The Social Framework of Individual Decisions

570+1 Experiments in (Un)Ethical Behavior

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
im Fach Psychologie

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Philipp Gerlach, B.Sc., M.Phil., M.A.

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät:
Prof. Dr. Bernhard Grimm

Gutachter
1. Prof. Dr. Ralph Hertwig
2. Prof. Dr. Dr. Kimmo Eriksson
3. Prof. Dr. Michaela Gummerum

Tag der Verteidigung: 20. Dezember 2017

For Thor

"Value" is that which one acts to gain and keep, "virtue" is the action by which one gains and keeps it. "Value" presupposes an answer to the question: of value to whom and for what? "Value" presupposes a standard, a purpose and the necessity of action in the face of an alternative. Where there are no alternatives, no values are possible.

Ayn Rand, 1957 (1997, p. 1012)

Der Mensch ist ein mittelmäßiger Egoist: auch der Klügste nimmt seine Gewohnheit wichtiger als seinen Vorteil. [Man is a mediocre egoist: even the most cunning takes his habits more important than his advantage.]

Friedrich Nietzsche, 1888 (1988, p. 118)

Contents

List of Figures

Chapter 4 (continued)

Appendices

LIST OF TABLES

Appendices to Chapter 4

ABSTRACT

English

When and why do people engage in (un)ethical behavior? This dissertation summarizes general theories and synthesizes experimental findings on (non)cooperation, (un)fairness, and (dis)honesty. To this end, Chapter 1 introduces experimental games as a paradigmatic tool for rigorously studying (un)ethical behavior. Exploring 100 meta-analyzed experimental games, Chapter 2 demonstrates that small changes in the framing of context—for example, referring to a social dilemma as a cooperative endeavor versus a competition—can have large and long-lasting effects on the participants' propensity to cooperate. Context framing also shapes beliefs about the cooperative behavior of interaction partner(s) as well as donations in nonstrategic allocation decisions, which are known to correlate with cooperation. Taken together, the results suggest that social norm theories provide a plausible explanation for cooperation in general and specifically, for its sensitivity to context framing. Chapter 3 uses social norm theories to explain why research on experimental games regularly suggests that economics students behave more selfishly than their peers. The concept of social norms is extended to include external sanctions: the willingness to enforce compliance in the form of costly punishment. The results of a relatively new experimental paradigm indicate that economics students and students of other majors were about equally likely to be concerned with fairness and had similar notions of fairness in the monetary allocation task. However, economics students made lower allocations themselves, expected others to make lower allocations, and were less willing to sanction the nonconforming behavior of others. Skepticism mediated their lower allocations, suggesting that the economics students behaved more selfishly because they expected others not to comply with a shared fairness norm. Chapter 4 shows that intrinsic sanctions (psychological costs; e.g., shame and guilt) can be sufficient for ethical behavior to emerge. Meta-analyzes on the basis of 470 experimental games provide answers to many of the ongoing debates on who behaves dishonestly and under what circumstances. The findings show that dishonest behavior depends on both situational factors, such as reward magnitude and externalities, and personal factors, such as gender and age, as well as on the experimental paradigm itself.

German

Warum und unter welchen Umständen verhalten sich Menschen ethisch (in-)korrekt? Die vorliegende Dissertation fasst allgemeine Theorien und experimentelle Befunde (nicht-)kooperativen, (un-)fairen und (un-)ehrlichen Verhaltens zusammen. Kapitel 1 führt hierzu experimentelle Spiele als Paradigma zur rigorosen Untersuchung (un-)ethischen Verhaltens ein. Auf der Basis von 100 meta-analytisch integrierten experimentellen Spielen zeigt Kapitel 2, dass kleine Änderungen in deren kontextuellen Rahmungen—beispielsweise, die Bezeichnung eines sozialen Dilemmas als kooperatives Unterfangen oder als Wettstreit—große und langanhaltende Auswirkungen auf die Kooperationsneigung der Teilnehmer haben kann. Kontextuelle Rahmungen verändern außerdem die Verhaltenserwartungen gegenüber anderen Teilnehmern sowie die Aufteilungen in nicht-strategischen Verteilungssituationen, welche mit Kooperation korrelieren. Zusammengenommen lassen die Ergebnisse Theorien sozialer Normen als plausible Erklärung für Kooperationsverhalten allgemein und spezifisch für den Effekt kontextueller Rahmung erscheinen. Kapitel 3 ergründet, warum Forschungsergebnisse experimenteller Spiele häufig zeigen, dass sich Studierende der Wirtschaftswissenschaften egoistischer als ihre Kommilitonen verhalten. Theorien sozialer Normen werden hierbei aufgegriffen und um externe Sanktionen erweitert, also um die Bereitschaft, Konformität mittels kostspieliger Strafe zu erzwingen. Es wird gezeigt, dass sich Studierende der Wirtschaftswissenschaften und anderer Fächer in ihren Aufteilungsentscheidungen von Geldbeträgen ähnlich häufig mit Fairness beschäftigen und zu ähnlichen Einschätzungen kommen, welche Aufteilung als fair gilt. Studierende der Wirtschaftswissenschaften jedoch teilen weniger großzügig und erwarten dies auch von anderen. Zudem sind sie weniger bereit, als unfair angesehene Aufteilungen zu sanktionieren. Die Ergebnisse deuten darauf hin, dass sich Studierende der Wirtschaftswissenschaften egoistischer verhalten, weil sie nicht daran glauben, dass sich andere Teilnehmer an eine grundsätzlich geteilte Fairnessnorm halten. Kapitel 4 zeigt, dass bereits intrinsische Sanktionen (wie Scham- und Schuldgefühle) ausreichen, damit sich Menschen ethisch korrekt verhalten. Meta-Analysen auf Basis von 470 experimentellen Spielen bieten Antworten zu den zahlreichen aktuellen Debatten, wer sich unter welchen Umständen unehrlich verhält. Es wird gezeigt, dass unehrliches Verhalten von zahlreichen Faktoren abhängt, wie situative Einflüssen (z.B. Anreiz und externe Effekten), persönliche Aspekte (z.B. Geschlecht und Alter), sowie das experimentelle Paradigma selbst.

Preface

This dissertation integrates insights from the social and behavioral sciences. I have always considered the growing isolation of these scientific disciplines as most unfortunate. The social and behavioral science disciplines investigate only different facets of the same overarching themes. Yet, academic inbreeding made their theories unrelated, their insights atomized. Borrowing from other disciplines is often discouraged, even punished. As a result much research is redundant and unconnected. I feel very fortunate to have escaped this bleak aspect of academic life as a member of the vivid Center for Adaptive Rationality. It was thanks to this unique research group that I could flourish as the interdisciplinary generalist I identify myself as.

There many people to whom I am pleased to acknowledge my cordial gratitude. I was fortunate to agonize over my dissertation under the supervision of Ralph Hertwig who encouraged me to dive into meta-analytical techniques. I also greatly appreciate the insightful debates with my co-author Kinneret Teodorescu. Working with "my interns" Mayu Amano, Bastian Jaeger, and Inge ter Laak proved enormously helpful and I gratefully esteem their contributions. I also highly acknowledge Susannah Goss and Anita Todd sifting through my manuscripts as well as Kimmo Eriksson and Michaela Gummerum for refereeing the final product. The partly valuable, partly entertaining, and always cynical comments of Christina Leuker, Alan Tump, and Veronika Zilker enlivened my office days. Illuminating remarks have been thankfully received from many people, including Stefan Herzog, Perke Jacobs, Yaakov Kareev, Tirza Lauterman, Tomás Lejarraga, Lucas Molleman, Thorsten Pachur, Malte Petersen, Timothy J. Pleskac, Georg Sator, and Jan K. Woike. I hesitate to add that all blunders and gaffes remain my own.

On a further note, I wish to extend my gratitude to the generous financial support from the German National Foundation (Studienstiftung) and to the extensive academic support from the Max Planck Institute for Human Development, including its International Research School on the Life Course. I also wish to thank the many experimenters from other institutes who kindly shared their primary data and who answered my daunting questions about their research.

Above all I want to thank Angela Gerlach, my cherished wife and true friend. What I owe to Angela extends far beyond the gratitude of any decent preface writing. Angela gave birth to our son in the early stage of her doctoral studies, just 4 month after work on this

dissertation began. Raising a child while working towards two PhDs meant coordination and cooperation on daily basis. We only managed thanks to Angela's resolute dedication and her great generosity—not to speak of the innumerable minor things, which seem trivial and self-evident. They are not.

