

Christoph Luetge, Hannes Rusch, & Matthias Uhl (eds.), *Experimental Ethics: Toward an Empirical Moral Philosophy*. New York: Palgrave, 2014. ISBN 9781137409799, \$100, Hbk.

It would be unkind but not inaccurate to say that most experimental philosophy is just psychology with worse methods and better theories. In *Experimental Ethics: Towards an Empirical Moral Philosophy*, Christoph Luetge, Hannes Rusch, and Matthias Uhl set out to make this comparison less invidious and more flattering. Their book has sixteen chapters, organized into five sections and bookended by the editors' own introduction and prospectus. Contributors hail from four countries (Germany, USA, Spain, and the United Kingdom) and five disciplines (philosophy, psychology, cognitive science, economics, and sociology). While the chapters are of mixed quality and originality, there are several fine contributions to the field. These especially include Stephan Wolf and Alexander Lenger's sophisticated attempt to operationalize the Rawlsian notion of a *veil of ignorance*, Nina Strohminger et al.'s survey of the methods available to experimental ethicists for studying *implicit morality*, Fernando Aguiar et al.'s exploration of the possibility of operationalizing *reflective equilibrium* in the lab, and Nikil Mukerji's careful defusing of three *debunking arguments* about the reliability of philosophical intuitions.

Part I introduces experimental philosophy as a promising but problematic methodology with several precedents in the history of philosophy and related fields. It begins with a reprint of Kwame Anthony Appiah's 2007 presidential address to the Eastern Division of the American Philosophical Association, followed by chapters authored by two of the editors (Luetge and Rusch, the latter of whom co-authors with Niklas Dworazik). Readers already familiar with the field should skip to the next section, but for newcomers these chapters will be a helpful introduction to the field.

The five chapters in Part II are case studies for experimental ethics. In chapter 5, Eric Schwitzgebel summarizes his research program on the moral behavior of ethicists; in a somewhat dismal series of papers he has shown that, with a few exceptions, professional ethicists are indistinguishable from other philosophers and professors. While such results might lead to skepticism about the effects of studying ethics, this is not the conclusion Schwitzgebel draws (indeed, to get evidence for such skepticism, one would have to randomize people to philosophical specializations and careers). Instead, he reflects on the tension between the epistemic and ethical values associated with the study of philosophy, pointing out that pressure to live in accordance with the values one advocates professionally may lead to motivated reasoning that obscures the moral truth. In Chapter 6, Verena Wagner offers an interpretation of some studies of the “Knobe effect.” Unfortunately, her familiarity with this literature is limited and out-of-date, leading her to focus on the red herring of blameworthiness. For better-informed interpretations of this literature, see Robinson et al. (2015) and Sauer (2014). In chapter 7, Ezio di Nucci summarizes some of his research on the trolley problem. This early research is a good starting point for the methodological refinements suggested in later chapters. For instance, because di Nucci employed categorical variables for both his predictors and his outcomes, he was only able to run an underpowered χ^2 test of independence; with richer measures, his plausible hypothesis about the doctrine of double effect could be better tested. In chapter 8, Wolf and Lenger describe a two-stage experiment on distributive justice, which suggests that “people tend to equalize income as suggested by Rawls’s difference principle” when behind an experimental veil of ignorance, but that when “the veil is lifted, individuals egoistically choose in line with their post-veil interests” (p. 95). They are aware of how denuded of personality people are supposed to be behind the veil of ignorance and admit that it is impossible to produce such

conditions in the lab. I heartily endorse their use of monetary incentives and measurement of behavior (rather than the verbal responses to hypothetical scenarios more typically recorded by experimental ethicists), though I worry that even with these methodological improvements the degree to which they can approximate the original position is very limited. In chapter 9, Ulrich Frey explores various potential explanations for the disconnect between people's stated values related to environmental protection and their behavior – a pressing question that is also relevant to the debate between motivational internalists and externalists.

Part III contains four chapters on methodology in experimental ethics, the crown jewel of which is chapter 10 by Nina Strohminger, Brendan Caldwell, Daryl Cameron, Jana Schaich Borg, and Walter Sinnott-Armstrong. They explore the strengths and weaknesses of three implicit judgment tasks (the Implicit Association Test, the affect misattribution procedure, and the process dissociation procedure), as well as eye-tracking and functional magnetic resonance imaging (fMRI). While there is some value to asking people to read, think about, and respond explicitly to questions about hypothetical scenarios as experimental ethicists have tended to do, such research is limited by people's introspective awareness of and willingness to sincerely express their judgments and preferences. Implicit measures provide a window into the unconscious and the intentionally obscured aspects of moral cognition and behavior. Naturally, none of these measures is perfect either, so Strohminger et al. conclude by calling for “a multi-method, integrated approach” (p. 146) that exploits the advantages of each method to make up for the weaknesses of others. In chapter 11, Martin Bruder and Attila Tanyi suggest a within-subjects methodology meant to prompt participants to “subject their initial responses to a thorough ‘test of reflection’” (p. 167), which would help to distinguish genuine intuitions from mere hunches. While I am unconvinced that the method they propose does the trick, I agree that

