TECHNISCHE UNIVERSITÄT MÜNCHEN
Peter Löscher-Stiftungslehrstuhl für Wirtschaftsethik
Fakultät für Wirtschaftswissenschaften

Using Experiments in Ethics

An Inquiry into the Dark Side of Competition

Johanna Jauernig

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)

genehmigten Dissertation.

Vorsitzende:	Prof. Dr. Alwine Mohnen
Prüfer der Dissertation:	1. Prof. Dr. Christoph Lütge
	2. Prof. Dr. Michael Kurschilgen

Die Dissertation wurde am 7. Juni 2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 16. August 2017 angenommen.

Acknowledgments

I would like to express my gratitude to my advisor, Christoph Lütge for his open-mindedness, which gave me the chance to engage in this exploratory topic. I also would like to thank him for the inspirational, instructive and eye-opening experiences I was able to gain during my doctoral studies at his chair. I would like to thank Alwine Mohnen and Michael Kurschilgen for taking on the tasks of reviewer and chairman. My thanks go to Matthias Uhl for introducing me to the realm of economic thought and the method of experimental ethics and for his substantial and kind help in many regards. Klaus Abbink has helped me make the research design stronger and more concise and for that I would like to thank him. Cordial thanks go to Hannes Rusch for checking the thesis' content and to Lea-ann Dunbar for making this thesis more readable. I also would like to thank my colleagues, former and present, for inspirational thoughts and a good time: Thanks to Julian Müller, Nikil Mukerji, Jan Gogoll and Michael von Grundherr. Gerhard Vollmer has introduced me to the philosophy of science, which helped me a great deal in writing this thesis and I thank him for that. I would like to very specially thank Wally Eichler for supporting me in manifold ways during the last years. Finally, thanks go to friends and family for their support during the last years.

Abstract

In two experimental studies I investigate detrimental effects of post-competition punishment. Results show that the punishment levels of winners and losers are equally high. An analysis of different target groups shows that losers are more likely to be spared by their peers, whereas winners face punishment from all sides. This provides evidence contrary to the long-standing *frustration aggression hypothesis*. In a second study, a voluntary commitment device with a low implementation probability is introduced. Results show that more than two thirds of subjects use the device and that their commitments convey true intentions. However, subjects' behavior depends on the signal they receive from their counterparts. They do not stick to their commitment if the other did not bind him- or herself. A better understanding of competition's detrimental effects helps us to mitigate these negative effects while still profiting from competition as a means of performance enhancement. Thus, our findings have important implications for competitive settings such as work environments. These experimental results are embedded in theoretical reflection on the use of economic experiments in the field of experimental ethics. This thesis argues that the economic method is especially suited to avoiding methodological issues in experimental philosophy, thus advancing the field.

Contents

Chapter 1

Why We Need Empirical Research in Ethics

The discipline of moral philosophy or ethics is traditionally a theoretical discipline. Especially normative ethics, to which we normally refer when using the term "ethics,"[1] is generally considered a classical armchair discipline. That means that the philosopher deals with ethical problems by means of his own thinking and reasoning. This approach still remains the philosophical mainstream that is explicitly defended by some contemporary philosophers. Timothy Williamson, for example, defends philosophy as an "armchair"-discipline, since no matter if it is approached from a rationalist or empiricist perspective, it can be easily pursued by one's own thinking (Williamson 2005). In this view systematic data gathering and analysis is neither necessary nor desirable. While Williamson acknowledges empirical influences on philosophical reflections, others completely deny the relevance of empirical research to philosophical investigations (Feldman 2010).

In this doctoral thesis I will challenge the armchair position in two ways: The first is a rather mild challenge, to which many ethicists would probably subscribe. It claims that ethics is not always free of the influence of empirical findings and can actually profit from this influence. This claim will be explicated in this chapter. The second challenge is a bit more bold: It claims that ethics needs to use the rather sophisticated method of experiments and especially economic experiments to be able to address certain issues that otherwise could not be effectively analyzed, let alone solved. This claim will be explained in detail in Chapter 2.

The first challenge states that some ethicists either base their theories on empiricism or have been influenced by empirical observations. This can be corroborated with the help of an example: Jeremy Bentham starts his book "Introduction to the Principles of Morals and Legislation," which is the foundation

[1] When defining the sub-disciplines of ethics the term *descriptive ethics* regularly appears. It is, however, not clear to which extent descriptive ethics differs from e.g. anthropology or moral psychology. Furthermore, few researchers call themselves "descriptive ethicists".

of his utilitarian ethics, with an observation on human nature. Chapter 1 begins with the following sentences: "Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne" (Bentham 1879, 1). From this observation he directly derives guidelines that will become the basis of his utilitarian principle which is the "greatest happiness of the greatest number" and at the same time the "measure of right and wrong"(Bentham 1776, A Fragment on Government, preface). Here we almost find a symbiosis of empiricism and moral principles. Even though the observation of nature has a huge influence on Bentham's moral theory, this kind of reasoning about empirical conditions can admittedly still be done from an armchair.

Bentham's example shows that some (moral) philosophers in the past and the present are taking empiricism into consideration. It is arguable that moral philosophers should attempt to integrate more empirical findings into their theories, otherwise they run the risk of missing out on crucial aspects of social reality. In his article "Is empirical research relevant to philosophical conclusion," Erik Angner (2013) argues along this line[2]. Angner dismantles Fred Feldman's argument, which states that empirical research is not relevant to philosophical questions. Therefore, so Feldman, philosophers have little reason to pay attention to the work of psychologists or economists and they do not run the risk of embarrassing themselves by ignoring empirical findings. Analyzing one of Feldman's own arguments about the nature of happiness, Angner shows that even Feldman relies on empirical facts. From the fact that even arguments intended to be purely theoretical rely on empirical work Angner derives the conclusion that an exchange between philosophy and science can provide mutual benefits.

This thesis will argue that ethics can profit from taking empirical findings into consideration. One advantage of integrating empirical research is that it can draw attention to new problems that were not obvious from the armchair position. Empirics sometimes also reveals counter-intuitive findings, which can even override ethical intuitions that have become deeply rooted in common sense. One example is the so called *four-eye principle*, which is said to strengthen compliance. It is therefore frequently applied in corporations, especially where due diligence is required as well as in ethically sensitive contexts, like situations which are prone to bribery. Empirical research has now discovered a new threat which reasoning alone had failed to detect: In some experiments research showed that a four-eye situation, in which each participant has the power to veto in favor of moral conduct, compliance to the moral norm was not strengthened but eroded (Dana et al. 2007). Contrary to the expectation, participants did not check each others behavior, nor did the presence of others pressure participants into norm abidance. On the contrary, it seemed to free them from their responsibility to comply with the norm, because somebody else could be held responsible. This phenomenon has been described as *diffusion of responsibility* or

[2]His arguments regard philosophy as an entire discipline. Therefore his argumentation also applies to the philosophical sub-field of ethics.

moral diffusion.

One of the experimental studies of this thesis provides another example of how empirical findings can address new ethical issues. For many decades the *frustration aggression hypothesis* formulated in 1939 by Dollard and colleagues was used to explain violence as a phenomenon of frustrated losers. Decades later, psychologists Bushman and Baumeister (1998) argued that this hypothesis became conventional view before being sufficiently tested. Findings of our recently conducted study (Jauernig et al. 2016) suggested that in competitive settings it is not just the losers who show elevated levels of aggressive behavior. Generally, winners and losers showed aggression at comparable levels. In some instances, when winners were grouped with other winners and losers with other losers, the winner's aggression stayed high, whereas the loser's aggression declined. This clearly contradicts the *frustration aggression hypothesis.* Knowing from which side aggressive behavior is to be expected is hugely relevant for ethics. For ethical approaches which take factors such as intention or the responsibilities attached to a certain social role into consideration, knowing who issued aggression can be important for moral judgment. A more technical objection comes from psychologist Steven Pinker. He points out that the *frustration aggression hypothesis* has inspired school programs to strengthen self-esteem in children in order to prevent violence. He calls this a "most popular delusion about violence" (Pinker 2011, 510–20). In addition, it could be argued that programs intended to prevent violence, but missing the right target group are a waste of money (often the tax payers') and therefore not ethically justifiable.

The structure of this thesis is as follows: In this chapter, I will attempt to integrate empirically informed ethics in the wider context of ethical approaches. The role of self-interest in providing a foundation for ethics as well as the capability of experimental ethics to address the issue of implementation will be a crucial part in this thesis' argumentation. I will also cast light on how empirics can be integrated into ethics and which problems occur, specifically the *naturalistic fallacy.* Finally in this chapter, I will argue that acknowledging empirical facts can be an ethical endeavor in itself. Chapter 2 focuses on the method of experimental ethics. I will set out what has been done in the field of experimental philosophy (experimental ethics being a sub-field) and illustrate the shortcomings of the methods which have been employed so far. After that I will briefly sum up the method of experimental economics and show how it can enrich research in experimental ethics. I will also address important ethical questions which have been neglected by experimental ethics up to now. The chapter concludes with the three uses of experimental ethics.

In Chapter 3 I will focus on the studies which constitute the main part of my thesis. I will provide an overview on the experimental research on competition, focus on the dark side of competition and show the research gap that occurs as well as how we attempt to address the open questions with our studies. Chapter 4 consists of the two published papers, which are the main part of my doctoral thesis. In Chapter 5 I will summarize the results of my concluded research and point out what further research there is to

be done. I conclude this thesis by embedding the findings into the methodological considerations on experimental ethics. I will attempt to show how the results of the two studies on competition are a step in the direction of advancing experimental ethics with economic method.