Chapter 1

General Introduction

History is rich with examples of ordinary people committing enormous atrocities. In the aftermath of the Second World War, for example, the public was shocked to learn of the exorbitantly unethical behavior average Germans had engaged in. How could innocuous family fathers join paramilitary death squads? How could ordinary middle-aged women guard extermination camps? How could well-mannered 15-year-old boys fight as Nazi guerrillas (e.g., Browning, 2001; Welzer, 2007)?

In the search for answers to these and similar questions, the social and behavioral sciences have launched a massive research program. Over the past 65 years, a great number of experimental and nonexperimental insights have contributed to our understanding of when and why people engage in (un)ethical behavior. Keeping track of the ever-accumulating knowledge seems daunting—if not impossible. Several researchers have therefore promoted the synthesis of existing knowledge over generating highly specialized theories and atomized empirical findings (e.g., Gintis, 2007; Pruitt & Kimmel, 1977; Weber, Kopelman, & Messick, 2004). This dissertation aims at such a synthesis. It integrates general theories and experimental findings from various disciplines—including anthropology, economics, psychology, and sociology. The overall goal is to combine these theories and findings to provide general insights into when and why people behave (un)ethically. To this end, I performed meta-analyses of 570 experiments and conducted one new experiment.

Before discussing the theories and experiments, their findings and implications, this chapter briefly sketches a general framework for how (un)ethical behavior can be studied through relatively rigorous experimental designs: experimental games. It argues that experimental games provide a particularly promising tool for connecting insights from the social and behavioral sciences, which—despite their shared interests in social phenomena, such as cooperation, fairness, and honesty—have worked all too often in parallel. The chapter closes with an overview of the remainder of this dissertation.

The Social Framework of Individual Decisions

Before decisions are made, situations are interpreted. The evaluation of options is not governed by simple "objective" facts. Rather, a decision (e.g., to shoot, to protest, to cooperate) is the result of the interplay of several layers of interpretations—from individual disposition (e.g., skills, beliefs, desires) to overarching social context (e.g., pogroms, wars, social dilemmas). When people approach new situations they typically search their memory for familiar mental models (e.g., exemplars, schemas, scripts). Once applied, mental models help people reduce the complexity of new situations by steering attention and providing meaning to what is observed (Baldwin, 1992; S. T. Fiske & Taylor, 1991, Chapters 4 & 5; Smith & Queller, 2001). Depending on the mental model, dissimilar options may be considered; and similar behaviors may be interpreted differently: One and the same behavior can thus be interpreted as appropriate or inappropriate, required or prohibited, marginal or imperative—depending on the interpretation.

What mental models exist and to what situations they apply is often socially learned (e.g., through imitation, [dis]aproval, explicit sanctions). This is true in particular for mental models that apply to social situations (Berger & Luckmann, 1966).

Social situations are typically not fixed states but rather dynamic processes, characterized by outcome interdependency (one or more people share influence over each other's fates; Kelley & Thibaut, 1978) and development over time (the outcomes of past behavior become the preconditions of future decisions). According to Coleman (1990), this social framework can be captured in the form of a simple diagram (Figure 1.1): Social situations of the past shape an individual's disposition in the present (Step 1 in Figure 1.1; e.g., by providing mental models, through socialization). Individual disposition then shapes the decision-making process, which manifest in concrete behaviors (Step 2). In their aggregated form, individual behaviors form new social situations (Step 3), which serve as the starting point for the process to repeat. To understand how social situations evolve (e.g., change, stabilize, escalate; Step 4) the intermediate steps (i.e., 1–3) must be identified.

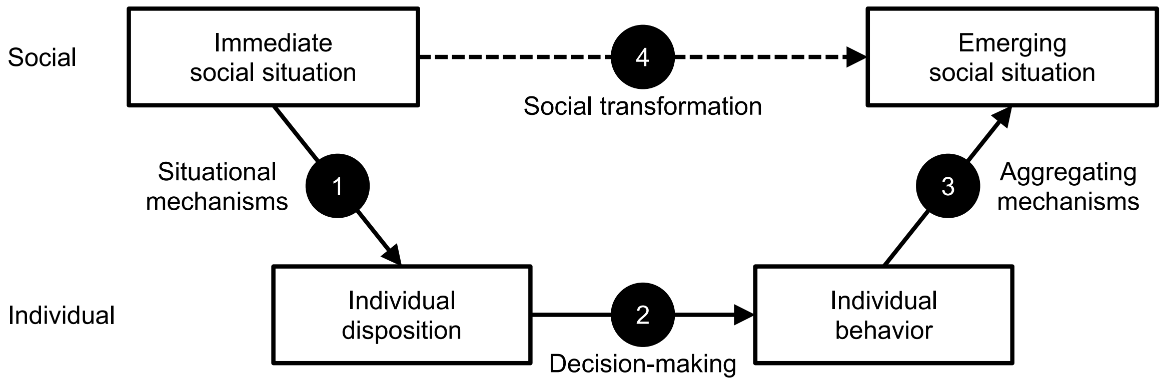


Figure 1.1. *Causal diagram for relating social and individual levels of explanation* (adapted from Coleman, 1990, p. 702). Social transformation (Step 4) must be explained on the basis of three underlying steps (1–3): situational mechanisms, which link social situations to individual disposition; decision-making, which links individual disposition to individual behavior; and aggregating mechanisms, which link individual behavior to social situations.

For example, to understand why the enormous atrocities that seemed unthinkable to ordinary Germans in 1933 became "reasonable" in 1943, the social transformation of Nazi Germany (Step 4) must be dissected as a dynamic social process. In this process the social mechanisms and the individual mechanisms of the social framework coevolve and mutually depend on each other: How did blaming "the" Jews for warmongering and causing economic misery affect the individual dispositions of Germans (Step 1)? How did such propaganda manifest in everyday decisions (Step 2)? And what did the resulting behavior signal to the Nazi elite and other figures (Step 3)?

Explaining facets of this social framework has been at the heart of the behavioral and social sciences. In pursuit of answers, the academic disciplines have adopted a loose division of labor. The behavioral sciences (e.g., psychology, cognitive science) have largely focused on understanding individual-level mechanisms, that is, situational mechanisms (Step 1) and decision-making (Step 2). The social sciences (e.g., economics, sociology), in contrast, have predominantly explored the mechanisms at the social level, that is, aggregating mechanisms (Step 3; e.g., how can individual ethical behavior be detrimental for a group?) and social transformation (Step 4; e.g., how do ethical rules evolve?). Unlike philosophical approaches—which have traditionally relied on introspection and abstract reasoning—the social and behavioral sciences have been interested primarily in understanding and explaining actual human behavior, including the conditions that foster (un)ethical behavior (e.g., when do people engage in compliance, corruption, and cooperation?).

Experimental Games as a Tool for Studying (Un)Ethical Behavior

Many of the social and behavioral scientific insights on when and why people engage in (un)ethical behavior have resulted from controlled experiments (e.g., Bandura, 1965; Darley & Latané, 1968; Milgram, 1974). An increasingly popular (e.g., Chapter 2) and highly rigorous experimental design incorporates experimental games (also known as economic games; Camerer, 2003; Camerer & Fehr, 2002). A paradigmatic example is the investment game (Berg, Dickhaut, & McCabe, 1995), in which a participant, the trustor, is endowed with an amount of money and can decide how much to transfer to another participant, the trustee. The transferred amount is multiplied (e.g., tripled) and becomes the endowment of the trustee. The trustee can then return none, some, or all of this endowment to the trustor. By observing behaviors of in the two roles, researchers can measure generalized trust (in the form of the behavior of the trustor) and generalized trustworthiness (in the form of the behavior of the trustee; e.g., Johnson & Mislin, 2011; Simpson & Eriksson, 2009). Because the trustor's decision hinges on the expected trustworthiness of the trustee (Buchan, Croson, & Solnick, 2008), the trustee's decision to reciprocate has an ethical component: The trustee's reciprocity likely increases the trustor's future transfers in investment-game-like situations—and a lack of reciprocity is likely to be detrimental to future trust.

