that as a matter of fact, philosophical experts *do not* display the same or similar patterns of intuitions associated with global personality traits as the folk. Some have tried *indirect strategies* to argue that it is not likely that the expertise defense will succeed in many philosophical domains. These strategies do not attempt to show directly that philosophical experts' intuitions vary in ways similar to the folk. Rather, these strategies attempt to show the ways in which expertise functions in philosophy is sufficiently dissimilar to the ways expertise functions in other domains. Because of the dissimilarity, we should not have much confidence that expertise in philosophy will dramatically reduce or eliminate the effects of extraneous factors. Weinberg, Gonnerman, Buckner, and Alexander (2010) give some detailed arguments that the ways philosophers could be experts do not neatly map onto known ways that expertise makes intuitions qualitatively better in other domains. To take one of their examples, the type of feedback offered in philosophy is often not the same as the feedback offered in other disciplines. In chess, it is fairly clear and immediately known when one makes an error. In philosophy, it is often not clear or immediately known when one makes an error (e.g., how long did it take for the Justified True Belief (JTB) account of knowledge to be thought wrong? Is it clear to everybody the JTB account is wrong?). Hence, the feedback in the two domains is sufficiently different to suggest that philosophers may not have the right kind of expertise (or learning environments) to insulate their intuitions from the effects of extraneous factors like personality.

Here, we adopt a *direct strategy* to question the expertise defense. Direct strategies provide evidence that experts display the same or similar types of effects as the folk. There is gathering evidence that philosophical experts sometimes behave in much the same way as the folk. For example, ethicists sometimes do not behave any better than non-ethicists (Schwitzgebel, 2009; Schwitzgebel & Rust, 2010a, 2010b). But even more direct, there is some evidence that extraneous influences such as the order in which scenarios are presented often influences philosophical experts in similar ways as in the folk (Schwitzgebel & Cushman, submitted for publication). Therefore, at least in some cases, expertise does not reduce the effects of some extraneous factors.

We add to this growing body of evidence suggesting that expertise in the free will debate does not reduce the effect of personality's influence. Research on expertise sometimes uses proxies wherein a person is considered an expert if they have certain credentials, extensive experience, or particular academic degrees. Unfortunately, such proxies can be weakly or even negatively related to verifiable expert performance (Ericsson & Lehman, 1996; Ericsson et al., 2007). We reasoned that a strong test of the expertise defense would go beyond potential confounds associated with the use of reputation or professional affiliations, requiring the assessment of verifiable differences in expert knowledge (e.g., performance on a validated psychometrically sound knowledge test). A test-based approach acknowledges both the potential variability and domain specificity of philosophical expertise. It is not uncommon to find philosophers who have spent the vast majority of their professional careers specializing in one or two narrow areas of philosophy. As such, many professional philosophers may not be particularly knowledgeable regarding the free will debate. So, on our interpretation of the expertise defense, simply being a professional philosopher does not mean that a person's intuitions about free will and moral responsibility should be unbiased by personality. Rather, those who are more specifically knowledgeable in the free will debate should be increasingly free of potential judgment biases related to extraneous features (i.e., personality traits).

As an essential step in the evaluation of the expertise defense, we sought to test whether personality influenced intuitions about free will and moral responsibility controlling for expert knowledge in the free will debate. We hypothesized that:

1. Extraversion will predict compatibilist judgments (Feltz & Cokely, 2009; Nadelhoffer et al., 2009).
2. Those showing verifiable expert knowledge in the free will debate will display a more incompatibilist pattern of intuitions than those who are not free will experts (consistent with the noted, recent discussions).
3. The contributions of extraversion and philosophical (free will) expertise will be at least partially independent (i.e., extraversion will continue to predict judgment biases in those with expert knowledge).

[7] We will leave aside the important objection to the expertise defense that philosophers, when they engage in the free will debate, are trying to capture something important about what it is to be human and human interactions. Indeed, some philosophers openly admit trying to capture the folk notion of some philosophically relevant issues (Jackson, 1998, see Stich and Weinberg (2001) for a discussion).

2. Experiment

2.1. Method

2.1.1. Participants

The experiment included 121 participants recruited via an online survey platform (www.unipark.de) managed at the Max Planck Institute for Human Development in Berlin. We advertised our study via academic philosophy and psychology mailing lists in Germany that are used by advanced students (undergraduate and graduate) and faculty in psychology and philosophy.