1.1 A New Approach to Ethics: Solving the Implementation Issue

The integration of empirical research into ethics may prove to be necessary in order to address one fundamental problem of morality. This problem reaches back into the history of moral thought: How can ethical claims be effective in guiding human interactions? This question has become especially relevant in modern times as will be seen in what follows. Ethics has been based on ultimate justifications: the idea of the good (Plato); divine truth (in the entire Christian ethics); or reason[3].

These ultimate justifications provided certainties that firmly anchored ethics in human consciousness, and were backed up by mechanisms of social sanction within the prevailing small-scale societies. These certainties, however, have been shaken by what Sigmund Freud called the three insults to humanity: First, Copernicus' cosmological insult that the earth is not in the center of the universe, second, Charles Darwin's biological insult that humans evolved from animals and third, Freud's own insult, the theory of the unconscious mind according to which humans can not even control their own psyche (Freud 2005). These revolutionary transformations in the history of thought have caused science to try and regain certainty about what is true and what is moral. In this endeavor, empirical research plays an important role. In the domain of sciences *fallibilism*, which is the rigorous testing of previously formulated hypotheses, became common practice. In this procedure the hypotheses arise from speculation and are subsequently subjected to reality checks. In the realm of philosophy, this methodological revolution of *fallibilism* is mirrored in the critical rationalism of Karl Popper (1980). His fundamental view is that humans are fallible, hence all positions claiming ultimate justification (i.e. dogmata) are to be rejected. Knowledge — just like science — evolves in a trial and error process, therefore it can never be certain. The aim of the trial and error process is to reject what proves to be wrong and to hold on to what has yet to be proven wrong. For ethics this means successively improving the rules in society that prove to be deficient. Experimental ethics comes into play exactly in testing moral theories (or conjectures) against the background of modern reality.

As far as ethics in general concerned, the 20th century is characterized by a series of different attempts to provide morality with a new foundation. Two of these attempts are *discourse ethics* in which ethical truths are generated in an ideal discourse following rules of a rational argument (Habermas 1983) and the renaissance of *contract theory*, in which only those rules apply, to which all individuals would consent in a

[3]Reason was believed to enable humans to see the truth and recognize what is morally right — especially with the rationalists such as Descartes but, of course, also Kant.)

hypothetical contract (originally Hobbes, but in the 20th century taken up by Rawls (1971)). The overarching question of these ethical theories is how morals can be substantiated without strong preconditions, i.e. assumptions that need their own justification. *Discourse ethics* has been criticized for being implicitly based on strong assumptions. The discourse, which serves as the basis for all norm justifications, is itself based on a moral norm, i.e. the assumption that each participant considers their own interests to be equal to the interests of all others. Therefore the discourse already relies on a moral norm and thus can not be the foundation (Hoerster 2014). Rawl's theory bypasses this criticism by means of the *veil of ignorance*: In this hypothetical situation, individuals don't yet know in which position they will end up after the veil is lifted. Therefore they have to take all possible societal positions into consideration.

This results in a principle for the rule setting process, according to which each rule is morally justified if it improves conditions for the ones worst off (*maximin rule*). This principle, however, has been criticized from an economic point of view, since it presumes radical risk aversion of the actors. This extreme risk aversion becomes an artificial precondition, because it cannot be found in humans (Binmore 2005).

One huge advantage of contractarian ethics[4] is that it is based on self-interest, which constitutes a weak and therefore superior precondition (Hoerster 2014, Lütge 2015). An ethics based on self-interest has yet another advantage — it can solve the implementation issue. As Homann has pointed out (Homann 2014) an ethics, which systematically focuses on individual moral conduct, fails to address this issue. Especially in complex modern societies[5] individuals run the risk of being exploited, if they act morally, which is due to the logic of the prisoner's dilemma. In such a dilemmatic situation individuals' rational choices lead to a suboptimal outcome. This outcome can not be changed by an individual, but only by changing the structure of the situation. Therefore appeals to moral conduct are not suitable to solve moral problems. From this starting point, Homann develops the *order ethics* approach, according to which morality must be backed by suitable rules and sanctions in order to prevail. Below this overarching order framework, however, people may pursue their self-interest, in line with Adam Smith's idea of the *invisible hand*.

Central to this order ethical approach is the idea that every good intention is subjected to various empirical conditions which need to be taken into consideration. These conditions influence the scope of action up to a degree of determination and can not be influenced by an individual (Homann 2014). Homann summarizes this relation with a practical syllogism:

(1) Normative principles

(2) Empirical conditions

(3) Normative judgment

If this syllogism is neglected, which means one is to jump directly from the normative principles to

[4] Contractarian ethics are ethics based on contract theory.

[5] Structural feature of modern societies is a large-scale anonymous environment lacking effective face-to-face checks to ensure moral conduct.

normative judgment about concrete issues while neglecting empirical conditions, ethics is bound to collide with the implementation problem at some point. "The road to hell is paved with good intentions," as Saint Bernard of Clairvaux put it in the 12th century. This can be interpreted as a claim to take interdependencies into account, in order to create desired consequences. The case of the fight against the black market for ivory in Africa provides an interesting example for a well intended policy measure having the opposite effect. In 2016, Kenya's government burned over 100 tons of ivory to deliver a statement against elephant poaching. This dramatic shortage of ivory increases prices thus making elephant hunting even more profitable (Biggs et al. 2016). The policy makers did not take the market mechanisms into consideration, which leads to a detrimental effect. Against this background, selling confiscated ivory legally could have better consequences for the elephant population (Benabou and Tirole 2010). This example shows that the contingencies of the world need to be considered in order to effectively translate moral principles into concrete rules. Empirical ethics addresses this issue.

1.2 Integrating Empirics into Ethics: An Ethical Endeavor in Itself

When integrating findings from empirical research disciplines such as psychology, economics or biology into ethics, there are generally two possibilities: Empirical findings can be viewed as a pool of data necessary to deal with specific issues of applied ethics. In order to deal with bioethical question for example — like the boundaries of human embryo research — ethicists need to draw on genetic knowledge. This often translates into interdisciplinary research in which biologists, physicians and ethicists jointly tackle issues which arise with by new technologies. Similar patterns apply to many other fields of research such as media ethics, ethics of science, data ethics or also business ethics, to name just a few. These research fields can be called *hyphen-ethics*. Homann (2014) claims that the ethical foundations of *hyphen-ethics* are essentially theoretical, therefore philosophical ethics (understood here as armchair ethics) upholding sovereign autonomy. In contrast to *hyphen-ethics* Homann proposes a new concept of integrating ethics and empirical research: Homann suggests a deep interdependence between ethics and science, so that empirical findings are incorporated into the DNA of ethics. This way the basic program of ethics in general (and not only with regard to fields of applied ethics) is re-written. He calls this "ethics with economic method" (Homann 2014, 19). It is important to clarify that this approach is not restricted to business ethics and the term "economic" transcends the scope of economics in a traditional sense (meaning the science of financial and product markets). The economic approach means looking at issues from the perspective of individual cost-benefit calculations, it being a matter of preference what individuals consider to be a benefit (in line with Gary Becker 1976). The economic approach to ethics also ensures that the question of implementation is accounted for, since it does not demand from an individual to

act against his or her interest on a long term basis. Individual preferences are much more important in this order-ethical approach than in traditional philosophical ethics, where normative principles very much evolve around the individual moral intention. Therefore empirical investigations about human preferences and human behavior play a much more prominent role.

Furthermore Homann (2014) and Lütge (2014) argue that in our modern world characterized by positive sum games, mutual benefits can theoretically be realized[6]. The paradigm of temperance that is still highly valid in traditional philosophical ethics no longer applies: Trade and exchange can be mutual advantageous, hence competition (as a structure feature of modernity) becomes more moral than distribution, since it improves overall welfare instead of managing scarcity. This welfare gain is a moral good in itself. Neglecting these mechanisms (be it with the best intentions at heart) can thus be a detriment to society. Along this line, I argue that in order to find out what will be a particularly beneficial (and implementable) rule or institution, we as ethicists need to closely and empirically investigate the world around us, i.e. human behavior. Otherwise we fail to provide ethics that is suited to address pressing moral issues and flexible enough to respond to new problems resulting from rapid technological progress.

1.3 "Is" and "Ought" and the Problems Connected with Experimental Ethics

By integrating empirical research in ethics we are confronted with a certain problem: If we derive an "ought" from an "is" we commit a *naturalistic fallacy* (Hume 1978). Normative principles, Hume explicates, can be derived only from normative statements and not from descriptive statements. That means if one looks at the world in a descriptive manor, one can not derive a normative claim from this observation without violating the laws of logic. Once an additional normative principle is introduced, observed behavior can be morally evaluated. As we have seen before, this creates problems because it is the goal of modern ethics to do without normative claims to begin with. To tackle the problem of committing a *naturalistic fallacy*, we have two strategies. The first shows that integrating empirical findings in ethics is nevertheless without alternative and the second consists in an auxiliary construction to bridge the gap between "is" and "ought."

Kant (1788) coined the ethical formula "ought implies can," meaning that an individual can only be obliged to perform an action if he or she is (logically) capable of performing it. Since human capabilities are essentially an empirical matter, this formula reinforces the importance of empiricism in ethics.

In defense of the freedom of science, biologist Bernard Davis twists the *naturalistic fallacy* and warns against what he calls the *moralistic fallacy* (Davis 1978). His argument goes as follows: Science only

[6]This is in contrast to a zero sum society where the gain of the one is the loss of the other.

discovers the realities of nature, it does not create them. By keeping science from this endeavor, these realities may be obscured, but not altered. Or as Davis puts it "...blocking off an area of inquiry on moral grounds fixes our knowledge in that area, it becomes in effect, an illogical effort to derive an 'is' from an 'ought' " (Davies 1978, 390). If science, however, is allowed to do its work, we might learn enough about the world to be enabled to address problems through social institutions. From an ethical point of view, his argument can be interpreted as a claim for a strong empirical component in ethics.