The appeal of experimental games is not only that they make it possible to measure concrete (un)ethical behaviors—rather than relying on vague or self-reported measures of (un)ethical behavior, such as via surveys. Experimental games are also particularly helpful in studying the mechanisms of (un)ethical behavior because one can focus on particular aspects of the social framework. Chapter 2, for example, aims to explain why situations that involve potential conflict (social dilemmas) but that are framed as cooperative endeavors prompt cooperation (the context framing effect). For instance, referring to a prisoner's dilemma as the "community game" is known to lead to more cooperation than when the same game is referred to as the "Wall Street game" (Gerlach & Jaeger, 2016). Chapter 2 argues that such context frames change individual dispositions (Step 1 in Figure 1.1) in the form of beliefs and/or normative standards. These dispositions then shape individual decision-making (Step 2) because, at least for some people, cooperation is a social norm that hinges on beliefs (what are other individuals going to do?) and normative standards (what is socially desirable?). In their aggregated form, the participants' decisions to cooperate make the entire interacting group profit (Step 3). Therefore, cooperative context frames may prompt interpretations that become true as a result of their behavioral consequences (Step 4; self-fulfilling prophecies; Merton, 1948).

Moreover, experimental games provide a taxonomy of social interaction scenarios, "a rough equivalent of the periodic table of elements in chemistry" (Camerer, 2003, p. 3). Through this taxonomy, scientific insights about human behavior can be formulated in a manner that is decipherable in the various social and behavioral science disciplines (Camerer & Fehr, 2002). Hence, experimental games introduce a common language, which allows the transfer of knowledge between the disciplines.

Another major benefit of experimental games is that they are repeatable encounters. Repeatability manifests in two forms. First, it allows researchers to study the dynamic aspects of social processes within a single experiment and without changing the parameters of the social situation. For instance, the investment game can be iterated over several rounds with all parameters remaining constant. In this way, the process of emerging or collapsing trust and trustworthiness (Step 4 in Figure 1.1) can be observed. Alternatively, iterated social dilemmas games, for example, allow one to analyze the process of emerging or collapsing cooperation (Step 4 in Figure 1.1). Second, the strictly defined parameters of experimental games make it possible to repeat the *same* experimental game in different places, with different population groups, at different times, in different languages, and so forth (unlike natural experiments or quasiexperiments; e.g., Henrich et al., 2001). Thanks to this second form of repeatability, experimental games allow researchers to compare the results of one experimental game to those of another. For example, a finding such as the context framing effect can be verified with a completely different population group. Because situations are highly comparable, quantitative syntheses (Chapter 2 and Chapter 4) can be used to straightforwardly test the reliability of specific findings and assess moderators of the findings, such as whether laboratory studies yield greater context framing effects than online experiments.

Naturally, experimental games also have several limitations. For example, the degree to which unethical behavior can be studied is highly restricted, due to ethical limitations on experiments with human subjects. Experimental games require outlining the possible options for participants and therefore provide limited insights into how mental models affect the awareness of behavioral alternatives. Moreover, experimental games are artificial situations and whether the observed behaviors generalize to the world beyond is an open, empirical question. I highlight more of these limitations throughout the following chapters (see Chapter 5 for a brief summary).

Furthermore, research on (un)ethical behavior with experimental games has also yielded partly inconsistent findings—despite the relatively stringent investigational control of experimental games. The heterogeneity in the findings makes synthesizing experimental results all the more relevant. Doing so, for example, would allow researchers to evaluate the robustness of findings and to clarify the conditions under which specific (un)ethical behaviors are more or less prevalent. Moreover, research with experimental games has generated a wide range of theories, which remained largely unconnected. For example, in their remarks about why people cooperate in social dilemmas (cf. Chapter 2), Smithson and Foddy (1999, p. 14) concluded that "theoretical integration has proven elusive and its prerequisite remain unclear." One reason might be that experimental game researchers have largely ignored theories derived from methods that are less stringent than those in experimental games. Kerr (1995), for example, criticized the ambivalence of experimental game research regarding social norm theories. He argued that social norms are widely recognized to explain cooperation in social dilemmas—even by experimental game researchers themselves. Yet, experiments explicitly testing accounts based on social norms are scarce. As the subsequent chapters point out, relatively general theories—whose origins often lay outside the realm of experimental game research—are indeed helpful for explaining a range of observations in experimental games. For example, social norm theories may help explain why cooperation (Chapter 2) and (dis)honest behavior are largely context dependent (Chapter 4; in particular Appendix 4.4) and even why a number of experimental game studies suggest that economics students behave more selfishly than other students (Chapter 3).

Overview of the Dissertation

In the following chapters I aim at integrating theories from several social and behavioral science disciplines and synthesizing findings from experimental games. The overall goal is to provide broad insights into when and why people behave (un)ethically. Chapter 2, Chapter 3, and Chapter 4 are largely self-contained and can be read in isolation from the remaining chapters. Chapter 5 concludes this dissertation with a summary of the main insights, major criticisms, and common threads contained herein.

In more detail, Chapter 2 assesses when and why people choose to cooperate. To this end, it focuses on cooperative dilemmas, in which there is tension between what is advantageous for the individual and what is good for all. These cooperative (or social) dilemmas are often studied in controlled experiments such as the public goods game or the prisoner's dilemma. Yet small changes in the context framing of these experiments—e.g., referring to a prisoner's dilemma as a "Community game" versus a "Wall Street game"—can have large effects on the participants' propensity to cooperate. To explain context framing effects, Chapter 2 summarizes three classes of theories on why humans cooperate in social dilemmas—social preference theories, group identity theories, and social norms theories—and it explains how each relates to context framing. Then meta-analytic techniques are used to integrate 100 experiments (totaling N = 9,740 participants) that manipulated the context frame. The results suggest that context framing can alter cooperation in the first round, later rounds, and even the last round. Context framing effects are stronger in experiments with flat-fees (than when outcomes are incentivized), in laboratory experiments (than in online experiments), in prisoner's dilemmas (than in public goods games), and in experiments that use priming techniques (than in experiments changing the description of the game). Context framing effects also influence the beliefs about the interaction partner's choices and donations in dictator games, which lack the strategic component of social dilemmas. Chapter 2 concludes that social norms theories provide a plausible explanation for the effects of context framing. It also discusses why context framing is frequently seen as a threat to experimental control and how it presents an opportunity to study the persistence of social preferences.

Chapter 3 extends the notion of cooperation by analyzing when and why people behave fairly. To this end, the concept of social norms is extended by explicit sanctions in the form of costly punishment. Chapter 3 also sheds lights on interpersonal differences in the interpretation of experimental games. In particular, it investigates why economics students behave more selfishly than other students in experimental games. By assessing the underlying motives that drive selfishness, it separates three potential explanatory mechanisms: economics students are less concerned with fairness when making allocation decisions; economics students have a different notion of what is fair in allocations; or economics students are more skeptical about other people's allocations, which in turn makes them less willing to comply with a shared fairness norm. The three mechanisms are then tested by inviting students from various disciplines to participate in a relatively novel experimental game—the third-party punishment game—and asking all participants to give reasons for their choices. Compared with students of other disciplines, economics students were about equally likely to mention fairness in their comments; had a similar notion of what was fair in the situation; however, they expected lower offers, made lower offers, and were less willing to enforce compliance with a fair allocation at a cost to themselves. The economics students' lower expectations mediated their allocation decisions, suggesting that economics students behave more selfishly because they expect others not to comply with the shared fairness norm. This conclusion was triangulated by the lower rates of costly enforcing compliance among economics students.

Chapter 4 shows that implicit and intrinsic sanctions—in the form of psychological costs, such as shame and guilt—can be sufficient for ethical behavior (honesty) to emerge: Over the past decade, a large and growing body of experimental research has analyzed the psychology of dishonest behavior. Yet the findings as to when people engage in (dis)honest behavior are to some extent unclear and even contradictory. A systematic analysis of the factors associated with dishonest behavior thus seems desirable. A meta-analysis reviews four of the most widely used experimental paradigms: sender–receiver games, die-roll tasks, coin-flip tasks, and matrix tasks. Data from 470 experiments are integrated (totaling $N = 30{,}043$ choices) to address many of the ongoing debates on who behaves dishonestly and under which circumstances. The findings show that dishonest behavior depends on both situational factors, such as reward magnitude and externalities, and personal factors, such as the participant's gender, age, and study major. Further, laboratory studies are associated with more dishonesty than field studies. To some extent, the different experimental paradigms come to different conclusions. For example, the rate of liars in die-roll tasks and matrix tasks is relatively similar, but participants in die-roll tasks lie to a considerably greater degree. Also found substantial evidence for publication bias in almost all measures of dishonest behavior is found. Future research on dishonesty would benefit from more representative participant pools and from clarifying why the different experimental paradigms yield different conclusions.

Finally, Chapter 5 briefly discusses the main insights of the previous chapters—especially in light of their theoretical/empirical contributions and their methodological limitations.

