TU Delft

Delft University of Technology

Ethics, morality, and game theory

Alfano, Mark; Rusch, Hannes; Uhl, Matthias

DOI
10.3390/g9020020

Publication date
2018

Document Version
Final published version

Published in
Games

Citation (APA)
Alfano, M., Rusch, H., & Uhl, M. (2018). Ethics, morality, and game theory. *Games*, *9*(2), Article 20. https://doi.org/10.3390/g9020020

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

games

MDPI

Editorial

Ethics, Morality, and Game Theory

Mark Alfano [1,2], Hannes Rusch [3,4] and Matthias Uhl [4,*]

1 Faculty of Technology, Policy and Management, Delft University of Technology, 2628 BX Delft, The Netherlands; mark.alfano@gmail.com
2 Institute for Religion and Critical Inquiry, Australian Catholic University, East Melbourne, VIC 3002, Australia
3 School of Business and Economics, Philipps University Marburg, 35037 Marburg, Germany; hannes.rusch@tum.de
4 TUM School of Governance, Technical University of Munich, 80333 Munich, Germany
* Correspondence: m.uhl@tum.de; Tel.: +49-89-907793-280

Received: 24 April 2018; Accepted: 25 April 2018; Published: 26 April 2018

Abstract: Ethics is a field in which the gap between words and actions looms large. Game theory and the empirical methods it inspires look at behavior instead of the lip service people sometimes pay to norms. We believe that this special issue comprises several illustrations of the fruitful application of this approach to ethics.

Keywords: morals; ethics; strategic interaction; game theory; behavior; economics; philosophy

1. The Relation between Game Theory and Ethics

Game theory rests on assumptions about players' preferences and their mutual expectations regarding everyone else's behavior. From these sets of assumptions, game theorists derive predictions about the outcomes of interactions. Ethics, by contrast, is a normative discipline informed by reflection on moral behavior and attitudes. The application of game theory to ethics dates back to 1954 when Richard Braithway lectured on the "Theory of games as a tool for the moral philosopher" [1]. Elements of game theoretical reasoning about ethical problems, however, can already be found in the works of Thomas Hobbes and David Hume [2].

Game theory enables philosophers to apply mathematical rigor to questions of ethics. This is not welcomed by all. Robert Solomon, for instance, has opined that the use of game theory has been a disaster in ethics [3]. He argues—quoting Aristotle as a witness—that formal thinking does not mix easily with the sensitivities essential to ethics. According to Solomon, the "rational" agent of game theory is a monster and thinking in terms of the maximization of self-interest, in contrast to doing the right thing unhesitantly, is to have one thought too many.

In a direct reply to Solomon, Ken Binmore, one of the protagonists of applying game theory to ethics, argues that his defense strategy will be identical to the innocent husband who is asked why he is beating his wife: Instead of justifying his misconduct, he insists that he is not beating his wife at all [4]. Binmore emphasizes that game theory is ethically neutral because it is merely a theory of consistent choice. He notes that it analyzes the results of people's actions given their motives, i.e., preferences, without reflection on what these motives ought to be. Thus, game theory can be used to model the behavior of altruistic agents as well as the behavior of selfish agents.

Indeed, the claim sometimes brought forward by philosophers that everybody knows that people who play the prisoner's dilemma actually manage to mutually cooperate is misleading. If participants in experiments systematically and knowingly choose what would not be a dominant strategy in terms of maximizing monetary payoffs, then they obviously care about something in addition to their own monetary payoffs. This, however, does not mean that the 'prisoner's dilemma is falsified' in any

sensible meaning of the term. It rather means that the game people play after having transformed the raw payoffs according to their preferences is no prisoner's dilemma. However, maybe nobody has ever claimed that they do.

Mixing up monetary payoffs and utilities is the cardinal fault of many critics of game theory. Of course, some economists may assume that people are selfishly maximizing their monetary payoffs, at least in anonymous market environments, when in fact they are not. Falsifying this empirically accessible assumption is a meaningful exercise, while falsifying an analytical tool is not. Game theorists are certainly the wrong parties to blame in this context.