Despite empirics being indispensable for ethics, there is still a gap between normative postulate and the objective message of science. According to Hans Albert (1991), it is essential to bridge the gap and make ethics accessible for scientific critique. He deems this possible with the help of so called *bridge principles*. In his book, he names two of these principles, the "postulate of feasibility" (normative claims must be implementable) and the "postulate of congruency" (normative claims must not contradict scientific findings), but acknowledges that more principles must be found (Albert 1991, 92). Albert claims that a complete separation of the normative and the factual is not possible, but that certain normative values are always intertwined with factual knowledge. There is not one single point in time at which our values can be chosen independently of what we know about the world. New ideas and experiences not only change our cognitive system, but also our values. Thus revised factual beliefs can prove incompatible with ethical convictions. These *bridge principles* can help draw attention to such inconsistencies. From the view of Albert's *critical rationalism*, moral postulates may not be treated as ultimate truths, but as hypotheses which remain subjected to critical objections — one important source of critique being findings from natural and social sciences. I would argue that these findings must be complemented by empirical investigations of ethicists themselves. The accomplishments to-date in this field as well as how to advance experimental ethics are the subject of the following chapter.

Chapter 2

Why Ethics Needs Experiments with Economic Method

At the beginning of Chapter 1, I stated that ethicists are already using empirical findings to inform their theories and argued that dealing with empirics is important for ethics because it helps discover new ethical issues. In traditional ethics, empirical findings may already have been used. These observations, however, often had little authority in and of themselves and, most of the time, were not systematically elicited. In this chapter I will challenge traditional armchair ethics even further by arguing that in some cases it is not sufficient to look at empirical findings, it is not even sufficient to systematically elicit empirical data with the help of vignette studies as was done in the early days of experimental philosophy. I will argue that it is the sophisticated method of experimental economics which is needed to address certain ethical problems.

2.1 Previous Research in Experimental Philosophy

For centuries, philosophy and natural (with the beginning of the modern age increasingly empirical) sciences were not yet clearly distinct from one another and the term *experimental philosophy* was frequently employed. In his survey article, Appiah (2014) gives some examples of classical philosophers simultaneously being natural scientists, one of the most prominent of them being René Descartes. Descartes' theory of how mind and body function together was influenced by both his philosophical and his physiological findings.

At the time when philosophy and the empirical sciences such as physics or biology parted their ways in the 19th century and new disciplines such as psychology came into existence, philosophy became widely viewed as a theoretical endeavor. In the 20th century, however, a new research field called *experimental philosophy* appeared. This new discipline deals with genuinely philosophical questions using the help of

empirical data that are gathered and analyzed. I will hence use the term *experimental philosophy* (X-Phi as it is frequently abbreviated) in this new sense.

The initial trigger for philosophers to engage in empirical work was an emerging interest in folk intuition. For centuries, philosophers were engaged in thought experiments and tested theories with the help of their own intuitions. If a theory was able to withstand all thought experiments it was considered a sign of quality. This methodological use of intuition became especially popular in the realm of analytic philosophy.[1] The first philosopher to actually systematically elicit folk intuition himself was probably the Norwegian Arne Naess (Appiah 2014, Naess 1938), who wanted to find out what ordinary people think about certain philosophical notions. The interviews he conducted with these ordinary people about their notions of truth resulted in a book titled " 'Truth' as conceived by those who are not professional philosophers." This book which appeared in 1938 can be considered as one of the first publications in experimental philosophy in the modern sense. Despite Naess' explicit appeal to philosophers to hence investigate people's intuitions, it was decades later that the young research field gathered momentum fueled by the findings of cognitive sciences and experimental psychology.

Since then a large number of papers on experimental philosophy have been published (on philpapers.org over a thousand are listed). Most of the X-Phi studies have in common that they elicit the intuition of ordinary people, just as Naess suggested. In his survey book on experimental philosophy, Nikil Mukerji (2016) categorizes the most relevant findings of experimental philosophy under the headlines (1)"knowledge", (2)"meaning", (3)"intention" and (4)"freedom of will." Studies in the first category investigate the traditional analysis of knowledge as *justified, true belief* (Ichikawa & Steup 2001), which is under attack not only by analytic philosophy, but also by experimental findings which particularly contribute to the debate about the reliability of epistemic intuitions. In the experimental analysis of meaning (2) laypeople's intuitions about the meaning of specific word classes are elicited. In the philosophical debate on what *intention* (3) means, experimental findings have contributed a great deal. The most relevant finding in this research field, the so called *Knobe effect*, will be explained in detail hereinafter. The debate on *freedom of will* (4) has long since been an essential part of ethics. Findings from experimental ethics (as well as findings from cognitive science and neuroscience) have shaken up the debate. Currently many X-Phi studies are trying to find out whether ordinary people hold a compatibilist view, meaning they consider freedom of will and determinism to be simultaneously valid, or an incompatibilist view meaning freedom of will and determinism cannot prevail at the same time. Results of this ongoing research struggle have important implications for the burden of proof.

I will now illustrate X-Phi research by providing two examples, the *Knobe effect* and the *trolley problem*. Both can be assigned to the subcategory of X-Phi, experimental ethics, which is of relevance for this thesis. Then I will address the methodological shortcomings of X-Phi and show how they can

[1] The structure of the method of analytic philosophy in contrast to experimental philosophy is set out in detail in Mukerji 2016.

be approached with the method of experimental economics. At the end of this chapter I will point out some research fields that are especially relevant a modern ethics that pays regard to institutional implementation, but have not yet been addressed by experimental philosophy. This chapter concludes with an analysis of the three main uses of experimental ethics.

One widely acknowledged study is Joshua Knobe's investigation of the attribution of intention (Knobe 2003). He presented participants with one of two vignettes in which a chairman decided to implement a new program. In both versions the program's implementation earns the company a lot of money, however in one version there are negative environmental side-effects and in the other version there are positive environmental side-effects. In both versions the chairman opts for the implementation of the program and explicitly states that he doesn't care about the environmental side-effects. After being presented with the vignette, participants are asked whether the chairman is harming or helping the environment intentionally. The results were astonishing. The majority of participants attribute intentionality to the actor when the environment is harmed, but not if it is helped. This asymmetry has since caused a great number of (partly experimental) literature with possible explanations. Its relevance for ethics is obvious since the attribution of intention is closely related to the attribution of moral responsibility. From an deontological point of view, intentionality is an important asset for the moral evaluation of an action. Even a consequentialist who evaluates an action by its consequences needs to acknowledge the findings. If the majority of people attribute intention and thus causation in an incoherent way, consequentialist judgment might be biased in other instances too. The *Knobe effect* also shows that laypeople are at least partially influenced by their normative judgments when they evaluate an action as intentional.

Another prominent issue of experimental philosophy is the *trolley problem*. This thought experiment originates from analytic philosophy. One of the first philosophers to extensively deal with the *trolley problem* was moral philosopher Judith Jarvis Thomson (1976). The *trolley problem* deals with the situation of a run-away railway trolley which threatens to kill five workers on the railroad tracks. A bystander could however pull a switch leading the trolley to other tracks where only one worker would be killed (this version of the dilemma is referred to as the *bystander at the switch*). Out of this dilemma arises the question of whether or not it is morally acceptable to pull the switch. In another prominent version of the trolley dilemma, the switch is replaced with a footbridge over the tracks on which a fat man is standing. The question is now, is it morally acceptable to push the fat man onto the tracks to stop the trolley, thus sacrificing him and saving the five workers (*fat man*). Both versions of the trolley problem were taken up by psychologist Marc Hauser, who conducted a large scale study eliciting moral intuitions to corroborate his thesis that humans have a universal moral instinct. The results showed that approximately 90% of people considered it morally acceptable for a bystander to intervene at the switch, but only 10% considered pushing the fat man onto the tracks an acceptable intervention. These findings have since lead to a myriad of experimental studies and theoretical analyses, questioning why a majority of

participants would pull a switch, but would not push the fat man.

The trolley literature, or *trolleyology* as it has been smugly called due to its extreme dimensions, has since dealt with the question which measures are acceptable to promote moral good (e.g. saving people). The so called *doctrine of the double effect*, for instance, explains the intuitional differences by distinguishing between an evil as a side effect and an evil as an intended means to an end (Di Nucci 2014). More recently *trolleyology* has also been used to analyze moral problems related to automated driving (Bonnefon et al. 2015). In experimental economics, the bystander dilemma has been used to group people into rule-minded (pulling the switch is not permitted) and outcome-minded (pulling the switch is permitted). Cornelissen and colleagues (2013) showed that the phenomenon of *moral balancing* is facilitated by an outcome-based mindset, whereas morally consistent behavior can be attributed to a rule-based mind-set. The stability of these mindsets and what implications they have for policy measures promoting ethical behavior remains a subject for further research.

2.2 Shortcomings of Experimental Ethics

In a meta-analysis of papers published on philpapers.org with are classified as "experimental philosophy," Hannes Rusch[2] found that the majority of studies consist of vignettes (64% in his analysis from 2014), while 32% used questionnaires or stimuli and only 2% economic games. Overall, 20% of studies did not use experimental manipulations, thus being strictly spoken observations or surveys. In most of the studies people's assessments are elicited as stated preferences as opposed to an incentive compatible elicitation of revealed preferences. This approach can be challenged. There are various reasons casting doubt on findings of studies, which are non-incentivised and/or use vignettes. These problems can be illustrated with the help of two X-Phi studies, one from the trolley literature and the other from the debate between compatibilists and non-compatibilists.