Chapter 2

Cooperation Needs Interpretation — A Meta-Analysis on Context Frames in Social Dilemma Games[I]

Cooperation is indispensable to human societies. Unlike other animals, humans frequently manage to cooperate in large groups, even with strangers (Fehr & Gächter, 2002; Trivers, 1971). In modern societies, for example, substantial numbers of people donate blood, vote in general elections, participate in consumer boycotts, and contribute financially to public broadcasting. Although the majority benefits from the existence of public goods—such as blood banks, democratic elections, consumer activism, and public broadcasting services—the individual costs of contribution can be high. These examples belong to a class of situations known as cooperative (or social) dilemmas: situations in which collective and individual gains diverge (Dawes, 1980; Liebrand, 1983). One potential consequence of this divergence of interests is that the public good is not established in the first place (the tragedy of the commons; Hardin, 1968). The question of what motivates people to cooperate, even at substantial costs to themselves, has attracted increasing research attention over the past 50 years and across a wide range of fields, including anthropology, ecology, economics, psychology, sociology, and political science (Figure 2.1).

[I] This chapter is based on: Gerlach, P., Jaeger, B. & Hertwig, R. (2017). Cooperation needs interpretation. A meta-analysis on context frames in social dilemma games. *Manuscript in revision at Psychological Bulletin.*

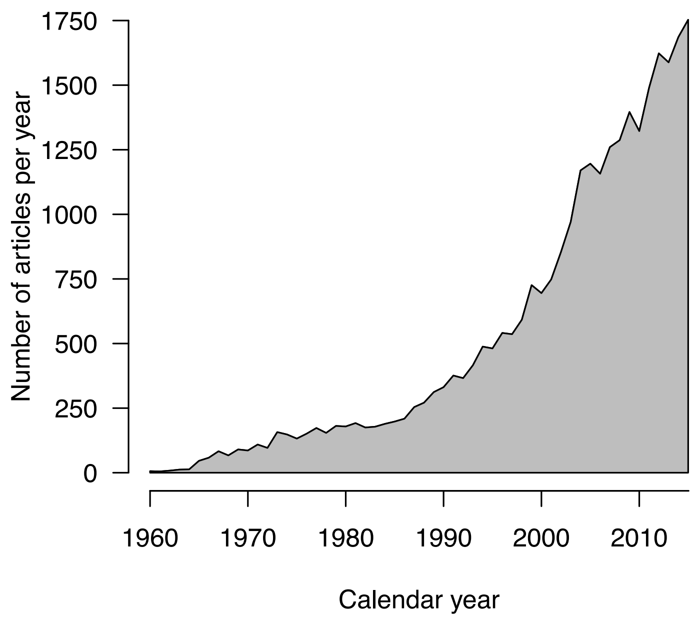


Figure 2.1. *Number of peer-reviewed publications on social dilemmas per year (1960–2015).* Data from an EBSCO Host (ebscohost.com) search for all published articles with one or more of the following keywords in the title or abstract: “social dilemma[s]”, “public good[s]”, “commons dilemma[s]”, and/or “prisoner[s][’][s] dilemma[s]”.

Controlled experiments on allocation decisions, also known as experimental games, offer a rigorous method to investigate public goods provision and other forms of cooperation (Camerer, 2003). Two frequently studied experimental paradigms that simulate social dilemmas are the public goods game and the prisoner’s dilemma. In public goods games, participants simultaneously and anonymously choose how much of their private savings to contribute to a common pool. Private savings are not shared with others and are paid out directly. In contrast, the common pool is multiplied by a factor greater than one and smaller than the number of participants and then equally divided between all participants—regardless of how much they contributed. All participants are thus tempted to privately save money and to profit from other participants’ contributions (Figure 2.2, left panel). In prisoner’s dilemmas, two or more anonymous participants simultaneously choose between cooperation and defection. Mutual cooperation leads to the highest collective gains. One-sided defection increases personal gains but to the detriment of the interaction partner. Thus, participants are tempted to defect (Figure 2.2, right panel). In public goods games and prisoner’s dilemmas alike, purely self-interested individuals will always defect, regardless of how other participants decide. As a consequence, the collective gains diminish.

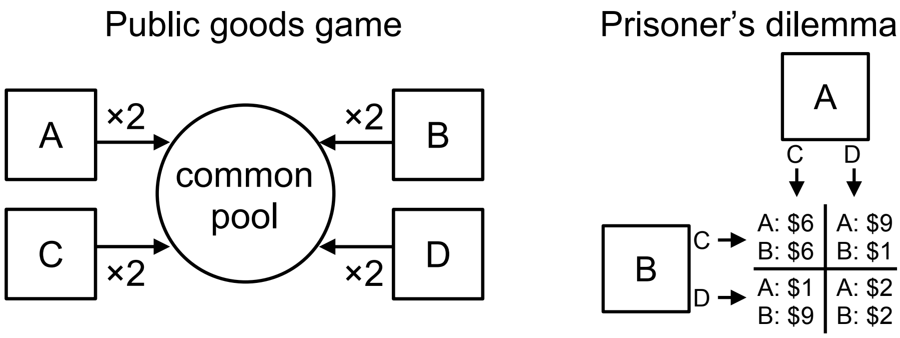


Figure 2.2. *Sample configurations of the two social dilemma games*. Left panel: In this public goods game, participants A, B, C, and D simultaneously decide how much to transfer to the common pool. Transfers are doubled and the common pool is then evenly distributed among all participants. Right panel: In this prisoner's dilemma, participants A and B simultaneously decide to cooperate (C) or defect (D). Mutual cooperation leads to both participants earning $6; mutual defection leads to both participants earning $2. A's one-sided defection earns her $9 to the detriment of B, who earns $1—and vice versa.

Despite the rigorous control that these social dilemma games allow, relatively small changes in the experimental setup are known to have large effects on participants' decisions. Nobel laureate Roth (1995) for example, observed that the "choice an individual makes is sometimes sensitive to the way it is presented, or 'framed,' in the sense that even theoretically equivalent choices may elicit different responses when presented differently" (p. 79). To the extent that researchers have analyzed framing effects in social dilemmas, the focus has been largely on valence framing: a class of situations in which the same information is presented as either *losses* or *gains* (Levin, Schneider, & Gaeth, 1998). In social dilemmas, such valence framing commonly involves the framing of property rights. For instance, public goods games can be framed as involving decisions on how much to either "give" to a common pool or "take" away from an existing common pool (commons dilemma frame). Yet framing can go much further than variation in valence. In the following, we focus on a type of framing that we call *context framing*: the communicative process of associating concepts to situations (e.g., a social dilemma) so that the situation is interpreted in the light of these concepts.[2]

[2] Terms that have been used interchangeably to refer to context framing include cultural framing (Wong & Hong, 2005); institutional framing (Elliott, Hayward, & Canon, 1998); label framing (Dufwenberg, Gächter, & Hennig-Schmidt, 2011); linguistic framing (Banerjee & Chakravarty, 2014); rhetorical framing (Cronk, 2007); semantic framing (Hagen & Hammerstein, 2006); and social framing (J. P. Carpenter, Burks, & Verhoogen, 2005; Dreber, Ellingsen, Johannesson, & Rand, 2013; Ellingsen, Johannesson, Mollerstrom, & Munkhammar, 2012).

For example, experimenters can preface prisoner's dilemmas by emphasizing either the competitive nature of the individual payoffs or the possible group advantage of cooperation. Highlighting one or the other has been shown to shift the cooperation rate from 13% to as much as 89% (Deutsch, 1957, 1958, 1960). Liberman, Samuels, and Ross (2004), who conducted one of the seminal investigations on context framing, found that cooperation rates among their participants doubled when the situation was framed as a "Community game" rather than a "Wall Street game." In a meta-analysis, competitive framing and cooperative framing were identified as two of the three most powerful predictors of cooperation (Sally, 1995).

Although all these investigations, including the meta-analysis, indicate that context framing plays an important role in cooperation decisions, none has clarified the mechanisms underlying this relationship. It is also unclear under which conditions context framing effects occur or when they may be attenuated or amplified. There is also the possibility of publication bias, the selective reporting of only experiments that found significant framing effects. Overall, evidence for context framing is anything but univocal: whereas some studies have reported strong framing effects (e.g., Deutsch, 1957, 1958, 1960; Liberman et al., 2004), others have failed to find any effect or reported mixed results (e.g., Brandts & Schwieren, 2009; Engel & Rand, 2014). Most of the experiments conducted to date had relatively small sample sizes, raising the question of how robust the context framing effects observed really are. In this chapter, we provide a comprehensive overview of studies on context framing and seek to explain the mechanism underlying context framing effects. To this end, we first outline three classes of theories on why humans cooperate in social dilemmas and explain how each relates to context framing. We then use meta-analytic techniques to integrate 100 experiments in which the context frame was manipulated. To explain the heterogeneity in the framing effects observed, we use meta-regression techniques that allow us to identify the moderators of those effects. Finally, we discuss the limitations and theoretical implications of our analysis, with particular regard to experimental practice and the ongoing scientific debate on context framing.