We do not believe that (the application of) game theory is a threat to ethics. On the contrary, we contend that it can offer deep and intriguing insights into questions central to ethics, inspire new questions, and provide the conceptual rigor required to make ethical questions accessible to empirical testing. Maybe most importantly, game theory seems to be the only reasonable candidate for the unifying language needed to address ethical questions interdisciplinarily, insofar as it provides a common ground for the social sciences and the evolutionary sciences.

Ethics is a field in which the gap between words and actions looms large. Game theory and the empirical methods it inspires look at behavior instead of the lip service people sometimes pay to norms. We believe that this special issue comprises several illustrations of the fruitful application of this approach to ethics.

2. Motivation for This Special Issue

In their Stanford Encyclopedia of Philosophy on experimental moral philosophy, Alfano and Loeb point out that the paradigmatic outcome measure of experimental philosophy is the survey [5]. In a recent paper, Joshua Knobe systematically reviews all empirical studies in the field of experimental philosophy over the last five years, finding that just over 10% aimed to provide evidence for or against a proposed conceptual analysis, while the vast majority of studies were interpreted as evidence for or against a proposed account of the cognitive and affective processes underlying the relevant phenomena [6]. In either case, though, the predominant outcome measures were surveys, questionnaires, and verbal reports.

This is all well and good, but we believe that, for several reasons, a fuller understanding of moral happenings within and between people would benefit from theoretical and methodological diversification [7]. First, people's motivations are to some extent opaque, even to themselves. They may not be well-placed to say how strongly they are motivated to pursue, promote, or protect a given value. Second, affective and motivational phenomena are expressed most directly in action, not in language. Finally, Likert scales and similar measures provide at best ordinal measures of variables of interest. In order to make our models more precise and our predictions easier to falsify, cardinal measures are needed.

Game theory and the related empirical fields of behavioral and experimental economics respond to each of these concerns. Instead of asking people how much they value something or how motivating they find a reason, we can put them in a position to act based on that value or reason. Instead of finding directional effects that cannot be more precisely quantified, we can specify utility functions. Instead of surveying people's responses based on their potentially blurry subjective representations of which consequences an action in a given scenario might entail, we can precisely specify their choice sets and define exact outcomes for all decisions that can be made.

Experimental philosophy thus has much to gain from engaging with game theory. At the same time, game theory, behavioral economics, and experimental economics can benefit from engagement with the long history of philosophical reflection on the nature of values, motives, moral reasoning, and so on. This history provides a ready supply of models and theories to be articulated and tested. The papers in this special issue showcase some of these prospects for interdisciplinary research.

3. Contents of the Issue

3.1. Theoretical Contributions

Sarkisian [8] and Alger and Weibull [9] analyze models of strategic interactions between agents who have moral and altruistic preferences. Sarkisian outlines conditions under which teams of moral or altruistic agents perform better than teams of egoists in a principal-agent setup.

Complementarily, Alger and Weibull study a range of strategic interactions between single agents whose utility functions reflect altruistic and moral preferences. Intriguingly, they find that, although altruism and morality can improve outcomes in some games, they may hamper cooperation in others.

Van Basshuysen [10] and Baumann and Friehe [11] devote their theoretical analyses to problems currently discussed in applied ethics. Van Basshuysen uses the framework of matching theory to study how policies regulating refugee distribution to host countries could be designed to take both host countries' and refugees' preferences into account. He identifies trade-offs between the fairness and the efficacy of different mechanisms, but also outlines ways in which these trade-offs can be circumvented.

Baumann and Friehe take on the question of how firms might react to consumers' increasing interest in the moral conduct of companies. Focusing on the possibility of adverse effects of this recent development, they find that an increased demand for 'fair' modes of trade and production may backfire and induce firms to covertly resort to less preferable modes of production.

3.2. Experimental Contributions

Della Valle and Ploner [12] investigate how group identity modulates retributive dishonesty as a reaction to being treated unfairly by a cooperation partner. They find that sharing a salient group membership prevents such acts of revenge, while vengeance does erupt when group memberships are distinct.