In the trolley debate some researchers take the view that there is no dilemma between pulling the switch and pushing the fat man which needs solving. They argue rather that neither interference is permitted (Di Nucci 2014, Thomson 2008) and introduce instead another version of the *bystander at the switch* dilemma presenting an option to direct the trolley onto a track where the bystander is standing, thus sacrificing his or herself (*bystander trilemma*). Thomson and others argue that it is not acceptable for the bystander to sacrifice one person in order to save five, if the bystander is not prepared to sacrifice him or herself for the same reason. This theoretical point has been questioned (or in Di Nucci's opinion backed, see Di Nucci 2013, 2014) empirically by presenting subjects with the self-sacrifice option in addition to the option of *bystander at the switch*. This leads to many fewer subjects (40%) choosing the option of sacrificing one person in comparison to the original *bystander at the switch*, in which case the majority

[2]Hannes Rusch presented his findings at the *Experimental Philosophy Group Germany* Conference in Bochum in November 2015. Results have not yet been published.

(90%) opted for sacrificing one person (Huebner and Hauser 2011). For the methodological point I want to make here another finding from the study is very interesting: 38% chose to sacrifice themselves in this hypothetical setting. This surprisingly high number of people willing to sacrifice themselves suggests that there are methodological issues in general with the trolley dilemmas. Self-sacrifice is very much against human nature, i.e. their instinct to survive. Although there are instances of self-sacrifices to save others[3], if the percentage of people willing to sacrifice themselves for others was as high as Huebner and Hauser (2011) suggest, the world would probably a very different place.

On the methodological level there are two issues which detract from the validity of these findings. (1) The study has not been incentivised, therefore it does not safeguard against subject's mere lip service, since the *experimenter demand effect* might lead them to give a morally desired answer. It is doubtful whether subjects would have chosen the sacrifice option if they had had to suffer actual consequence, which in the extreme situation would have been sacrificing their life. We therefore can not rule out the possibility that there is a difference between words and actions, in economic terms "cheap talk." The studies remain questionable even if the counter-argument is presented that these studies are meant to elicit peoples' hypothetical decision, which mirrors their ethical mind-sets: On the one hand if we accept this argument, then we give up any claims to external validity (what is the value of an ethical mindset, if it can not be expected to be reflected in an individual's actions?). On the other hand the internal validity is questionable as well, since we can not rule out that the *experimenter demand effect* might be differently pronounced in various versions. In his study, Di Nucci's (2013) changed the order in which the *bystander-dilemma* and the three-track *bystander trilemma* was presented to the subjects. This caused huge differences in the results. Subjects, so it seems, are quite sensitive to *order effects*, which casts doubt on the internal validity of these kind of trolley experiments. In the absence of incentives subjects might be especially sensitive to moral norms made salient in the vignette.

In the second example that illustrates the short-comings of the most commonly used methods of X-Phi, vignettes are also employed. This is only one example of problems that many vignette studies face. Within the debate on determinism and freedom of will, John Turri (2017) published results to corroborate *natural compatibilism*, meaning that people hold agents morally responsible for their actions even if they were unable to behave differently. Turri's study aims at avoiding the main methodological short-comings of similar prior studies. These short-comings included long and complicated stimuli, provocative stimuli that could produce emotionally interfered results, and lack of assessment as to whether or not subjects understood variables in the intended way. By conducting six experiments with a total of 909 subjects, Turri carefully avoided the above mentioned errors. However, one methodological issue remains casting doubt on the study's findings. I will briefly explain the issue with the help of Turri's first study which

[3]Examples of self-sacrifice can be found in war situations, where soldiers intentionally let themselves be killed in order to save others. In my view, suicide bombing does not count as a self-sacrifice, since according to Islamic ideology suicide is rewarded (the family gets rewarded by the terror organization and the bomber is rewarded in his afterlife). Thus it is not a selfless act.

presents subject's with the following story[4]:

A man promised to deliver a package by 4 pm. He just got on the freeway. Given current traffic conditions, it is physically impossible that he will deliver the package by 4 pm. As a matter of physics, it is literally impossible that he can make it by 4 pm.

After having read the text, participants were asked (1) if the man could still deliver the package by 4 pm. and (2) if the man has the responsibility to deliver the package by 4 pm.? Results[5] showed that a large majority of subjects agreed that the man could not deliver the package on time, however they still held him responsible for the delay. This result is seen as a support for the view that people are natural compatibilists. However, this conclusion could be problematic: The story says nothing about how the man planed his delivery trip. We only know that he made a promise that due to traffic conditions he will not be able to keep it. This creates a huge information gap which can be filled individually by the participants. They could assume that the man made the promise carelessly without taking traffic into consideration. They could also assume that planning was careful, but that the man did not stick to his plan because his priorities changed (maybe he didn't get around to leaving in time). All of these assumptions may lead to the attribution of responsibility for his delay. Of course, we can not know which additional information subjects may assume, but the vignette is sufficiently vague that we can not be sure. Thus, the results may not necessarily support *natural compatibilism*. Even tough subjects admit that in the depicted situation a delivery by 4 pm. is not possible, they may hold the man responsible for an assumed prior action which brought him in the situation of being inevitably late.

I did not choose this study as an example for methodological sloppiness. Precisely because the study is cautiously and carefully designed, it shows that in many vignette studies a certain vagueness remains, reducing the experimenter's control over the independent variables. Where there is vagueness we can not really be certain subjects interpret the text or which aspects trigger their decisions. From literature on the *framing effect* we know that variations in framing can systematically reverse preferences, thus violating rationality requirements of consistency and coherence (Tversky & Kahneman 1985). Rich setting are particularly vulnerable to *framing effects*. To circumvent results to be confounded by *framing effects*, in experimental economics, it is widespread practice to use settings that are neutrally framed but explained in detail in order to have a maximum of control over the variables.

2.3 Enriching Experimental Ethics with Economic Experiments

In the last chapter I presented the main questions investigated by experimental philosophy and explained the methodological problems that occur when using unincentivised vignette studies (the most frequently

[4]He actually uses two different stories as a robustness check, but since results were the same for both studies, I chose one story for reasons of convenience.

[5]The two different stories were not analyzed separately. The analysis collapsed across this factor. This feature of Turri's analysis is, however, not relevant for the point I want to make.

applied method). When experimental philosophers cooperate with researchers from other disciplines, most often the collaboration is with psychologists. For this reason the methodological shortcomings of psychological experiments are to be addressed as well. I will show how X-Phi, and, to an even greater extent, experimental ethics can profit from the method of experimental economics. First, I will address the methodological characteristics of experimental economics, showing how they help to circumvent methodological problems in X-Phi and the overlapping field of experimental (moral) psychology. In a second step, I will draw attention to some important research fields that have been neglected by experimental ethics so far.

A crucial difference between economic and psychological experiments is the aspect of payment. In psychological experiments participants are usually payed a flat fee or receive credits for participating. In economic experiments, monetary incentives are used, whereas in X-Phi there are often no monetary incentives at all (see Knobe 2003). In economics, however, most often subjects' payment is not fixed but rather depends on a clearly stated mechanism usually involving a combination of performance and chance. In most laboratories there is a fixed fee that subjects receive for showing up. The additional performance-based part of the payment is designed so that on average, subjects receive a certain amount of money per hour. The amount depends on the respective laboratory's policy. There are several advantages to this practice: (1) by providing monetary incentives, subjects are compensated for their cognitive effort and thus are likely to make a greater effort. If cognitive effort is considered a scarce resource that people allocate strategically, compensation can ensure that this effort is actually invested in order to avoid making judgment errors, thus variance is also reduced (Hertwig & Ortmann 2001, Smith 1976, Davis & Holt 1993, 25). (2) Incentives also have the advantage of avoiding the "hypothetical bias." Studies have shown that it makes a huge difference, if people make hypothetical choices as opposed to choices with real consequences. A meta-study of Harrison and Rutstroem (2008) revealed that subjects' willingness to donate money to a good cause (e.g. environmental conservation) decreased drastically if they were actually asked to pay the money. True willingness-to-pay is considerably less than the stated willingness-to-pay, which in many cases just turns out to be cheap talk.

Cheap talk is caused by social desirability, which is one aspect of the *experimenter demand effect*. Subjects chose what they think meets a social norm. In the laboratory interaction, the experimenter (unintentionally) takes the role of the "other," whose assumed expectations the subject wants to meet. In the case of ethically relevant decisions, this effect is even more pronounced. Since the subject is not only confronted with social expectations, but also with his or her own (personal) expectations. Most people want to maintain a favorable self-image and choosing what is perceived to be a morally good option is one way to gain self-reassurance about moral character. It remains questionable whether or not monetary incentives completely repress image concerns, both personal and in regard to others. We do, however, have reason to believe that these image concerns are drastically mitigated as shown by studies

like Harrison and Rutstroem's (2008).

Overconfidence is another bias we can mitigate with the implementation of monetary incentives (3). People generally tend to consider their own performance above average[6] (see for example Moore & Healy 2008). Studies have shown that participants rate their performance in a task lower and therefore more realistically if their guess is incentivised (Blavatskyy 2009). It is likely that this more realistic self-assessment can also be translated to the moral domain.

There is a second aspect which characterizes the implementation of economic experiments: the fact that no deception is used. Famous psychological experiments like the Milgram experiment (Milgram 1963) show that deceiving subjects is common practice in experimental social psychology (for an overview of the frequency of deception applied in psychological experiments see Hertwig & Ortmann 2001). The Milgram experiment, in which subjects are lead to believe that a participant is being tortured in an adjoining room, constitutes drastic example of deception. Less extreme examples of deception, for example the *hot sauce paradigm* (Lieberman 1999), in which aggression is determined by the amount of Tabasco one subject allocates to another subject (who does not actually exist), are common practice. In economics, studies using deception can not be published for a couple of reasons. I believe these reasons against deception are also of relevance for experimental ethics, leading to the conclusion that deception should also be avoided in X-Phi. The most important reason for avoiding deception is that it lowers experimental control (1). As I have criticized with regard to vignette studies, vagueness may lead to speculations we can not control. If subjects know that they could be deceived about the procedure of the experiments, the room for speculation about the "real" meaning of the experiment increases drastically. As experimenters, we want subjects to react to incentives and make choices accordingly. As soon as subjects attempt to discern the purpose of the experiment and detect deception (many psychological designs explicitly try to obscure what is being measured), choices get distorted.