How to Theorize Cooperation and Context Framing Effects

Theories on why humans cooperate in social dilemmas can be classified into at least three broad classes: social preference theories, group identity theories, and social norms theories (for a more conceptual treatment of the three classes of theories, see Appendix 2.1).

Social Preference Theories

Social preference theories—the most common explanations for human cooperation—assume that at least some people are interested in maximizing not only their own payoffs but also, to some degree, the payoffs of their partners, and/or at achieving some relation between their own payoffs and those of their partners (e.g., Andreoni & Miller, 2002; Bolton, 1991; Bolton & Ockenfels, 2000; Dehue, McClintock, & Liebrand, 1993; Fehr & Schmidt, 2006; Levine, 1998; Liebrand & McClintock, 1988; Loewenstein, Thompson, & Bazerman, 1989). Social preference theories commonly assume that these interests are stable and intrinsic characteristics of the individual that are faithfully expressed across time and contexts (Camerer & Thaler, 1995; Loomes, 1999). Although few social preference theories would exclude the possibility that context frames can influence behavior, they do not regularly address context framing effects (for an exception, see Andreoni, 1990, 1995; Andreoni & Miller, 2002). Social preference theories assume that people base their decisions on expected utilities, with description invariance as a bedrock principle: context framing is the "surface structure" whereas payoffs make up the "deep structure" on the basis of which individuals decide (Wagenaar, Keren, & Lichtenstein, 1988; see also Cooper & Kagel, 2003). As long as the deep structure remains intact, choices remain the same. Social preference theories therefore cannot *directly* account for context framing effects.

If beliefs are taken into account, however, then social preference theories can *indirectly* account for context framing effects (Camerer & Fehr, 2002, footnote 6; Dreber et al., 2013; Ellingsen et al., 2012; Fischbacher & Gächter, 2010). Beliefs are individuals' expectations about their partners' choices (so called first-order beliefs). Where do such beliefs come from? In the absence of more reliable knowledge, such as a partner's promise to cooperate, individuals may use external cues to inform their beliefs. Context frames could provide such external cues. For example, conditional cooperators describe people who prefer mutual cooperation to one-sided defection and mutual defection (e.g., Battigalli & Dufwenberg, 2009; Cubitt, Drouvelis, & Gächter, 2011; Fischbacher, Gächter, & Fehr, 2001). In social dilemma games, conditional cooperators are unwilling to cooperate unless they have sufficient reason to believe that their partner(s) will also cooperate. Framing a social dilemma as the "Community game" rather than the "Wall

Street game" could change a conditional cooperator's belief in the likelihood of her partner cooperating and thus influence her decision to cooperate without preferences having changed.

More generally, social preference theories argue that beliefs can become self-fulfilling: If people with beliefs-dependent social preferences interact, and if context framing gives them sufficient reason to change their beliefs, then decisions can also change. In this sense, context frames serve as coordination mechanisms (similar to focal points; see Schelling, 1960).

Group Identity Theories

A competing class of theories argues that social preferences themselves are not cast in stone but flexible and to some degree dependent on the context (e.g. Akerlof & Kranton, 2000; Bacharach, 1999, 2006; Bacharach & Bernasconi, 1997; Balliet, Parks, & Joireman, 2009; Balliet, Wu, & De Dreu, 2014; Brewer & Kramer, 1986; Gold & Sugden, 2007a, 2007b; Kramer & Brewer, 1984; Sugden, 1993, 2000, 2015; Wit & Wilke, 1992). According to this view, individuals first categorize all interaction partners, including themselves, into groups, and then either identify with the group or not. Thus, partners are either categorized as belonging to one's ingroup or outgroup. Group identity theories argue that context frames can steer attention and initiate group identification among people. For example, an emphasis on shared goals might elicit ingroup identification in the form of empathy and therefore prompt individuals to maximize collective instead of individual gains—the latter being at the cost of their partners. Alternatively, an emphasis on conflicting goals might lead individuals to see their partners as outgroup members, prompting them to override their social preferences and maximize their own payoffs rather than collective payoffs.

Social Norms Theories

A third class of theories argues that cooperation is a social norm and the result of a rule-based decision-making process (Bardsley, 2010; Bettenhausen & Murnighan, 1991; Bicchieri, 2006; Bicchieri & Xiao, 2009; Bicchieri & Zhang, 2012; Biel & Thøgersen, 2007; Elster, 1989; Kallgren, Reno, & Cialdini, 2000; von Borgstede, Dahlstrand, & Biel, 1999; Young, 2003). In this process, individuals first categorize situations as an exemplar of a class of situations—for example, "This social dilemma resembles situations of class A." They then rely on behavioral rules associated with that class of situations—say, cooperation in situations of class A. Social norms theories assume that people prefer to follow the rule if two conditions are fulfilled: The individual thinks (1) that others will also follow the rule and (2) that following the rule is what people ought to do (i.e., is socially appropriate or socially desirable). For example, individuals may see mutual cooperation in social dilemmas as the "right thing to do." Context frames can change the categorization of situations such that different behavioral rules are invoked. These behavioral rules in turn trigger different beliefs and/or normative standards.

Two aspects of social norms theories are worth further consideration. First, social norms theories can account for two principal routes through which context framing may affect cooperation: context framing can (1) invoke beliefs and (2) induce changes in normative standards. Changes in either beliefs or normative standards can lead to cooperation. However, beliefs and normative standards are often strongly associated (Eriksson, Strimling, & Coultas, 2014). Second, beliefs are somewhat more broadly defined in social norms theories than they are in social preference theories. In social norms theories, beliefs refer to what a reference group of other people would generally do (Bicchieri, 2006). In social preference theories, beliefs are limited to what a specific assigned partner is actually going to do (first-order beliefs). The definition of beliefs in social norms theories can include the narrow definition of beliefs in social preference theories. We return to these definitions of beliefs in the Discussion. For now, let us conclude that both classes of theories account for framing-induced changes in beliefs about the partner's choices in the context of social dilemma games.

Interim Summary

Social preference theories, group identity theories, and social norms theories offer different accounts for why humans cooperate in social dilemmas. Each class of theory assumes a different mechanism to underlie the relationship between context framing and propensity to cooperate. Social preference theories and social norms theories argue that context frames invoke beliefs, whereas group identity theories are mute about beliefs. Social norms theories account for frame-induced perceptions of normative standards, whereas social preference theories and group identity theories say nothing about the perception of normative standards.

Operationalization

The predicted effects of context framing can be empirically tested. Participants' beliefs about their interaction partner's choices are frequently measured in social dilemma games. In contrast, frame-induced changes in the perceived normative standards are rarely elicited. We therefore used an indirect measure: context framing effects in dictator games (Forsythe, Horowitz, Savin, & Sefton, 1994; see also Kahneman, Knetsch, & Thaler, 1986). In dictator games, one participant decides how to split an amount of money between herself and another participant, then the games ends. This single split is the only transaction in the game and it alone determines the payoffs. Investigations suggest that a person's dictator game donation positively correlates with her propensity to cooperate in social dilemmas (Peysakhovich, Nowak, & Rand, 2014). Nonetheless, dictator games lack the strategic component of social dilemmas. Therefore, first-order beliefs about how the assigned partner will act in this specific case (in the sense of social preference theories) are irrelevant. Social preference theories thus do not account for context framing effects in dictator games. In contrast, group identity theories and social norms theories can both account for context framing effects in dictator games—although they disagree about the underlying mechanism. Whereas group identity theories assume that context frames induce different preferences, social norms theories assume that context frames can induce either changes in beliefs (more broadly defined) and/or changes in the perception of normative standards. Table 2.1 summarizes these suggested mechanisms and the context framing effects accounted for by the three classes of theories.

Table 2.1.