2.1.2. Procedure and materials

Participants were tested in an online experimental 'e-lab' survey. All materials were presented in German. Participants first completed the extraversion scale from the NEO-PI-R (Costa & McCrae, 1992, German version: Ostendorf & Angleitner, 2004). According to the NEO-PI-R, each of the five global personality factors can be itemized into six different facets. In personality psychology, a facet is a narrower and more specific division of the personality trait. The global personality trait extraversion consists of six facets including warmth, gregariousness, assertiveness, activity, excitement seeking, and positive emotion. This instrument therefore provided both a general measure of extraversion as well as a finer grained assessment of the contributions of six key dimensions. This higher fidelity test enabled an exploratory assessment of more specific and proximal mechanisms that might mediate previously observed relations.

After completing the NEO-PI-R, participants completed the cognitive reflection task (CRT). The CRT is correlated with general cognitive abilities (e.g., working memory capacity; see Cokely, Kelley, and Gilchrist (2006)) and is designed to measure a cognitive style linked to more careful and thorough information processing during decision making (Cokely & Kelley, 2009; Frederick, 2005; see also Baron, 1985). Next, participants read the following scenario (translated in German):

> Most respected neuroscientists are convinced that eventually we will figure out exactly how all of our decisions and actions are entirely caused. For instance, they think that whenever we are trying to decide what to do, the decision we end up making is completely caused by the specific chemical reactions and neural processes occurring in our brains. The neuroscientists are also convinced that these chemical reactions and neural processes are completely caused by our current situation and the earlier events in our lives, and that these earlier events were also completely caused by even earlier events, eventually going all the way back to events that occurred before we were born.
>
> So, if these neuroscientists are right, then once specific earlier events have occurred in a person's life, these events will definitely cause specific later events to occur. For instance, once specific chemical reactions and neural processes occur in the person's brain, they will definitely cause the person to make the specific decision he or she makes. So, once specific earlier events have occurred in a person's life, these events will definitely cause specific later events to occur.
>
> For example, one day a person named John decides to kill a shop owner, because he needs money and does it. Once the specific thoughts, desires, and plans occur in John's mind, they will definitely cause his decision to kill a shop owner.[8] (Nahmias et al., 2007).

Participants were asked to rate their level of agreement with the following statements on a scale from 1 (absolutely disagree) to 7 (absolutely agree):

1. John is morally responsible for his action.
2. John did it because of his own free will.
3. John's decision was up to him.

Finally, participants were presented with the *Free Will Skill Test* (for information on reliability and validity see Section 4). This test contained 10 statements such as "Two important opinions in the debate about free will and determinism are called compatibilism and incompatibilism" (Appendix A). Participants were asked to indicate whether the statement was true, false or that they did not know the answer. A correct answer counted as one point, the wrong answer as a minus point, and "I don't know" as zero points, providing a correction for participants guessing. Finally, demographic information was collected (i.e., sex and age).

3. Results

3.1. Extraversion

Responses to the three prompts ($Q1$–$Q3$) were highly correlated (.477, .519, .670; $p < 0.05$ *for all)* and so a composite free will score was calculated ($Score = (Q1 + Q2 + Q3)/3$). Warmth was the only facet related to compatibilistic judgments in a

[8] This scenario is a modification of scenarios used by Nahmias, Coates, and Kvaran (2007). It is a combination of what they call "reductionistic" and "non-reductionistic" scenarios. We used this combination partially because the current version is easier to understand in German. Further, it allowed for a test of the generalizability of extraversion's predictive power using new stimulus materials.

simple regression. It explained 6% ($p = 0.005$) of the variance (see model 1 in Table 1).[9] A Bonferoni correction indicated that the relationship between warmth and compatibilist judgments remained reliable correcting for the exploratory analyses on each of the personality facets.[10] As predicted by Hypothesis 1, extraversion—or more specifically warmth, which is an important facet thereof—was systematically related to compatibilist intuitions. Importantly, with respect to the effect of personality on judgments of free will and moral responsibility, we found no reliable difference between folk and expert intuitions.

3.2. Free will skill and cognitive refection

The Free Will Skill Test predicted judgments to the three prompts ($M = 1.5$, *range*: 0–10, $SD = 1.3$) explaining 9% ($p = 0.000$) of the variance (see model 2 in Table 1). As predicted by Hypothesis 2, greater philosophical knowledge was associated with more incompatibilist intuitions. The cognitive reflection task was not reliably related to intuitions in this scenario ($p = .255$).

3.3. Full model

To test Hypothesis 3, we analyzed whether warmth explained additional variance controlling for the Free Will Skill Test score in a stepwise forward regression (Table 1). The full model explained 14% ($p = 0.000$) of the variance in participants' judgments (see model 4 in Table 1). As predicted, when controlling for expert knowledge, warmth continued to predict a moderate amount of unique judgment variance (about 5%, $p = 0.01$; see Table 1). There was no interaction with gender ($p = .56$) or age ($p = .37$). There was no higher order interaction among Free Will Skill, Warmth, and judgment ($F < 1$).