There is ample evidence that the use of deception corrupts subject's choices. Subjects that quickly become suspicious of deception may contaminate the data of the experiment in which they are involved. Subjects that learn about deception in the obligatory debriefing may contaminate the data of future experiments. Various studies show that previous experience of deception increases participant's suspicion making data noisy (Krupat and Garonzik 1994), and even causing subjects to lye to the experimenters (Newberry 1973), and ignore events such as the epileptic seizure of a fellow subject, because they assume it to be a charade (MacCoun & Kerr 1987; for an overview see Hertwig & Ortmann 2001). None of these behaviors are desirable in an experiment setting. For this reason economic experiments avoid having a hidden meaning. Instructions explaining the whole procedure of the experiment are presented to all participants. Often these instructions are read aloud in the lab to install common knowledge among subjects.

[6] Johnson & Fowler (2011) even consider overconfidence as an evolutionary stable and advantageous trait.

There is also another good reason why deception should be avoided: It preserves the subject pool as a common resource for all researchers using the associated Lab (2). Hertwig and Ortman (2001) compare the subject pool to a public good which everyone can enjoy. Clear rules establish trust among subjects, re-use of subjects is not a problem, experienced subjects might even be advantageous in some instances, since they are acquainted with abstract tasks. However, even single instances of deception can erode the established trust, thus contaminating the pool.

Both methodological points, the use of incentives and the no-deception policy, provide a maximum of control. At the same time the control of the variables allows for causal inference, which is essential to better understand the underlying mechanism of ethically relevant behavioral phenomena. However, there is no perfect way to elicit an object as unique and variable as human decisions. Therefore, in economics, experiments are subjected to a rigorous examination of internal validity in the peer-review process. Adopting this methodological rigor in developing experimental designs can bring experimental ethics decisively forward.

In addition to the methodological issues, which plead for the use of economic method in experimental ethics, there are also content-related issues. In the quite young discipline of experimental ethics, most of the research questions evolved around individual intuitions regarding ethical dilemmas. Large parts of ethics are yet to be explored from an experimental point of view. In social contexts, many ethical issues can not be effectively addressed on an individual level. Therefore institutional ethics or order ethics claim to tackle moral problems on the level of rules and institutions. Research on how people make decisions in specific social situations is therefore highly relevant for (1) assessing what behavioral phenomena there are, (2) understanding how these phenomena work and (3) testing measures to alter undesirable outcomes.

Questions of social behavior have already been addressed in experimental and behavioral economics as well as in experimental social psychology. These findings, even though often highly relevant for ethics, have not yet been further explored by experimental ethicists. There are only a few very recent attempts to bring together both fields. In a special issue on “Behavioral Business Ethics,” Bernd Irlenbusch and Albena Neschen (2014) show how important experimental economics is for business ethics, illustrating their claim with the help of economic literature on corruption. In the field of economics an increasing amount of research has been published that explicitly deals with moral issues regarding individual moral quality (see e.g. Dana et al. 2007 on the erosion of fairness within a moral wiggle room) or the moral quality of a framework (see e.g. Falk & Czech on the acceptance of detrimental effects on others in bargaining situations). While these articles generated a wide response within the economic community, they have not been systematically explored by philosophers.

It seems that in business ethics research, the method of experimental economics is beginning to be

acknowledged[7]. However, this method can also be used to shed light on much more fundamental ethical phenomena, such as the erosion of trust, preemptive retaliation, or hypocrisy. For all those (arbitrarily picked) issues, empirical ethicists can build on experiments by economists (see Berg et al. 1995 for the trust game, Abbink & de Haan 2014 on the first strike game, Lönnqvist et al. 2014 on hypocrisy). So far there is very little independent research in the field of social and political experimental ethics (or more generally philosophy).

The two studies presented in this thesis shall contribute to bringing together economic experiments and social experimental ethics. Whereas the first study investigates the fundamental behavioral mechanism regarding aggression in competitive settings, the second study uses the findings in a way that is relevant to business ethics by investigating voluntary agreements as a means to mitigate aggression between competitors.

Before I set out the studies in detail in Chapters 3 and 4, it is important to note that experimental ethics shall not be incorporated into experimental economics. While I am strongly in favor of the rigorous adoption of the economic method, this new approach to experimental ethics shall still profit from the conceptual clarity and analytical power of philosophical reasoning. As Appiah puts it: "Neither questionnaires nor brain scans are likely to settle debates between deontologists and consequentialists, or compatibilists and incompatibilists" (Appiah 2014, 20). Still, experimental data can show us which questions are relevant, where the burden of proof lies, and which measures help to mitigate ethically undesired behavior.

2.4 Uses of Experiments in Ethics

In 1986, Alvin Roth formulated the different uses of experimental economics with the frequently quoted triad: speaking to theorists (1), searching for facts (2) and whispering in the ears of prince (3) (Roth 1986). Roth characterizes these three uses as different dialogs: Experimenters can enter into dialog with theorists (1), with each other (2), or with policy makers (3). He remarks, however, that these dialogs often overlap meaning that most experiments simultaneously serve more that one purpose. Following Roth's categorization, I will explore the three different uses of experimental ethics, illustrate these uses with examples of existing research in the field, and show how the field of experimental ethics could be advanced.

Many experiments which have been conducted by philosophical experimentalists from the X-Phi community are covered by the first use. Early studies attemptig to justify the use of experiments in ethics address the non-experimentalist mainstream community of philosophical theorists.[8] One example of

[7] In addition to the mentioned spacial issue of Journal for Business, Economics & Ethics on Behavioral Business Ethics, 15 (3) 2014, a volume with the same title was issued in 2012 (Rutledge) edited by David Cremer and Ann Tenbrunsel.

[8] I intentionally avoided using the term "theoretical philosophers" which would have excluded the sub-disciplines of practical philosophy such as ethics, political philosophy or philosophy of law. Theoretical as well as practical philosophy, however, formulate theories which can be addressed by experiments.

these early studies is the *trolley problem* in which both experimentalists as well as analytical philosophers analyze how people react in moral dilemmas. Ethical theories are in fact a fruitful source for experimental ethics and an exchange can be mutually beneficial. Experimentalists can derive their hypotheses from ethical theories helping them better understand human behavior in ethically relevant situations. These experimental findings can assist in examining theories for their implementability, which, as was explained in Chapter 1, is a crucial issue. In a specific philosophical discourse, experiments eliciting folk intuition can also help determine where the burden of proof lays. An example for such discourse would be the argument between compatibilists and non-compatibilists, in which experimental results are used to help clarify which point of view carries the burden of proof. In the future, experimental ethics should do more than simply analyze ethical theories. Other disciplines such as psychology or evolutionary biology have sophisticated theories at their disposal (moral psychology, for example, has already profited from these theories). Findings from these disciplines need to be integrated into experimental ethics, helping to develop a basis for new ethical experiments. Opening up experimental ethics to other disciplines can help us better understand how humans function as moral actors and, for example, which mindset promotes moral conduct.

Gaining a better understanding of the underlying mechanisms of morally relevant behavior[9] in humans is also a main objective of the second use of experimental ethics — searching for facts. In this young discipline, much of the research is exploratory by nature. Experimental ethicists are only beginning to understand which mindsets or situations presuppose a certain (morally relevant) outcome. Various disciplines such as psychology, sociology, economics, biology, history and of course philosophy can inspire investigations into morally relevant behavior. Our first study (Jauernig et al. 2016) is an example of such an exploratory endeavor. We began by questioning which effects competitive situations have on human behavior and happiness perception. Our conjectures were influenced by several theories and hypotheses from psychology (the *frustration aggression hypotheses*) and economics (*group identity theory*), but due to the lack of (experimental) examination on the research question, no standard theory could be tested.

In the field of experimental philosophy research objectives in ethics have been quite limited in number (see. Mukerji 2016). In my opinion, experimental ethicists could be more adventurous in their search for facts.[10] Just as theories can inform experiments, experimental findings can be the foundation for new theories. Especially in political or social philosophy, of which X-Phi has been quite negligent, experiments can be used to advance the fields.

The third use of experimental ethics, dialog with policy makers, is especially relevant for the neglected fields of political and social philosophy. Some experiments aiming at testing theories or searching for facts

[9] I speak of *morally relevant behavior* to include every decision of an actor which could have an effect on others. I do not limit the use of the term to behavior, which is explicitly intended as moral, as probably deontologists would. In my view the complexities of human interactions and the complexities of the world in which humans interact ask for this wider concept of moral relevance.

[10] As e.g. Vollmer (2011) suggests, science does not only need strict assessment of hypotheses, but also daring in formulating these hypothesis.

might as well have obvious implications for policy makers. The phenomenon of *diffusion of responsibility*, for example, could present interesting exploratory findings that teach us about fundamental human behavior, and at the same time have a huge influence on the set up of corporate processes that involve delicate decisions. Other experiments explicitly aim at testing to which extent a certain rule, mechanism or institution is suited to alter ethically undesired behavior. Our second study (Jauernig et al. 2017) is one example of such an experimental test of an institution. We examined whether or not the introduction of a self-binding device decreases punishment behavior after competition.