The Three Classes of Theories and the Mechanisms and Framing Effects Accounted for by Each

	Class of theory		
	Social preference	Group identity	Social norms
Mechanism accounted for:			
Beliefs	Yes	—	Yes
Social preferences	—	Yes	—
Perception of normative standards	—	—	Yes
Framing effect accounted for:			
Cooperation (social dilemmas)	Yes	Yes	Yes
Beliefs (social dilemmas)	Yes	—	Yes
Donations (dictator games)	—	Yes	Yes

Methods

Literature Search

To test the three classes of theories and to provide a statistical synthesis of the available literature on context framing, we searched the databases of Google Scholar (scholar.google.com), EBSCO Host (search.ebscohost.com) and Web of Knowledge (apps.webofknowledge.com) in January 2017 for all articles containing the keywords *game[s]* and *social dilemma[s]* in combination with *connotation[s]*, *focal point[s]*, *frame[s]*, *framing*, *game description[s]*, *label[s]*, *label[l]ing*, *metaphor[s]*, *prime[s]*, *priming*, and *salience*. We retrieved the first 2,500 hits per search; the searches covered all available fields (title, abstract, etc.) and all types of articles (published and unpublished) to minimize any bias arising from only significant results being published. We also looked for replications and unpublished manuscripts in PsychFileDrawer (psychfiledrawer.org) and RePEc (repec.org).

Selection Criteria

Overall we screened 7,568 articles. Of these articles, 154 were on context framing. Of these, 57 reported experimental investigations involving social dilemma games and/or dictator games (Figure 2.3).

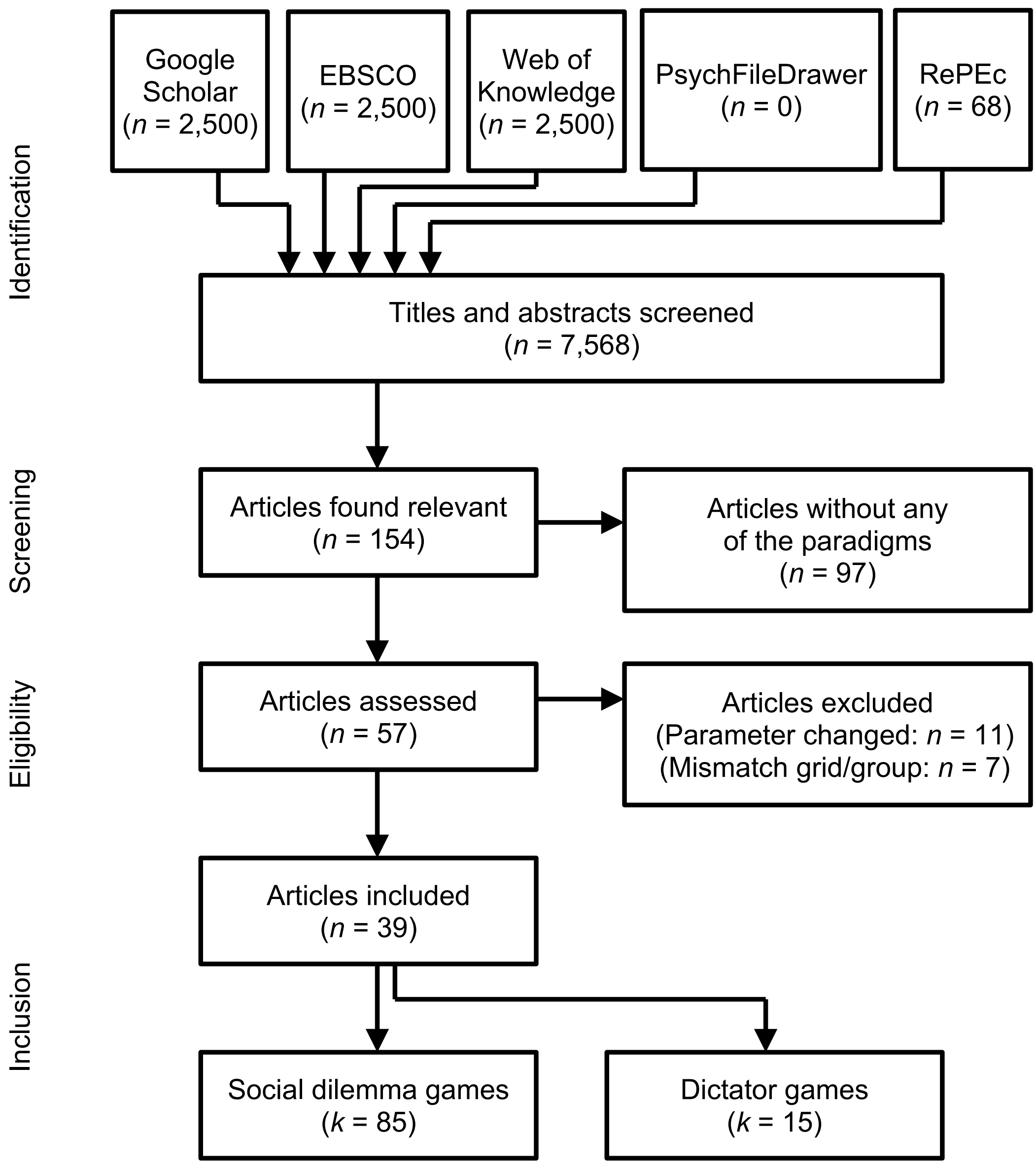


Figure 2.3. *PRISMA flow diagram describing the selection of relevant articles.* n = number of identified articles; k = number of identified experiments.

To ensure that the social dilemma games and dictator games included in our analysis were comparable, we used two inclusion criteria. First, we included only experiments that manipulated the context frame but whose other parameters were exactly the same. For example, payoff conditions had to be similar in the contrasted conditions. Assignment of participants to different parameters in addition to different context frames would have confounded the effect of context framing with the effect of other variables. This criterion led to the exclusion of 11 articles. More details can be found in Appendix 2.2.

Second, we included only experiments with comparable context frames. To obtain comparable context frames two psychology graduate students used grid-group analysis to code all frames (Douglas, 1970; Thompson, Ellis, & Wildavsky, 1990). Grid-group analysis is a general framework to classify social situations according to two dimensions: the group dimension describes the extent to which the individual is absorbed and sustained by group membership; the grid dimension refers to the degree to which the individual's behavior is circumscribed by prohibitions and prescriptions. On the basis of these two dimensions, all context frames were classified as referring to one of four concepts: *cooperative* frames (high group) referred to the individual's common interests; *competitive* frames (low group) referred to discord and rivalry among the individuals; *moral* frames (high grid) referred to how one ought to behave, for example, in terms of fairness, religious appeals, and responsibility for each other's payoffs; and *generic* frames (low grid) referred to concepts that lacked any such moral connotation. Generic frames were the supposedly "neutral," "abstract," "baseline," "control," "clean," "individualistic," "non-associative," "non-framed," "no label," "no frame," "standard," and "unprimed" conditions. The two independent coders agreed on the categorization of all but one frame (Cohen's κ = 0.99). Further, they decided to drop six articles because the frames could not be classified according to grid-group analysis. In sum, seven articles were excluded from further analysis at this stage (cf. Appendix 2.2).

If the above criteria were fulfilled but data needed for the meta-analysis were not reported in the article, we generated the data from the figures or—if this was not possible—contacted the corresponding authors. Overall, our inclusion criteria resulted in a pool of 39 articles (with N = 9,740 observations) reporting a total of 100 primary experiments: 85 social dilemma games (n = 7,721) and 15 dictator games (n = 2,019). All experiments involved samples of healthy adults, windfall money, and random assignment to the context frames. A tabular overview of all experiments integrated is provided in Appendix 2.3. The full database is available online at https://osf.io/27u8y/.

Coding

Context frames were implemented in several different ways. Some experiments implemented framing via the *title*, referring to the dilemma as, for instance, a "Community game" versus a "Wall Street game" (Liberman et al., 2004). Other experiments framed the *options* differently. For instance, participants chose between putting money in envelopes labeled "mine" and "community box" versus "free rider" and "community box" (Torsvik, Molander, Tjøtta, & Kobbeltvedt, 2011). Some experiments *embedded* dilemmas in coherent stories. Here, participants first read an extensive narrative and then faced a choice situation that was linked to the plot. For instance, "investment managers" decided about "investing" in a joint "stock," while "representatives" decided about "contributing" to a "social event" (Pillutla & Chen, 1999). Other experiments *added statements* to the instructions. For example, a public goods game either included the statement that the research was "about community co-management of protected areas" or did not include this statement (Bouma & Ansink, 2013). A subtler source of context framing is *priming*, a technique used to implicitly activate mental representations associated with the displayed stimuli. Primed contexts were never directly related to the description of the social dilemma itself. Instead, they resulted from a supposedly unrelated task, such as constructing grammatically correct sentences from randomly ordered words with moral content.