4. Discussion

Results indicated that extraversion was a reliable predictor of compatibilist judgments (H_1). This finding replicates previous research (Feltz & Cokely, 2009) and extends results to compatibilist judgments using a different scenario, in a different culture, across a wide range of ages (see also Cokely & Feltz, 2009a, 2009b). One specific facet of extraversion—i.e., warmth—was found to be primarily responsible for extraversion's predictive power. Results also provide evidence supporting some recent hypotheses suggesting that many people who are knowledgeable about the free will debate are incompatibilists (H_2). Results further indicated that both extraversion and expert knowledge were reliable and non-redundant, non-interactive predictors of judgment bias in a paradigmatic free will case.

4.1. A reliable and valid measure of philosophical expertise

Beyond the assessment of the domain general skills of logic and reasoning (Cokely & Feltz, 2009a, 2009b; Livengood, Sytsma, Feltz, Scheines, & Machery, 2010), the assessment of philosophical expertise involves the reliable measurement of some domain specific expert philosophical knowledge or skill (as in the Free Will Skill Test). The development of such tests is a new, necessary, and challenging endeavor for experimental philosophers and psychologists. Here, in contrast to all previous studies that have used self-report or other proxies (e.g., professional position) to assess philosophical expertise, we used a performance based test. Such performance based measures are typically superior to self-report skill measures in a number of aspects (e.g., objective and verifiable assessment of skill; difficult to cheat; difficult to fake).[11] However, one crucial question remains: Is the test valid and reliable?

Given the aims and results of the current study there is good reason to think the Free Will Skill Test meets or exceeds the standards of classical test theory used for the development of psychometric ability tests. To validate this claim and provide further evidence on the psychometric quality of our new instrument, we conducted a validation study with 44 philosophy graduate students (age: *mean* = 23, $SD = 2.76$; 16 *females*). Participants were invited via a classroom listserv to take the Free Will Skill Test and to complete a short survey. Among the moderately skilled sample used in our validation study (*range*: *0–9*, *mean* = 2.9, $SD = 2.2$), we did not observe any distributional skew (the Kolmogorov–Smirnov test was not significant $p = .14$); the test was relatively normally distributed.[12] Further analysis indicated that the instrument had a very high test–retest-correlation, $r = .99$ ($p = 0.000$) over short time intervals (10–30 min). The test had a Cronbach's Alpha of .75, which is above the conventional adequacy threshold for psychological instruments, providing further evidence of reliability and indicating high internal consistency. We also documented evidence of convergent validity as the scores showed substantial correlations with free will relevant self-rated knowledge ($r = .5$, $p = .001$), estimated number of papers read (*rho* = .492, $p = 0.001$), lectures

[9] Warmth has the highest correlation of all facets with extraversion ($r = .59$) in the Ten Item Personality Inventory (Gosling, Bentfrow, & Swann, 2003).

[10] Because the regressions involved some exploratory analyses, the chance of falsely detecting relations was increased (i.e., Type I error). Because it is simple, common, and conservative we used a Bonferoni correction. Given our six exploratory regressions, one for each facet, the standard alpha (.05) was divided by the number of regressions conducted (0.0083 > 0.001). See Abdi (2007) for more on this and other common corrections.

[11] We did not find any evidence of cheating. For example, there was no correlation between those who took longer on the measure and one's subsequent performance. Longer completion times should be expected if participants were looking up answers online, etc. The time someone needed for the test was uncorrelated with the number of correct answers s/he gave ($p = .923$ for Pearson's r; $p = .894$ for Spearman's Rho).

[12] In the main experiment, results were positively skewed as our sample included a large number of participants who had very little knowledge or training in philosophy. Lower scores were anticipated as few people were expected to know what "compatibilism" means as philosophers use it (along with terms like "determinism" and "moral responsibility") (Nahmias et al., 2006).

Table 1
Explained variance of the different predictors.

No.	Model	Predictor	*t* (int)	*p* (t)	Δr^2	ΔF	*p* (F)
1	Simple Regression	Free will score	31	0.001	0.09	12.4	0.001
2	Simple Regression	Warmth	3	0.002	0.06	8.3	0.005
3	Stepwise Regression	(a) Free will score	31	0.001	0.09	12.4	0.001
		(b) Warmth	3	0.001	0.05	6.6	0.011
4	Full Model	Both variables	3	0.001	0.14	9.8	0.001

Please note that absolute *r* and *F* are indicated by Δr^2 and ΔF for models that are not stepwise (1, 2, and 4).

attended (*rho* = .327, *p* = 0.032), and years one had been interested in the debate (*rho* = .385, *p* = 0.011). Finally, the primary study provided evidence of criterion validity as we observed a reliable and moderately strong relationship between expert knowledge and incompatibalist intuitions. While further research promises additional opportunities to refine and improve this test, these results provide a foundation of converging evidence that the current Free Will Skill Test is a valid and reliable measure of expert knowledge in the free will debate.