As the word "dialog" indicates, ideas for laboratory testing of possible policy measures can also come from corporate practitioners (regarding corporate regulations) or policy makers (regarding public institutions). Despite the fact that it is not easy for practitioners and scientists to find a common language, such a cooperation could prove advantageous in advancing experimental ethics as a discipline that has a impact on human interactions. In experimental economics, extended research exists on corruption (see e.g. Lambsdorff 2012) and tax compliance (see for an overview Torgler 2002). These findings inform us in a very specific way about how our institutions should be shaped. Experimental ethics needs to take such findings into consideration in order to move on from an exclusively theoretical perspective in order to make an impact on society.

Chapter 3

Investigating the Dark Side of Competition

3.1 Concluded Experimental Investigations on Competition

In the wide discipline of economics competition is a well researched key concept. However, the phenomenon of competition has also been studied in other fields such as literature (Richter 2012), sociology (Wetzel 2013), psychology (Garcia et al. 2013) and ethics (Lütge 2014). In addition to theoretical analysis, competition has also been studied experimentally, which is of special interest for this thesis. Laboratory studies investigating competition can be found in economics as well as in psychology. In these laboratory studies, for instance, competition's effect on effort and creativity has been studied. While there seems to be a positive correlation between competition and performance (Bonner et al. 2000), there is mixed evidence about the effect competition has on creativity (see e.g. Erat & Gneezy 2016). An additional research field deals with gender differences in competitive behavior. Research shows that women shy away from competition much more so that men do (Niederle & Vesterlund 2007). Even if women's competence is equal to men's in non-competitive settings, women perform less well than men in competitive situations (Gneezy et al. 2003).

An integral part of experimental research on competition has been devoted to explore the downside of competition. In a series of studies, Harbring and Irlenbusch (2008, 2011) investigated sabotage in contests. They found that people are prone to strike against competitors in order to gain a competitive advantage. Other findings indicate that competition increases dishonesty (Faravelli et al. 2015). In bargaining situations participants become more reckless regarding the damage they impose on uninvolved third parties. This externality increases with the number of bargaining partners (Falk & Czech 2013). Most of the studies investigating the downside of competition have looked at competitors' behavior within the competitive situation. Therefore every decision or action to harm fellow competitors serves the goal

of having an advantage over the others. This behavior appears to be strategic and the rationale behind it is easy to understand from the perspective of the competitor who wants to win.

3.2 A Research Gap: Detrimental Non-strategic Effects of Competition

If we investigate participants' harmful behavior within competition, we must assume strategic motives. We don't know, however, which role spiteful intentions (called "other-regarding preferences" in economic terms), might play. Behavioral preferences such as spite have long been neglected in economic literature, since the concept of spite can not be measured in monetary terms. However, spite seems to provide a psychological utility referred to as the "pleasure of being nasty" (Abbink & Sadrieh 2009). Spiteful behavior in competitive settings can be detrimental to the competitors' perception of fairness. As a result, spiteful behavior can erode adherence to rules and thereby endanger the positive effects of competition on the aggregate level.

In order to be able to distill spiteful motives in competitive situations, strategic elements must be ruled out. This has been done in several psychological studies which indicate that there is, in fact, spiteful behavior, which has nothing to do with a competitive advantage in the "game." In their experiment, Muller and colleagues investigated whether the subjects who were better off or worse off showed more aggressive behavior (Muller et al. 2012). They presented subjects with inscrutable perception tasks and subsequently (and randomly) attributing a score to each subject. This score was either higher or lower in comparison to a partner who did not actually exist. This procedure generated "winners" and "losers." In a second stage, the same subjects engaged in a reaction task lasting for a few rounds. When a subject won a round, he could subject his (presumed) partner to noise via headphones as punishment. In a similar study, the noise punishment was replaced by administering hot sauce (Liebermann et al. 1999). Both ways of measuring aggression are frequently employed as an indicator for aggression in psychological experiments. Results show that randomly picked winners, i.e. the ones better off, were significantly more aggressive than the ones worse off.

In another study, the effects of competition on aggression were investigated with the help of video games (Adachi & Willoughby 2011). Each subject had to engage in a video game that was either violent or competitive. Subsequently aggression was measured with the *hot sauce paradigm*. The researchers found that video game violence was not sufficient to elevate aggressive behavior. Competitive games, however, produced greater levels of aggression.

Both studies indicate that competition sparks spiteful behavior which has nothing to do with gaining a competitive advantage in the "game." In the following studies which are the core of this thesis, non-strategic motives for aggression after competition are dismantled to get a better understanding of the

dark side of competition. In our experimental design, competition was modeled as a one shot interaction of subjects. Punishment was then elicited in a second stage after the outcome of competition was unalterable. This design enabled us rule out strategic elements. In accordance with the paradigm of experimental economics, our design differs from Muller et al. 2012 and Adachi and Willoughby 2011 in two basic ways. First, to introduce competition into the experiment, no deception was used, thus subjects actually competed against each other. Second, punishment, which in psychological terms is often referred to as aggression, was elicited by money burning. Money burning consists of destroying subject's monetary endowment without any personal monetary gain.

Chapter 4

Two Experimental Studies on the Dark Side of Competition

This chapter consists of the two studies which are the core of this doctoral thesis. Since both studies are published, copyrights lie with the respective journals. Therefore, in the following sections, the research questions of each studiy are summarized and links to the articles are provided.

4.1 Study 1: "Competition Induced Punishment of Winners and Losers: Who is the Target?"

In the first study, which was published in the Journal of Economic Psychology (Jauernig et al. 2016, DOI: 10.1016/j.joep.2016.08.004), an experimental work horse was established. The work horse consists of a competitive stage and a money burning stage. The following research questions were addressed using baseline treatment and two experimental manipulations: Do subjects burn each others' money after competition? Do winners or losers punish more? Does punishment persist if target groups are manipulated? Which role do *group effects* play, i.e. does punishment persist if we match winners with other winners and losers with other losers? Does clearly winning or losing have any influence on the punishment behavior of competitors. How is subjects' happiness influenced by post-competition punishment?

4.2 Study 2: "Voluntary Agreements Between Competitors: Trick or Truth?"

The second study was published in the Journal of Business Economics (Jauernig et al. 2017, DOI: 10.1007/s11573-017-0862-8) and uses the same work horse as in Jauernig et al. (2016). In a new treat-

ment, we introduced a stochastically implemented self-binding device between competition stage and the money burning stage. This enabled us to better understand voluntary agreements between competitors. These voluntary agreements are also an important tool of corporate social responsibility. With the introduction of the self-binding device we were able to address the following research questions: Are voluntary commitments between competitors honest or trick signals? Does voluntarily entering into a commitment ultimately predict moral behavior?

Chapter 5

Conclusion

Both studies provided us with a variety of results that will be summed up in this chapter. After that, I will draw attention to some possibilities for further research in investigating the detrimental effects of competition. The last part of the conclusion illustrates how these findings advance the field of experimental ethics.

5.1 Summarized Results of the Studies

At first glance the results of the first study (Jauernig et al. 2016) appear to be non-results. In our baseline treatment comparing post-competition behavior in winners and losers, we found no differences in punishment levels. This finding is in contrast to Muller et al. (2012) and Dollard et al. (1939). We analyzed data from the the target's perspective since we were interested in target groups of post-competition punishment. After competition, 40% of both winners and losers were punished by their former opponent. On average, 20.11 of 100 Experimental Currency Units (ECU) were destroyed. The groups of winners and losers did not significantly differ from each other regarding the percentage of punishers as well as the level of punishment. These results did not change, when a rematching took place after the competition such that a winner could be punished by a loser, against which he or she did not compete, and a loser could be punished by a winner, against which he or she did not compete. This finding excludes one possible explanation for post-competition aggression: personal feud meaning that a winner is punished as a revenge by the loser he defeated in competition and a loser is punished because of a continuing dominance and aggression by the winner, who defeated her in competition. However, the results show that subjects do not punish their direct opponent more or less severely. Therefore, it seems that from the perspective of the respective outgroup, winners and losers were lumped together. This conjecture was tested with another manipulation, where winners were matched with other winners and losers were matched with other losers for the punishment stage. If outgroup effects trigger punishment,

we would expect a significant decrease of punishment levels.

In fact the results presented a differentiated picture. Losers were punished less severely by ingroup than by outgroup members. Winners, however, were punished similarly by ingroup as well as by outgroup members. This result implies that it does not matter for winners whom they face — they are subjected to equally high punishment rates as well as punishment levels. For losers, on the other hand, it does matter who they encounter, since they have less to fear from their peers. This indicates that group effects can not explain the punishment behavior we observe (at least not for winners). Further analyses show that subjects of a given group (winners or losers) are in fact not at all lumped together. Analyzing score differences we found the following results: Losers get punished more severely with winners' increasing score difference. That means that winners who won clearly feel somehow entitled to punish losers more severely. Winners' punishment by losers, however is not predicted by score differences. So losers did not take their own performance in competition into consideration when punishing winners. The whole picture shifts when we look at the ingroup constellations: Within the loser group the clearer a loser was defeated in competition the more he or she punished another loser after the re-matching. As far as winners within their group are concerned, no correlation between score and punishment behavior could be found. Generally these findings suggest that in certain situations the individual performance in competition is decisive for punishment behavior. Therefore group effects can not be the main driver for punishment.

To get a more complete picture of post-competition effects, we asked subjects several questions after decisions were made. Since this part of the investigation was essentially exploratory, we did not previously formulate hypotheses. Nevertheless the results gave us interesting insights. We wanted to find out how subjects' happiness is effected by post-competition punishment and whether subjects anticipate this influence. Therefore subjects were not only asked about their own feelings, but also had to asses their partner's feelings in the punishment stage. Results show that self-assessment of happiness is very well predicted by score difference: the clearer the victory, the happier participants are; the clearer the defeat, the lesser less happy subjects are. These results show that subjects took an intrinsic value from their (good) performance. The binary outcome, being either winner or loser, does not predict happiness very well. The level of punishment is influencing subjects' happiness in a negative way: the higher the punishment, the more happiness is decreased. Overall, punishment exercised a weaker influence on subject's happiness than performance in competition. This is an interesting result since only punishment has an influence on subjects' money.