Grossman and van der Weele [13] put the hypothesis under scrutiny that forcing participants to make quick decisions about charitable giving will result in higher donations. Intriguingly, they find no evidence of such an effect. Their study thus casts some doubt on the reliability of a frequently used experimental method and indicates, more generally, that dual-process reasoning may not be as readily observed as is often assumed.

In a cross-cultural study, Hennig-Schmidt and Walkowitz [14] analyze 65 hours of transcribed negotiations over the division of an amount of money between groups with asymmetric disagreement points. Their cornucopia of results offers insights into, among other things, how stated distributive preferences depend both on the parties' outside options and on how those options change over time.

The literature review contributed by de Boer [15], finally, completes our special issue. It offers a selective, while broad, illustrative introduction to the literature on context sensitivities of social preferences.

Acknowledgments: We thank all our authors for choosing this special issue as the place to share their work; we further thank all reviewers involved in this project for their diligence and effort; excellent technical assistance by *Games*' editorial staff is also gratefully acknowledged.

Conflicts of Interest: The guest editors declare no conflict of interest.

References

1. Braithwaite, R.B. *Theory of Games as a Tool for the Moral Philosopher. An Inaugural Lecture Delivered in Cambridge on 2 December 1954*; Cambridge University Press: Cambridge, UK, 1955; ISBN 9780521043076.
2. Cavagnetto, S.; Gahir, B. Game Theory—Its Applications to Ethical Decision Making. *CRIS Bull. Centre Res. Interdiscip. Study* **2014**, 73–91. [CrossRef]
3. Solomon, R.C. Game Theory as a Model for Business and Business Ethics. *Bus. Ethics Q.* **1999**, *9*, 11–19. [CrossRef]
4. Binmore, K. Game Theory and Business Ethics. *Bus. Ethics Q.* **1999**, *9*, 31–35. [CrossRef]

5. Alfano, M.; Loeb, D. Experimental Moral Philosophy. In *The Stanford Encyclopedia of Philosophy*, Fall 2017 ed.; Zalta, E.N., Ed.; Stanford University: Stanford, CA, USA, 2017; Available online: https://plato.stanford.edu/archives/fall2017/entries/experimental-moral/ (accessed on 24 April 2018).
6. Knobe, J. Experimental Philosophy Is Cognitive Science. In *A Companion to Experimental Philosophy*, 1st ed.; Sytsma, J., Buckwalter, W., Eds.; Wiley Blackwell: Chichester, UK, 2016; pp. 37–52. ISBN 9781119099048.
7. Lütge, C.; Rusch, H.; Uhl, M. (Eds.) *Experimental Ethics*; Palgrave Macmillan: Basingstoke, UK, 2014; ISBN 9781137409799.
8. Sarkisian, R. Team Incentives under Moral and Altruistic Preferences: Which Team to Choose? *Games* **2017**, *8*, 37. [CrossRef]
9. Alger, I.; Weibull, J. Strategic Behavior of Moralists and Altruists. *Games* **2017**, *8*, 38. [CrossRef]
10. Van Basshuysen, P. Towards a Fair Distribution Mechanism for Asylum. *Games* **2017**, *8*, 41. [CrossRef]
11. Baumann, F.; Friehe, T. On Adverse Effects of Consumers' Attaching Greater Importance to Firms' Ethical Conduct. *Games* **2017**, *8*, 39. [CrossRef]
12. Della Valle, N.; Ploner, M. Reacting to Unfairness: Group Identity and Dishonest Behavior. *Games* **2017**, *8*, 28. [CrossRef]
13. Grossman, Z.; van der Weele, J. Dual-Process Reasoning in Charitable Giving: Learning from Non-Results. *Games* **2017**, *8*, 36. [CrossRef]
14. Hennig-Schmidt, H.; Walkowitz, G. Moral Entitlements and Aspiration Formation in Asymmetric Bargaining: Experimental Evidence from Germany and China. *Games* **2017**, *8*, 44. [CrossRef]
15. De Boer, J. Social Preferences and Context Sensitivity. *Games* **2017**, *8*, 43. [CrossRef]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).