4.2. Independent contributions of extraversion and philosophical skill

We documented that the effect of personality on intuitions about free will and moral responsibility persisted for those who had some expert knowledge about the free will debate. This result directly challenges one important aspect of the expertise defense concerning intuitions about free will and moral responsibility. This aspect of the expertise defense states that in virtue of experts increased knowledge of the debates, they will be insulated from seemingly irrelevant features such as personality's influence on their intuitions. But this empirical claim appears to be largely false for our sample including highly skilled individuals—the magnitude of the influence of personality was similar for both the folk and for those who have extensive knowledge about the free will.

Those who favor the expertise defense have a variety of responses to our data. To name just a few, one could argue that the Free Will Skill instrument is not sufficiently rigorous (but see the previous section for test reliability and validity information). One could argue that our data only apply to intuitions about free will and moral responsibility. One could argue that the most elite professional philosophers engaged in the free will debate are not influenced in these ways (a restriction of range argument). Or one might argue that knowledge is the wrong (or an incomplete) measure for philosophical expertise. Rather, philosophers may obtain a particular set of cognitive skills that makes them less susceptible to the effects of extraneous factors such as personality (Williamson, 2007, in press; see also Weinberg et al., 2010). For example, expert philosophers may not only be knowledgeable of the free will debate, they may also have the ability to focus on relevant factors of scenarios and thus can avoid influenced by irrelevant factors. As such, our study does not *show* that "experts" necessarily had the relevant constellation of expert abilities. That is, perhaps we incorrectly operationalized expertise.

We agree that these are all live, possible responses to our interpretation of the data. However, these objections rest on the *possibility* of their truth. And possibility is cheap (Feltz, 2008). In contrast, the current data suggest not only that it is possible that those with expert knowledge show bias, but that it is *actual* in some important cases and in verified, representative samples. Those who are extraverted and have some expert knowledge about the free will debate are more likely to be compatibilists about some paradigmatic cases than those who are knowledgeable and not extraverted. At a minimum, our results suggest that one important factor in expertise (reproducible and verifiable knowledge) does not reduce the effects of at least one extraneous factor (personality) on judgments of freedom and moral responsibility.[13]

The gathering data puts the critic of the expertise defense in a rhetorically strong position. If either indirect strategies or direct strategies against the expertise defense are correct, then the expertise defense does not shield expert philosophers from some extraneous factors. While determining burdens of proof is notoriously difficult to do, we can simply assume that experimentalists who challenge the expertise defense have the burden of proof. After all, there is something to be said for important, long, and storied traditions such as expert use of intuitions. However, given our (and other's) data indicating expertise does not make a difference, the burden of proof levied against the experimentalists is largely satisfied (see Williamson (in press) for a burden of proof argument against the experimentalist). Those who wish to maintain the expertise defense must provide some actual evidence that expertise (however operationalized) makes a difference relevant to supporting the expertise defense (Alexander & Weinberg, 2007). Absent that evidence, and with the growing evidence that calls into doubt the expertise defense, the expertise defense loses most of its insulating power.

4.3. Sources of disagreement and agreement

The current results may help us understand why some philosophical debates never seem to go away. Studies suggest that individual differences in personality are partly or even largely heritable (Bouchard, 1994; Jang, Livesley, & Vernon, 1996). If

[13] Current research in our labs is investigating the relation of personality to expert, philosohpcially relevant cognitive abilities.

differences in personality are heritable, and personality influences one's philosophically relevant intuitions, then it stands to reason that differences in philosophically relevant intuitions may also be heritable.[14] If intuitions provide evidence for one's philosophical view, then it seems as if the philosophical view one ends up endorsing will be partially a function of what personality one has inherited. As a result, many philosophical debates would continue through the ages and few would be resolved by reason alone.