As far as happiness of their counterparts is concerned, subjects did not guess correctly: They did not anticipate the effect performance in competition exercises on happiness, but perceived the competition outcome as a driver for happiness. As far as the influence punishment levels exercise on happiness is concerned subjects completely overestimated it — by a factor of three.

In the second study (Jauernig et al. 2017) is was our aim to test a measure to mitigate punishment.

Results show that after competition more than two thirds of subjects (74.6%) use a self-binding option, which is implemented with 20% probability. Thus we generate a situation in which subjects expressed their intention not to punish. However, in 80% of cases they still get the chance to punish after they learn about their counterpart's intention (to commit him- or herself or not). Therefore treatments arise endogenously depending on the choice of the subject one is matched with.

The punishment bound subjects inflicted on other bound subjects was negligibly small — only 8% of subjects punished their counterparts. The average punishment level decreased to 1.20 ECU. This result shows that subjects commit themselves with true intentions and not only to trick the opponent into a false sense of security. If we look at the behavior of self-committers who encounter subjects who opted against commitment, the picture looks quite different. Non-committers were punished by committers by 14.72 ECU, which is a significant increase. Overall, results show that self-committers' cooperation depends on the signal they receive from their counterpart.

If we look at behavior of non-committers, results are, again, quite different: 60.9% of non-committers who encountered subjects who did commit themselves punished them by an average of 20.61 ECU. This is significantly higher than the average punishment that committers inflict on other committers (1.20 ECU). Therefore self-commitment seems to be a predictor of moral behavior. In this study the outcome of competition did not influence subjects' propensity to choose the commitment option or not.

5.2 Further Research in the Field

With our mainly exploratory studies we provide a first insight into this new research field of detrimental effects after competition. Therefore many aspects remain to be investigated. In what follows I will outline some ideas for further research.

In our research we define competition as a task against another actor where performance is measured comparatively, ultimately leading to positional results, about which participants are informed. So far we investigated these different aspects as a unity. However, in order to better understand the dark side of competition, we must disentangle the different aspects. In our studies we induced competition by a simple calculation task. This task can be experimentally manipulated. An interesting question is whether or not the degree of competitiveness of participants depends on the nature of the task. In this regard, it is interesting to investigate which kind of tasks subjects find especially important to be good at. One possible hypothesis is that competition gets fiercer in a specific task, if subjects' self-worth hinges on their performance in the respective task.

Further studies can also investigate which influence the presence of an opponent in competition has on subjects' punishment behavior. This can be done by comparing a setting, in which subjects compete against each other to a setting, in which subjects solved tasks against the clock. In the latter setting

only inherent pressure can instigate performance. This question is very interesting from a psychological point of view since it shows to which extent subjects have internalized meritocracy. These findings have implications for institutions in which competition is introduced to enhance performance.

Another aspect worth investigating is the meritocratic aspect itself. Is it my own performance, i.e. being better or worse than others, which triggers aggression? This can be investigated with the help of two different control treatments. First, by introducing noise to the competitive process the feedback about ones own performance could be obscured. This leaves participants to know that there was a competition between subjects, but not how they performed relative to others. The second control treatment eliminates performance completely by randomly assigning the labels "winner" and "loser" to participants. This "lottery" could also help finding out to which extent the attached labels induce punishment behavior.

Overall, shedding light on the aspects of competition which might trigger ethically undesired behavior is a first step to avoiding this detriment. If we can point out the aspect which induces punishment behavior in participants, we might be able to create settings in which the advantages of competition are preserved, whereas detrimental effects can be mitigated.

Another research question, which could be addressed with the help of our working horse comes from psychology. Studies show, that the combination of high power and low status in a subject causes especially high levels of aggression against others (Fast et al. 2012). Fast and colleagues modeled status with the help of attributed roles in a fictitious corporate setting — some participants were "Idea Producers" and the others were "Workers". In our design differences in status can be induced by competition, winners having earned higher status than losers. As in Fast et al. (2012) power can be modeled by systematic outcome control. In our setting this can be implemented by giving only one of the two subjects who competed against each other the chance to punish the other. Thus, the chosen one, just like in a dictator game, holds all the power. Since this experiment would be conducted as a one shot game, he or she had nothing to fear in return. Understanding the interaction of status and power in competitive settings can be an important asset to mitigate detrimental effects — for example in corporations where asymmetric status and power relations are the usual case.

5.3 Lessons Learned for the Field of Experimental Ethics

In the introduction, I argued that the discipline of experimental ethics could profit a great deal from integrating economic experiments into its methodological repertoire. The two studies described here are examples for this endeavor. Jauernig et al. (2016) investigates the punishment behavior after competition. Thereby the main question is against whom the punishment is targeted and how this punishment can be explained. Results show that in general winner's and loser's punishment levels are equally high. This speaks against the prevailing *frustration aggression hypothesis*. Winners' higher punishment in the

ingroup setting provides further evidence against the hypothesis.

In science it is essential to know exactly what the problem is. Economic experiments can help pinpoint these problems as has been shown with regard to the phenomenon of *moral diffusion*. Only if we have profound knowledge about the issues we are dealing with, can we analyze the problem with the appropriate means and hopefully find implementable solutions. As the case of the prevailing *frustration aggression hypothesis* inspiring violence prevention measures shows, lack of evidence let policies for decades to fail as Pinker (2011) pointed out. This case indicates that ethicists should not only rely on evidence from other sciences, they also need to engage in empirical investigations themselves to find out which issues they are exactly dealing with. For that enterprise, economic experiments are especially suited since with the help of these experiments variables can be tightly controlled. Thus, causality between variables can be investigated with the help of different treatments. In our study, we manipulated target groups and kept all other variables constant. Results showed a significant differences between the punishment losers directed against winners and the punishment losers directed against other losers. That way we know that the target groups triggered a change in behavior.

Another advantage of our method is that we did not ask participants about their punishment choices in a hypothetical setting. As is customary in economic experiments, choices had real implications, subjects knew that if they punish their counterpart the respective participant is likely be going home with less money. With this method we avoid the *hypothetical bias* which confounds many X-Phi studies.

Whereas the first study investigates detrimental effects of competition on a fundamental level, aiming to better understand the phenomenon as such, the second study is tied more closely to a context of business ethics. As far as business ethics is concerned, the experimental method is especially suited: Business ethics basically investigates interaction problems in complex large-scale societies. As has been argued in Chapter 1, these problems need to be tackled systematically at the level of rules, since appeals to individual morality are bound to fail due to the logic of the prisoner's dilemma. With the help of behavioral experiments we can not only get a better understanding of these interaction problems, we can also test institutional solutions. For these tests, incentivised choices in the lab are especially important, since they reveal how incentive compatible the respective institution is.

In Jauernig et al. (2017) a policy measure to mitigate competition-induced punishment is introduced and tested. Voluntary agreements between corporate actors are a frequently applied means in Corporate Social Responsibility to commit to moral conduct. The different participating corporations are, however, often competitors to one another. Our experimental design mirrors this setting, which enables us to test the efficacy of self-commitment in competitive environments. Another advantage of the lab is that we can tailor our variables to the features we want to capture. Voluntary agreements rarely consist of completely binding contracts. Law enforcement penalizing any infringement would be way too expensive and in CSR goals there is often a certain vagueness which makes enforcement even harder. What is

more, stakeholders need sufficient knowledge about the corporation's transgression in order to be able to sanction the cooperation. Bottom line, even if a corporation enters a voluntary agreement it still has plenty of chances to circumvent the requirements.

We modeled all these factors with the help of a stochastic self-binding device, committers were bound to their choice with a low probability. This low probability represents a corporate context, in which for instance a Non Governmental Organization (NGO) observes the corporation. The NGO can not completely monitor the corporation. In some instances, however, an NGO might be able to pressure a corporation into compliance (by imposing the threat of a consumer boycott for example). In this case the corporation can not break its commitment without suffering serious damage. Against this background our findings are quite interesting. Results show that this rather weakly binding commitment device suffices in most cases. Participants generally adhered to their commitments. This is especially the case if their interaction partner submitted to the same commitment. These findings show that "soft" institutions can also be quite successful, which is a huge advantage because they save enforcement costs. However, our findings also sound a note of caution: If the other competitor has not expressed his willingness to commit himself, subjects feel no longer bound to their promise. In these constellations punishment increases drastically. This shows that self-binding devices with a low implementation probability work within homogeneous groups of self-committers. In interactions with non-committers we obviously need stronger sanctions to successfully implement new ethical standards.

My doctoral thesis is only a first step in exploring experimental economics from an ethicist's view. Many issues are yet to be investigated. I believe that if experimental findings from other disciplines are integrated into ethics as well as ethicists themselves conduct experiments using the economic method, ethics can close the gap between normative principles and normative judgment in concrete issues as Homann postulated. Thus we as ethicists can provide feasible and implementable solutions for the ethical problems we face.

References

Abbink, K., & Sadrieh, A. (2009). The pleasure of being nasty. Economics Letters, 105(3), 306–308.

Abbink, K., & de Haan, T., 2014: Trust on the Brink of Armageddon: The First-strike Game. European Economic Review 67, 190–196.

Adachi, P. J. C., & Willoughby, T., (2011). The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence? Psychology of Violence 1 (4), 259–274.

Albert, H. (1991). Traktat über kritische Vernunft (Vol. 9). Mohr Siebeck.