We predicted that, relative to generic frames (low grid), cooperative frames (high group) and moral frames (high grid) would enhance cooperation, whereas competitive frames (low group) would reduce cooperation. We expected to observe the purest context framing effects in the *first* round of the social dilemmas, in which the partner's choices had not yet been experienced. Our analyses therefore focused on the first round. To assess the persistence of context framing effects, we also analyzed *all* rounds of repeated social dilemmas as well as the *last* round of repeated social dilemmas.

Calculation of Effect Sizes

To enable quantitative integration, we calculated the mean contribution in each treatment as the fraction of the maximum possible contribution. Mean contributions could range from 0% to 100%, with 0% indicating that nobody cooperated or transferred anything and 100% indicating that everyone cooperated or transferred the full amount available. Let frame 1 be the context frame whose mean cooperation or transfer rate $\overline{C}_1$ was contrasted with the mean cooperation or transfer rate $\overline{C}_2$ of frame 2. Further, let SD_1 be the standard deviation and n_1 the sample size of frame 1 and SD_2 be the standard deviation and n_2 the sample size of frame 2. The effect size D represents the average context framing effect in percentage points, calculated as the mean difference between $\overline{C}_2$ and $\overline{C}_1$:

$$D = \overline{C}_2 - \overline{C}_1 \tag{2.1}$$

Its variance V_D is given by

$$V_D = \frac{SD_2^2}{n_2} + \frac{SD_1^2}{n_1} \tag{2.2}$$

We estimated the overall effect sizes by means of random effects models and 95% confidence intervals (95% CIs) per contrast. *Combined* random effects represent the weighted mean effect size across all contrasts.

Estimation of Heterogeneity

Two indicators of heterogeneity were used to examine variations in the effect size distributions. The I^2 statistic is the between-study heterogeneity independent of the number of experiments (Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006). Alternatively, between-study heterogeneity can be estimated using the τ^2 statistic and Q tests to assess the null hypothesis of homogeneity among the effect sizes, $\tau^2 = 0$. Rejection of the null hypothesis suggests that sampling error alone accounts for the difference between the effect sizes.

Evaluation of Robustness

To identify and counteract the risk of selective reporting of significant results only (publication bias), we employed the trim-and-fill correction method. This iterative algorithm adds studies to the analyses until the observations are symmetrically distributed around the average effect size (Duval & Tweedie, 2000). Assuming symmetry may, however, be inappropriate, as heterogeneity between the studies can reflect the "true" effect (Schwarzer, Carpenter, & Rücker, 2010; Terrin, Schmid, Lau, & Olkin, 2003). To assess the robustness of framing effects, we therefore also calculated Copas selection models as a parametric statistical alternative to trim-and-fill (Copas & Shi, 2000, 2001; for comparisons of the two methods: Rücker, Carpenter, & Schwarzer, 2011).

For all analyses, we used the statistical software R (R Development Core Team, 2008) and the default sensitivity parameters of the package meta (Schwarzer, 2007) and its add-on copas (J. Carpenter, Rücker, & Schwarzer, 2009).

RESULTS

Main Effects

First round. In the first round, every contrast yielded a context framing effect in the predicted direction (Figure 2.4). On average, framing effects increased the cooperation rate by D = 9%, from C = 42% to C = 51% (Table 2.2). Overall, there was substantial heterogeneity across experiments (I^2 = 66%, τ^2 = 95, p < .001), with the magnitude of the framing effects varying strongly between the primary experiments: the context framing effect ranged from D = 77% in the predicted direction to D = −21% in the opposite direction (Figure 2.4). We return to the reasons for the between-study heterogeneity in the section titled Moderator Analyses. When combined, using the classification of the grid-group analysis, the average framing effect ranged from D = 6% (generic vs. moral) to D = 36% (competitive vs. moral; Table 2.2). Competitive vs. moral frames led to substantially greater framing effects than any other contrast (all between group comparisons: $Q(1)$ > 9.54, p < .003). None of the other contrasts differed in the magnitude of the context framing effect (all between group comparisons: $Q(1)$ < 2.94, p > .085).

Table 2.2.

Average Rates of Cooperation in the First Round per Contrast

Contrast (Frame 1 vs. frame 2)	k	Cooperation rate: Frame 1	Cooperation rate: Frame 2	D	z	p
Competitive vs. cooperative	29	41.6%	52.1%	10.8%	3.85	< .001
Competitive vs. generic	10	35.1%	48.1%	12.6%	3.90	< .001
Competitive vs. moral	1	32.2%	68.2%	36.0%	5.26	< .001
Generic vs. cooperative	28	43.4%	50.9%	7.5%	2.88	.004
Generic vs. moral	11	43.3%	49.0%	6.2%	3.25	.001
Combined random effect	78	41.7%	51.0%	9.2%	6.13	< .001

Note. For experiments with a larger standard error, smaller weights were applied in the calculation of D. Therefore the difference in the cooperation rates between Frame 1 and Frame 2 need not be equal to D.