If something like the preceding argument is correct, then it offers an intriguing problem for much of traditional 20th century analytic philosophy. This tradition largely relies on intuitions about particular cases as evidence for or against a particular theory. As others have argued (Alexander & Weinberg, 2007), finding that personality systematically influences philosophically relevant intuitions could not have been discovered from the armchair. Hence, philosophy needs to change. Many (if not most) philosophers need to leave the armchair and consult more closely with empirical scientists to be sure that they have the best evidence and that they are not influenced by seemingly extraneous factors. Second, and perhaps more radically, in order to solve philosophical debates it may be necessary to change the way that philosophy is conducted. Progress has been made in the sciences by standardizing the way things are done: conventions are put in place, standards of evidence are created, and general methodologies are endorsed and assessed. While these types of things do not guarantee the elimination of controversies (even about those very standards, e.g., Gigerenzer, 2004), they do allow a general and self-conscious way for theorists to proceed, change, and evaluate progress. In this way, the influence of seemingly extraneous factors on theoretical pursuits could be minimized or at least measured and monitored. Such efforts may yield appropriate resources so that the influence of seemingly irrelevant factors (including personality) is minimized. Hence, philosophy may need to undergo a transformation to solve its own problem.[15] How radical this transformation needs to be remains to be seen (e.g., Kornblith, 1998; see Alexander & Weinberg, 2007 for a discussion).

The current findings also offer suggestions about how to approach differences in fundamental philosophical debates. If personality influence people's intuitions, that would be an important consideration when debating a contentious philosophical point. Rather than assuming that someone with dissenting viewpoints is irrational, has false beliefs, or otherwise engages in sub-optimal reasoning, it may be better (or more accurate) to recognize that they have a different personality and thus a different intuitional grounding. Once this is recognized, there is an opportunity to develop more fruitful discussions regarding the sources and potential resolution of disagreements. In short, this insight may offer new ways for reconciliation both in philosophy and in other fundamental theoretical or ideological debates (see Haidt & Graham, 2007 for similar remarks about morality).

The interplay of expertise and personality seems to also have important implications for a number of applied domains. For example, results suggest that even highly skilled professionals such as lawyers, judges, ethicists, and philosophers may not be "immune" to the influence of their different personalities: Some professionals will show biases. While it is clear that biased minds often make better decisions (Gigerenzer & Brighton, 2009), biases can also lead to errors. Indeed, related dynamics can be of great concern for professional practices that involve novices as might be the case when a jury is largely composed of extraverts. Such a jury may be more willing to hold a person morally responsible for an action even if the person could not have done anything to prevent the action from coming about (see Nadelhoffer (2006) for a similar argument in intentional action). It may be desirable for a jury to consist partially or completely of extraverts. However, it is not obviously true. Philosophical arguments need to be forwarded and empirical evidence needs to be evaluated. But before we could even have a philosophical debate about whether a jury should be composed of extraverts, we should have a reasonable idea about how extraverts (or those with other personalities) tend to make court-relevant decisions. Importantly for traditional philosophy, understanding how extraverts make decisions is not something that can be done from the armchair.

These considerations offer the intriguing possibility that personality may allow us to determine whose intuitions about free will and moral responsibility ought to be trusted. However, caution needs to be taken when thinking about whose intuitions ought to be trusted. To a large extend, whether intuitions ought to be trusted depends on what intuitions are to be trusted for. As the above discussion suggests, perhaps extraverts' intuitions about freedom and moral responsibility ought to be trusted more (or less) in courtroom decisions. To illustrate, extraverts may be less likely to evaluate excusing conditions for bad actions appropriately. In such cases, we may have reason to prefer intuitions of non-extraverts because they are less likely to be biased towards attributing freedom and moral responsibility even if excusing conditions are present.

But what about trusting extravert's intuitions for philosophical uses? Again, it depends on what project one is engaged in (see Stich (2010) and Feltz and Cokely (in press-a), Feltz and Cokely (in press-b), for more detailed discussion). If one is interested in doing conceptual analysis, then intuitions of extraverts could be trusted to reflect the concepts that groups of people have (of course, the same goes for non-extraverts). After all, extraverts are expressing *their* intuitions about the relation of

[14] Of note, there are three different ways for genotype and environment to covary (Plomin, DeFries, & Loehlin, 1977). The first is "active" where one chooses environments according to one's genes. For example, each new generation of philosophers, because of their personality, finds some positions and examples intuitively appealing and adopts them. The second is "reactive" where one's reaction to an environment is partially influenced by one's genes. For example, each new generation of philosophers will likely adopt a view that is consistent with previous philosophers with similar personalities because they react positively to views that are the same as their own. The last is "passive" where one's environment is mostly formed by people that share in part the same genes. For example, if one is the child of a philosopher, one will probably not only share half of one's genes, but also some philosophy books. Depending on which type of covariance is at play in philosophical debates, genes could offer even more of an explanatory role.