it would be a significant methodological improvement to include measurements before and after reflection (and to distinguish between solitary reflection and dialogic reflection). In chapter 12, Andreas Bunge and Alexander Skulmowski explore the methodology of designing institutions in such a way that people find it easier to do what they do or would reflectively judge to be the right thing – an approach that I have dubbed *moral technology* (Alfano 2013). As they put it, “Carefully constructed institutions avoid creating conflicts between multiple psychological systems of moral judgment” (p. 181). They also emphasize “how little resemblance filling out a survey form involving more or less contrived scenarios bears to making a moral judgment in everyday life” (p. 184) and advocate more ecologically valid methods, including “immersive virtual environment technology” (p. 186). In chapter 13, Fernando Aguiar, Antonio Gaitán, and Blanca Rodríguez-López argue that experimental ethicists should conduct behavioral studies like those familiar in experimental economics. These studies employ tailored scripts, repeated trials (to get within-subjects data and allow for learning), and financial incentives (to promote engagement and sincerity).

Part IV includes two critical reflections on the state of experimental ethics. The first, by Jacob Rosenthal, is an ill-informed dud, but chapter 15 by Mukerji does an excellent job of charitably engaging with and defusing three empirically-motivated arguments against the use of the method of cases to test and refine moral principles: “the argument from disagreement, the argument from framing effects, and debunking explanations” (p. 227). Regarding disagreement, Mukerji helpfully points out that studies to date have tended to focus on casuistic controversies, not easy cases, and that “there are many cases on which we should reasonably expect no disagreement at all (even among philosophers)” (p. 232). Regarding framing effects, Mukerji objects that the existence of these may indicate that participants are responsive to the reasons

embedded in hypothetical scenarios – just not to all of them at once. Indeed, I contend that framing effects show that people *do* respond to reasons, including moral reasons (Koralus & Alfano forthcoming). Regarding debunking arguments that point to the emotional sources of moral judgments, Mukerji plausibly denies that emotions are *ipso facto* unreasonable. Cognitive theories of moral emotions, such as Roeser’s (2011) *affectual intuitionism*, support this contention.

Part V includes two chapters and a brief prospectus by the editors. I mentioned above that certain questions about the effects of training in professional ethics could only be answered by randomizing people to philosophical specialization. Julian Müller’s proposal in chapter 16 is even more ambitious. He argues that certain questions in empirical ethics can only be answered by conducting large-scale social experiments, such as Startup Cities. “A Startup City is essentially a newly founded city that is part of a larger entity like a modern democratic state or a Union like the EU but has considerably more freedom to test different socio-economic policy schemes, while adhering to some minimal standard of human rights and free exit” (p. 261). This is a truly radical proposal, though one with the gold-standard precedent of Plotinus’ plan to found Platonopolis. The editors, drawing on the previous chapters, conclude the book by identifying what they consider the five most pressing problems in contemporary experimental ethics: 1) inadequate methodological rigor, 2) lack of clarity regarding the relation between experimental philosophy and armchair philosophy, 3) lack of acceptance of experimental methods among some philosophers, 4) over-reliance on verbal responses to hypothetical scenarios, 5) lack of integration with other disciplines. I agree with all five, especially items 1 and 4. I would add as item 6 the fact that experimental philosophy continues to have a “woman problem.” Of the twenty-six contributors to this volume, only four are women (15%). This is in

line with the problematic gender ratio identified by Peggy DesAutels (2015) in other collections of experimental philosophy. Given that two of the four best papers in this volume have female authors, the quality of work done by women in the field of experimental ethics cannot account for the disparity. If future editors of volumes in experimental philosophy do not make a conscious effort to promote gender diversity, experimental philosophy may improve its methods, but it will still be unkind but not inaccurate to say that it is just social science with less diversity and better theories.

References

- Alfano, M. (2013). *Character as Moral Fiction*. Cambridge University Press.
- DesAutels, P. (2015). [Review of J. Knobe & S. Nichols (eds.), *Experimental Philosophy, Volume 2*.] *Notre Dame Philosophy Reviews*.
- Koralus, P. & Alfano, M. (forthcoming). Reasons-based moral judgment and the erotetic theory. In J.-F. Bonnefon & B. Trémolière (eds.), *Moral Inference*. Psychology Press.
- Robinson, B., Stey P., & Alfano, M. (2015). Reversing the side-effect effect: The power of salient norms. *Philosophical Studies*, 172(1): 177-206.
- Roeser, S. (2011). *Moral Emotions and Intuitions*. Basingstoke: Palgrave.
- Sauer, H. (2014). It's the Knobe effect, stupid! *Review of Philosophy and Psychology*, 5(4): 485-503.

MARK ALFANO

Ethics & Philosophy of Technology

Delft University of Technology

Jaffalaan 5, B4.250

2628 BX Delft

The Netherlands