Angner, E. (2013). Is Empirical Research Relevant to Philosophical Conclusions?. Res Philosophica, 90(3), 365–385.

Appiah, K. A. (2014). Experimental Philosophy. In Experimental Ethics., 7–25. Palgrave Macmillan UK.

Becker, G. S. (1976). The economic approach to human behavior Chicago: University of Chicago.

Bénabou, R., & Tirole, J. (2010). Individual and corporate social responsibility. Economica, 77(305), 1–19.

Bentham, J. (1879.) An introduction to the principles of morals and legislation. Oxford.

Bentham, J., Burns, J. H., & Hart, H. L. A. (1776). A fragment on government (p. 108). Cambridge: Cambridge University Press.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. Games and economic behavior, 10(1), 122–142.

Biggs, D., Holden, M. H., Braczkowski, A. R., & Possingham, H. P. (2016). Track the impact of Kenya's ivory burn. Nature, 534(7606), 179–179.

Binmore, K. (2005). Natural justice. Oxford University Press.

Blavatskyy, P. R. (2009). Betting on own knowledge: Experimental test of overconfidence. Journal of Risk and Uncertainty, 38(1), 39–49.

Bonnefon, J., Shariff, A., & Rahwan, I. (2015). Autonomous vehicles need experimental ethics: Are we ready for utilitarian cars? Computers and Society.

Bonner, S. E., Hastie, R., Sprinkle, G. B., & Young, S. M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. Journal of Management Accounting Research, 12(1), 19–64.

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence?. Journal of personality and social psychology, 75(1), 219.

Cornelissen, G., Bashshur, M. R., Rode, J., & Le Menestrel, M. (2013). Rules or consequences? The role of ethical mind-sets in moral dynamics. Psychological Science, 24(4), 482–488.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. Economic Theory, 33(1), 67–80.

Davis, D. D., & Holt, C. A. (1993). Experimental economics. Princeton university press.

Davis, B. B. (1978). The moralistic fallacy. Nature, 272, 390.

Di Nucci, E. (2014). Trolleys and double effect in experimental ethics. In Experimental ethics (80–93). Palgrave Macmillan UK.

Di Nucci, E. (2013). Self-sacrifice and the trolley problem. Philosophical Psychology, 26(5), 662–672.

Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., & Sears, R. R. (1939). Frustration and aggression. Yale University Press.

Erat, S., & Gneezy, U. (2016). Incentives for creativity. Experimental Economics, 19(2), 269–280.

Falk, A., & Szech, N. (2013). Morals and markets. Science, 340(6133), 707–711.

Faravelli, M., Friesen, L., & Gangadharan, L. (2015). Selection, tournaments, and dishonesty. Journal of Economic Behavior & Organization, 110, 160–175.

Fast, N. J., Halevy, N., & Galinsky, A. D. (2012). The destructive nature of power without status. Journal of Experimental Social Psychology, 48(1), 391–394.

Feldman, F. (2010). What is this thing called happiness?. OUP Oxford.

Freud, S., & Freud, S. (2005). Eine Schwierigkeit der Psychoanalyse. In Gesammelte Werke: XII: Werke aus den Jahren 1917-1920 (pp. 3–12).

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition a social comparison perspective. Perspectives on Psychological Science, 8(6), 634–650.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. The Quarterly Journal of Economics, 118(3), 1049–1074.

Habermas, J. (1983). Moralbewußtsein und kommunikatives Handeln (Vol. 422). Frankfurt: Suhrkamp.

Harrison, G. W., & Rutström, E. E. (2008). Experimental evidence on the existence of hypothetical bias in value elicitation methods. Handbook of experimental economics results, 1, 752–767.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. Behavioral and Brain Sciences, 24(03), 383–403.

Hoerster, N. (2014). Wie lässt sich Moral begründen? (Vol. 6148). CH Beck.

Hume, D. (1978). 1739. A treatise of human nature. London: John Noon.

Ichikawa, J. J., & Steup, M. (2001). The analysis of knowledge. visited at the 24th of May 2017 on https://plato.stanford.edu/archives/spr2017/entries/knowledge-analysis/.

Irlenbusch, B., & Neschen, A. (2014). Zur Bedeutung der Experimentellen Wirtschaftsforschung fuer die Wirtschaftsethik?Dargestellt an Beitraegen zur Korruptionsforschung. Zeitschrift für Wirtschafts-und Unternehmensethik, 15(3), 303–318.

Harbring, C., & Irlenbusch, B., (2011). Sabotage in Tournaments: Evidence from a Laboratory Experiment. Management Science 57 (4), 611–627.

Harbring, C., & Irlenbusch, B. (2008). How many winners are good to have? Journal of Economic Behavior & Organization, 65(3-4), 682–702.

Homann, K. (2014). Sollen und Können. Grenzen und Bedingungen der Individualmoral, Wien: Ibera.

Huebner, B., & Hauser, M. D. (2011). Moral judgments about altruistic self-sacrifice: When philosophical and folk intuitions clash. Philosophical Psychology, 24(1), 73–94.

Jauernig, J., Uhl, M., & Luetge, C. (2017). Voluntary agreements between competitors: trick or truth?. Journal of Business Economics.

Jauernig, J., Uhl, M., & Luetge, C. (2016). Competition-induced punishment of winners and losers: Who is the target?. Journal of Economic Psychology, 57, 13–25.

Johnson, D. D., & Fowler, J. H. (2011). The evolution of overconfidence. Nature, 477(7364), 317-320.

Kant, I. (1788). Kritik der praktischen Vernunft, Kritik der Urteilskraft. In: Königlich Preußische Akademie der Wissenschaften. Gesammelte Schriften, Berlin. [Akademie Ausgabe: AA V].

Knobe, J. (2003). Intentional action and side effects in ordinary language. Analysis, 63(279), 190–194.

Krupat, E., & Garonzik, R. (1994). Subjects' expectations and the search for alternatives to deception in social psychology. British Journal of Social Psychology, 33(2), 211–222.

Lambsdorff, J. G. (2012). Behavioral and Experimental Economics as a Guidance to Anticorruption. 279?299 in: D. Serra, L. Wantchekon (eds.), New Advances in Experimental Research on Corruption. Research in Experimental Economics.

Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. A. (1999). A hot new way to measure aggression: Hot sauce allocation. Aggressive Behavior, 25(5), 331–348.

Lönnqvist, J.-E., B. Irlenbusch, & G. Walkowitz (2014). Moral hypocrisy: impression management or self-deception? Journal of Experimental Social Psychology 55, 53–62.

Lütge, C. (2014). Ethik des Wettbewerbs: Über Konkurrenz und Moral (Vol. 6159). CH Beck.

Lütge, C. (2015). Order ethics or moral surplus: What holds a society together?. Lexington Books.

MacCoun, R. J., & Kerr, N. L. (1987). Suspicion in the psychological laboratory: Kelman's prophecy revisited. American Psychologist, 42(2), 199a.

Milgram, S. (1963). Behavioral Study of obedience. The Journal of abnormal and social psychology, 67(4), 371.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. Psychological review, 115(2), 502.

Mukerji, N. (2016). Einführung in die experimentelle Philosophie. Wilhelm Fink.

Muller, D., Bushman, B. J., Subra, B., & Ceaux, E. (2012). Are people more aggressive when they are worse off or better off than others?. Social Psychological and Personality Science, 3(6), 754–759.

Naess, A. (1938). "Truth" as Conceived by Those Who Are Not Professional Philosophers (Skrifter Utgitt av Det Norske Videnskaps-Akademi I Oslo Il. Hist.-Filos. Klass 1938 No. 4). Oslo: I Komisjon Hos Jacob Dybwad.

Newberry, B. H. (1973). Truth telling in subjects with information about experiments: Who is being deceived?. Journal of Personality and Social Psychology, 25(3), 369.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. The Quarterly Journal of Economics, 122(3), 1067–1101.

Pinker, S. (2011). The better angels of our nature: The decline of violence in history and its causes. Penguin UK.

Popper, K. R. (1980). Die offene Gesellschaft und ihre Feinde II: Falsche Propheten. Hegel, Marx und die Folgen, 6. Francke Verlag München.

Rawls, J. (1971). A theory of justice. Harvard university press.

Richter, S. (2012). Mensch und Markt. Warum wir den Wettbewerb fürchten und ihn trotzdem brauchen, Murmann.

Roth, A. E. (1986). Laboratory experimentation in economics. Economics and Philosophy, 2(02), 245–273.

Smith, V. L. (1976). Experimental economics: Induced value theory. The American Economic Review, 66(2), 274–279.

Thomson, J. J. (2008). Turning the trolley. Philosophy & Public Affairs, 36(4), 359–374.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. The Monist, 59(2), 204–217.

Torgler, B. (2002). Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. Journal of Economic Surveys, 16(5), 657–683.

Turri, J. (2017). Compatibilism can be natural. Consciousness and Cognition, 51, 68–81.

Tversky, A., & Kahneman, D. (1985). The framing of decisions and the psychology of choice. In Environmental Impact Assessment, Technology Assessment, and Risk Analysis (pp. 107–129). Springer Berlin Heidelberg.

Vollmer, G. (2011): Kühne Vermutungen - strenge Kritik. Lässt sich die naturwissenschaftliche Methode verallgemeinern? In Guardini Stiftung: Trigon, Band 9. Berlin: BWV, 137–153.

Wetzel, D. J. (2013). Soziologie des Wettbewerbs: eine kultur-und wirtschaftssoziologische Analyse der Marktgesellschaft. Springer-Verlag.

Pinker, S. (2011). The better angels of our nature: The decline of violence in history and its causes. Penguin UK.

Williamson, T. (2005, June). I-The Presidential Address: Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking. In Proceedings of the Aristotelian society (Vol. 105, No. 1, pp. 1–23). The Oxford University Press.