[15] See a related discussion on the Experimental Philosophy Blog: http://experimentalphilosophy.typepad.com/experimental_philosophy/2010/02/xphi-its-all-about-not-being-a-sucker.html.

determinism with free will and moral responsibility. And one could reasonably assume that the intuitions tend to reflect those people's concept of freedom and moral responsibility.[16]

However, it is much less clear that extravert's intuitions are to be trusted for what can be called "Neo-Platonic" projects (Stich, 2010). Neo-Platonic projects attempt to discover the truth about some non-conceptual or non-linguistic philosophically relevant phenomenon. For example, some theorists want to know what the relation between determinism and freedom *is* and not some people's *conception* of that relation. In determining the nature of this relation, we take it that personality is extraneous to the truth of the content of the intuition one has about that relation. To the extent that the truth of the content varies with personality, we should not be confident that the intuitions are truthfully reflecting the nature of that relation. One may think that because warm people are generally more interested in other human beings they are better suited to make judgments about human agency. But it is not clear why we should give that general attitude preference when it concerns the question of whether free will and moral responsibility are compatible with the truth of determinism. It seems equally plausible that non-extraverts may appropriately weigh the impact of determinism on whether one is free or morally responsible without being caught up in desires to hold people morally responsible.

4.4. Directions for future research: How does extraversion influence judgments?

Because global personality traits can be too blunt to measure individual differences that may exist, studying the individual facets of global personality traits is an important step in understanding the proximal judgment processes and psychology involved in generating these intuitions.[17] In our study, we found that only warmth predicted compatibilist intuitions. Other facets of extraversion were unrelated to compatibilist intuitions.

That warmth is largely responsible for the relationship between personality and compatibilist intuitions converges with, and could help explain, the results of many experiments about free will. Warmth is considered an important facet in human social interaction and is defined as interest in friendliness towards others (Costa & McCrae, 1992). Because warm people tend to be more compassionate, care more about other *people*, and are more interested in *human interactions* than non-warm people, it seems that warm people's traits (i.e., other-directedness and compassionate attitudes) could make them more inclined to give compatibilist judgments in the types of cases we used in our experiment (i.e., highly affective cases). For example, in our experiment, warm people may be especially empathetic to the wife. Because extraverts are especially sensitive to the harming of the wife, they may be especially interested in holding the husband morally responsible.[18] This general pattern is what we find in almost all of the relevant studies about free will conducted so far—when the studies are about human beings many more people give compatibilist friendly intuitions compared to studies about concretely described non-human but human-like creatures.[19] There is a large body of evidence suggesting that in many cases people have compatibilist judgments when their judgments are made about specific people (Nahmias et al., 2004, 2005, 2006; Nichols & Knobe, 2007; Feltz, Cokely, & Nadelhoffer, 2009). However, people have less compatibilist friendly intuitions when the studies involve abstract questions (Nichols & Knobe, 2007), when they are about alternative universes (Roskies & Nichols, 2008), or when they involve bypassing of "human agency" (Nahmias et al., 2007). In this way, the identification of warmth as the main facet underlying the relations between extraversion and compatibilist intuitions unifies previous research and provides some insights into the underlying psychological mechanisms that may give rise to the observed judgment biases.

5. Conclusion

We have put the expertise defense to the test and find that after controlling for verifiable expert knowledge of philosophical debates concerning free will and moral responsibility, extraversion continues to predict compatibilist judgment biases. These results directly challenge the expertise defense and give us new clues about what processes generate some intuitions concerning freedom and moral responsibility. These results are also likely to be important for a host of applied domains such as law, medicine, and vocation selection, and further highlight the critical role that empirical research can play in understanding ancient and fundamental philosophical issues. As has been said before, given the stakes, we had better go after "the important questions with everything we've got" (Knobe & Nichols, 2008). Philosophy cannot continue from the armchair alone—philosophy must become a substantially empirical enterprise (Feltz & Cokely, in press-a, in press-b).

[16] Intuitions may not reflect people's concept (Machery, 2009). In that case, these intuitions could reflect people's conceptions (Cushman & Mele, 2008).

[17] For example, Ziegler, Danay, Schölmerich, and Bühner (2010) found that the global trait extraversion is unable to predict performance in a University exam. However, an examination of the facets of extraversion reveals a different pattern. The facets gregariousness (−.14), activity (.14), and excitement seeking (−.18) are all related to performance on the exam. However, some of the facets point in different directions and cancel each other out at the global level. Hence, in the case of a university exam, the global trait fails to detect theoretically relevant individual differences. Something similar may be true for the use of global personality traits in the assessment of other philosophically relevant intuitions.

[18] We would like to thank Dan Jones for this helpful illustration.

[19] One exception to this general pattern is found in Knobe and Nichols (2008) where people judged that a man who cheated on his taxes was not morally responsible. However, the results about this particular case are mixed with some results indicating that people judge the man as free and morally responsible (for a more detailed discussion, see Feltz, Cokely, & Nadelhoffer, 2008).

Appendix A

1. Well-known counterexamples for the PAP are called the Frankfurt cases. (true).
2. Arthur Schopenhauer said that there is definitely free will. (false).
3. Two important opinions in the debate about free will and determinism are called compatibilism and incompatibilism. (true).
4. PAP stands for the principle of alternate personalities. (false).
5. One frequently used argument for the freedom of choice is the experiment from Benjamin Libet. (false).
6. One well-known believer in free will was Jean Paul Satre. (true).
7. The classical Trolley Problem is about two trains on a collision course. (false).
8. William James suggested that there could be soft determinism. (true).
9. One argument in the field of moral philosophy is Moore's open statement argument. (false).
10. The Stockholm's interpretation sees quantum physics as an argument against determinism. (false).

References

Abdi, H. (2007). Bonferroni and Sidak corrections for multiple comparisons. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics*. Thousand Oaks, CA: Sage.

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

Baron, J. (1985). *Rationality and intelligence*. New York, NY: Cambridge University Press.

Bouchard, T. J. (1994). Genes, environment, and personality. *Science, 264*, 1700–1701.

Cokely, E. T., & Feltz, A. (2009a). Individual differences, judgment biases, and theory-of-mind: Deconstructing the intentional action side effect asymmetry. *Journal of Research in Personality, 43*, 18–24. doi:10.1016/j.jrp. 2008.10.007.

Cokely, E. T., & Feltz, A. (2009b). Adaptive variation in folk judgment and philosophical intuition. *Consciousness and Cognition, 18*, 355–357. doi:10.1016/j.concog.2009.01.001.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4*, 20–33.

Cokely, E. T., Kelley, C. M., & Gilchrist, A. H. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review, 13*, 991–997.

Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Cover, J. A., & O'Leary-Hawthorne, J. (1996). Free agency and materialism. In J. Jordan & D. Howard-Snyder (Eds.), *Faith, freedom, and rationality* (pp. 47–71). Lanham: Roman and Littlefield.

Cushman, F., & Mele, A. (2008). Intentional action: Two-and-a-half folk concepts? In J. Knobe & S. Nichols (Eds.), *Experimental philosophy* (pp. 171–188). Oxford: Oxford University Press.

Dennett, D. (1984). I could not have done otherwise: So what? *Journal of Philosophy, 81*, 553–565.

Ekstrom, L. (2002). Libertarianism and Frankfurt-style cases. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 309–322). New York: Oxford University Press.

Ericsson, A., & Lehman, A. (1996). Expert and exceptional performance: Evidence on maximal adaptions on task constraints. *Annual Review of Psychology, 47*, 273–305.

Ericsson, A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review, 85*, 114–121.

Feltz, A. (2008). Problems with the appeal to intuition in epistemology. *Philosophical Explorations, 11*, 131–141.

Feltz, A., & Cokely, E.T. (2007). An anomaly in intentional action ascription: More evidence of folk diversity. In D.S. McNamara, G. Trafton, (Eds.), *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (p.1748). Mahwah, NJ: Erlbaum.

Feltz, A., & Cokely, E.T. (in press-a). Individual differences in theory-of-mind judgments: Order effects and side effects. *Philosophical Psychology*. doi:10.1080/09515089.2011.556611.

Feltz, A., & Cokely, E.T. (in press-b). The philosophical personality argument. *Philosophical Studies*. doi:10.1007/s11098-011-9731-4.

Feltz, A., & Bishop, M. (2010). The proper role of intuitions in epistemology. In M. Milkowski & K. Talmont-Kaminski (Eds.), *Beyond description: Normativity in naturalised philosophy* (pp. 101–122). London: College Publications.

Feltz, A., & Cokely, E. T. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences intuitions about compatibilism and incompatibilism. *Consciousness and Cognition, 18*, 342–350. doi:10.1016/j.concog.2008.08.001.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 24–42.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics, 33*, 587–606.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences? *Topics in Cognitive Science, 1*, 107–143.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504–528.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98–116.

Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613–628.

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

Jang, K., Livesley, W., & Vernon, P. (1996). Heritability of the big five personality dimensions and their facets: A twin study. *Journal of Personality, 64*, 577–591.

Kane, R. (1996). *The significance of free will*. New York: Oxford University Press.

Kane, R. (1999). Responsibility, luck, and chance: Reflections on free will and indeterminism. *Journal of Philosophy, 96*, 217–240.

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

Knobe, J., & Nichols, S. (2008). An experimental philosophy manifesto. In J. Knobe & S. Nichols (Eds.), *Experimental philosophy* (pp. 3–14). New York: Oxford University Press.

Kornblith, H. (1998). The role of intuition in philosophical inquiry: An account with no unnatural ingredients. In M. DePaul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 129–141). Lanham: Rowman and Littlefield.

Livengood, J., Sytsma, J., Feltz, A., Scheines, R., & Machery, E. (2010). Philosophical temperament. *Philosophical Psychology, 23*, 313–330.

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

Lycan, W. (2004). Free will and the burden of proof. In A. O'Hear (Ed.), *Minds and person: Royal Institute of Philosophy supplement* (pp. 107–122). Cambridge: Cambridge 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*, 1–12.

McCann, H. (2005). Intentional action and intending: Recent empirical studies. *Philosophical Psychology, 18*, 737–748.
Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. *Philosophical Explorations, 9*, 203–220.
Nadelhoffer, T., Kvaran, T., & Nahmias, E. (2009). Temperament and intuition: A commentary on Feltz and Cokely. *Consciousness and Cognition, 18*, 351–355.
Nahmias, E., Coates, J., & Kvaran, T. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy, 31*, 214–242.
Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2004). The phenomenology of free will. *The Journal of Consciousness Studies, 11*, 162–179.
Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology, 18*, 561–584.
Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research, 73*, 28–53.
Nichols, S. (2004). The folk psychology of free will: Fits and starts. *Mind and Language, 19*, 473–502.
Nichols, S. (2007). The rise of compatibilism: A case study in the quantitative history of philosophy. *Midwest Studies in Philosophy, 31*, 260–270.
Nichols, S., & Folds-Bennett, T. (2003). Are children moral objectivists? Children's judgment about moral and response-dependent properties. *Cognition, 90*, B23–B32.
Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuition. *Nous, 41*, 663–685.
Ostendorf, F., & Angleitner, A. (2004). *NEO-PI-R. NEO Persönlichkeitsinventar nach Costa und McCrae. Revidierte Fassung. (NEO-PI-R. NEO personality inventory)*. Göttingen: Hogrefe.
Pereboom, D. (1995). Determinism al dente. *Nous, 29*, 21–45.
Pink, T. (2004). *Free will: A very short introduction*. New York: Oxford University Press.
Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*(2), 309–322.
Roskies, A., & Nichols, S. (2008). Bringing moral responsibility down to earth. *Journal of Philosophy, 105*, 371–388.
Schwitzgebel, E., & Cushman, F. (submitted for publication). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers.
Schwitzgebel, E. (2009). Do ethicists steal more books? *Philosophical Psychology, 22*, 711–725.
Schwitzgebel, E., & Rust, J. (2010a). Do ethicists and political philosophers vote more often than other professors? *Review of Philosophy and Psychology, 1*, 189–199.
Schwitzgebel, E., & Rust, J. (2010b). The moral behavior of ethicists: Peer opinion. *Mind, 118*, 1043–1069.
Smilansky, S. (2002). Free will, fundamental dualism, and the centrality of illusion. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 489–505). New York: Oxford University Press.
Smith, M. (1994). *The moral problem*. Malden: Blackwell.
Sosa, E. (2007a). Experimental philosophy and philosophical intuition. *Philosophical Studies, 132*, 99–107.
Sosa, E. (2007b). Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien, 74*, 51–67.
Stanley, J. (2005). *Knowledge and practical interests*. New York: Oxford University Press.
Stich, S. (2010). Experimental philosophy and the bankruptcy of "The Great Tradition." In *Experimental philosophy workshop*, Wroclaw, Poland.
Stich, S., & Weinberg, J. (2001). Jackson's empirical assumptions. *Philosophy and Phenomenological Research, 62*, 637–643.
Strawson, G. (1986). *Freedom and belief*. Oxford: Oxford University Press.
Weinberg, J. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy, 31*, 318–343.
Weinberg, J., Gonnerman, C., Buckner, C., & Alexander, J. (2010). Are philosophers expert intuiters? *Philosophical Psychology, 23*, 332–355.
Weinberg, J., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*, 429–460.
Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell.
Williamson, T. (in press). Philosophical expertise and the burden of proof. *Metaphilosophy*.
Wolf, S. (1990). *Freedom within reason*. New York: Oxford University Press.
Ziegler, M., Danay, E., Schölmerich, F., & Bühner, M. (2010). Predicting academic success with the big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality, 24*, 1–15.