Frame 1 vs. frame 2	Frame type	Framed content	Frame 1 n	Frame 1 C	Frame 1 (SD)	Frame 2 n	Frame 2 C	Frame 2 (SD)	D in %	95% CI	Weight
Competitive vs. cooperative											
Deu60b vs. Deu60a	Embedding	"Win as much money as you can" vs. "consider yourself to be partners"	16	12.5%	(34.2)	23	89.1%	(34.4)	76.6	[54.8; 98.5]	1.0%
Deu60d vs. Deu60c	Embedding	"Win as much money as you can" vs. "consider yourself to be partners"	21	23.8%	(43.6)	17	76.5%	(43.7)	52.7	[24.7; 80.6]	0.8%
Lib04a	Title	"Wall Street game" vs. "community game"	24	33.3%	(48.2)	24	70.8%	(46.4)	37.5	[10.7; 64.3]	0.8%
Lib04b	Title	"Market game" vs. "community game"	20	20.0%	(41.0)	20	55.0%	(51.0)	35.0	[6.3; 63.7]	0.7%
Bat99b	Embedding	"Business transaction" between "parties" vs. "social exchange" between "participants"	10	20.0%	(42.2)	10	50.0%	(52.7)	30.0	[−11.8; 71.8]	0.4%
Bat99a	Embedding	"Business transaction" between "parties" vs. "social exchange" between "participants"	10	20.0%	(42.2)	10	40.0%	(51.6)	20.0	[−21.3; 61.3]	0.4%
Eng14c vs. Eng14b	Embedding	Participants are "competing" vs. form a "team" to "successfully coordinate"	52	57.3%	(49.9)	50	76.8%	(43.1)	19.5	[1.4; 37.6]	1.3%
Ber14e vs. Ber14d	Title	"Wall Street game" vs. "Community game"	24	37.2%	(34.7)	24	56.0%	(35.4)	18.8	[−1.1; 38.6]	1.1%
Ell12	Title	"Stock market game" vs. "community game"	110	26.4%	(44.3)	118	44.9%	(50.0)	18.6	[6.3; 30.8]	1.7%
Eng14c vs. Eng14a	Embedding	Participants are "competing" vs. form a "team" to "protect itself against damage"	52	57.3%	(49.9)	52	72.8%	(44.8)	15.5	[−2.7; 33.8]	1.2%
Bra09i vs. Bra09h	Title	"Stock−exchange game" vs. "Community game"	10	22.8%	(26.1)	10	33.8%	(29.9)	11.0	[−13.6; 35.6]	0.9%
Bat99c	Embedding	"Business transaction" between "parties" vs. "social exchange" between "participants"	10	70.0%	(48.3)	10	80.0%	(42.2)	10.0	[−29.7; 49.7]	0.4%
Eng14f vs. Eng14d	Embedding	Participants are "competing" vs. form a "team" to "protect itself against damage"	50	38.6%	(49.0)	51	48.0%	(50.5)	9.4	[−10.0; 28.8]	1.2%
Haa15b	Title	"Banker game" vs. "community game"	18	33.3%	(48.5)	33	42.4%	(50.2)	9.1	[−19.1; 37.3]	0.8%
Ran14e	Embedding	Participants "compete with ... opponents" vs. "contribute ... to common project"	218	53.8%	(38.8)	210	62.5%	(40.0)	8.7	[1.2; 16.2]	2.1%
Dre13	Title	"Profit game" vs. "community game"	202	57.9%	(49.5)	198	65.2%	(47.8)	7.2	[−2.3; 16.8]	1.9%
Pil99	Embedding	"Investing" in an "investment corporation" vs. "donating" to "a social event"	97	32.1%	(24.9)	97	39.1%	(24.9)	7.0	[−0.1; 14.0]	2.1%
Haa15c	Title	"Banker game" vs. "community game"	28	16.6%	(39.0)	32	23.1%	(42.0)	6.4	[−14.1; 26.9]	1.1%
Haa15e	Title	"Banker game" vs. "community game"	25	40.0%	(50.0)	22	45.5%	(51.0)	5.5	[−23.5; 34.4]	0.7%
Con14b	Embedding	Participants "compete with ... opponents" vs. "contribute ... to common project"	175	54.8%	(41.6)	186	59.2%	(40.7)	4.4	[−4.1; 12.9]	2.0%
Ran14d	Embedding	Participants "compete with ... opponents" vs. "contribute ... to common project"	148	61.7%	(41.3)	168	65.8%	(41.1)	4.1	[−5.0; 13.2]	1.9%
Eng14f vs. Eng14e	Embedding	Participants are "competing" vs. form a "team" to "successfully coordinate"	50	38.6%	(49.0)	50	38.6%	(49.0)	0.0	[−19.2; 19.2]	1.2%
Bra09c vs. Bra09b	Title	"Stock−exchange game" vs. "Community game"	26	33.3%	(13.7)	20	32.9%	(16.6)	−0.4	[−9.3; 8.6]	1.9%
Haa15d	Title	"Banker game" vs. "community game"	27	32.1%	(48.0)	34	31.4%	(47.5)	−0.7	[−24.8; 23.4]	0.9%
Haa15a	Title	"Banker game" vs. "community game"	26	32.8%	(48.5)	26	32.1%	(47.1)	−0.7	[−26.7; 25.3]	0.8%
Ber14b vs. Ber14a	Title	"Wall Street game" vs. "Community game"	12	57.4%	(27.9)	12	55.9%	(37.3)	−1.5	[−27.8; 24.9]	0.8%
Haa15f	Title	"Banker game" vs. "community game"	17	51.0%	(51.4)	30	43.3%	(50.4)	−7.6	[−38.0; 22.7]	0.7%
Con14a	Embedding	Participants "compete with ... opponents" vs. "contribute ... to common project"	204	64.9%	(37.0)	186	55.6%	(39.2)	−9.3	[−16.9; −1.7]	2.1%
Bra09g vs. Bra09f	Title	"Stock−exchange game" vs. "Community game"	10	27.4%	(31.5)	12	14.5%	(14.9)	−12.9	[−34.1; 8.4]	1.1%
Random effect			**1692**			**1735**			**10.8**	**[5.3; 16.4]**	**34.1%**
Heterogeneity: $I^2 = 70.3\%$, $\tau^2 = 126.9$, $p < 0.01$											
Competitive vs. generic											
Bra09i	Title	"Stock−exchange game"	10	22.8%	(26.1)	6	55.1%	(36.0)	32.3	[−0.8; 65.3]	0.6%
Sme03b	Priming	Scrambled sentence task with might−related words, e.g., "assertive", "dominant"	25	32.2%	(17.0)	17	57.2%	(25.0)	25.0	[11.4; 38.6]	1.6%
Deu60b	Embedding	"Win as much money as you can" and "do better than the other"	16	12.5%	(34.2)	26	35.9%	(48.5)	23.4	[−1.7; 48.5]	0.9%
Deu60d	Embedding	"Win as much money as you can" and "do better than the other"	21	23.8%	(43.6)	11	41.2%	(52.2)	17.4	[−18.7; 53.4]	0.5%
Eng14c	Embedding	Participants are "competing with another ... in a market"	52	57.3%	(49.9)	20	74.3%	(44.4)	17.0	[−6.7; 40.7]	0.9%
Eng14f	Embedding	Participants are "competing with another ... in a market"	50	38.6%	(49.0)	20	53.5%	(51.0)	14.9	[−11.2; 41.1]	0.8%
Bra09c	Title	"Stock−exchange game"	26	33.3%	(13.7)	17	42.5%	(13.7)	9.2	[0.9; 17.6]	2.0%
Ber14e	Title	"Wall Street game"	24	37.2%	(34.7)	15	41.6%	(29.1)	4.4	[−15.8; 24.7]	1.1%
Bra09g	Title	"Stock−exchange game"	10	27.4%	(31.5)	16	28.8%	(23.8)	1.4	[−21.4; 24.1]	1.0%
Ber14b	Title	"Wall Street game"	12	57.4%	(27.9)	15	52.4%	(29.9)	−5.0	[−26.9; 16.9]	1.0%
Random effect			**246**			**163**			**12.6**	**[6.3; 19.0]**	**10.5%**
Heterogeneity: $I^2 = 11.3\%$, $\tau^2 = 12.01$, $p = 0.34$											
Competitive vs. moral											
Sme03b vs. Sme03a	Priming	Scrambled sentence task with might−related words vs. moral−related words	25	32.2%	(17.0)	26	68.2%	(30.2)	36.0	[22.6; 49.4]	1.6%
Random effect			**25**			**26**			**36.0**	**[22.6; 49.4]**	**1.6%**
Heterogeneity: Not applicable											

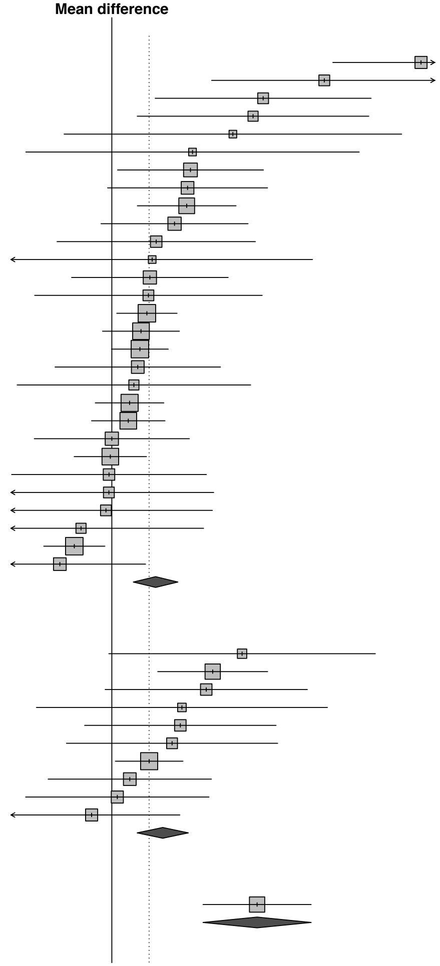


Figure 2.4. *In the first round, context framing effects were observed for all contrasts*. Forest plots: Larger squares indicate smaller standard errors and therefore greater weights on the random effects models; horizontal bars indicate the 95% confidence intervals. *n:* number of observations per framing condition (see Appendix 2.2 for details on the coding and see Appendix 2.3 for more details on the integrated studies).

Generic vs. cooperative											
Deu60a	Embedding	"Consider yourself to be partners" who are "interested" in each others' "welfare"	26	35.9%	(48.5)	23	89.1%	(34.4)	53.2	[29.9; 76.6]	1.0%
Deu60c	Embedding	"Consider yourself to be partners" who are "interested" in each others' "welfare"	11	41.2%	(52.2)	17	76.5%	(43.7)	35.3	[−1.9; 72.5]	0.5%
Ell98	Priming	Statements stressing the benefits of cooperative businesses strategies	24	58.3%	(50.4)	24	91.7%	(28.2)	33.3	[10.2; 56.4]	1.0%
Tor11b	Embedding	Money could be put in an envelope labelled "free rider" or "community box"	25	44.0%	(30.6)	25	72.8%	(34.1)	28.8	[10.9; 46.7]	1.3%
Reg04a	Options	Options between "free ride on" or "cooperate with" the "community"	20	34.4%	(37.1)	20	55.1%	(41.5)	20.7	[−3.7; 45.1]	0.9%
Bos15	Options	Options between "to contribute" and "not to contribute"	47	2.0%	(14.6)	42	21.0%	(41.5)	19.0	[5.8; 32.2]	1.6%
Duf11c	Title	"Community experiment"	24	34.2%	(26.9)	24	51.9%	(39.0)	17.7	[−1.2; 36.6]	1.2%
LaB14a	Priming	Text emphasising common goals of EU−citizens	29	57.4%	(21.1)	27	73.9%	(19.2)	16.6	[6.0; 27.1]	1.8%
LaB14b	Priming	Text emphasising common goals of EU−citizens	63	50.0%	(18.6)	61	64.8%	(15.6)	14.9	[8.8; 20.9]	2.2%
Ber14d	Title	"Community game"	15	41.6%	(29.1)	24	56.0%	(35.4)	14.3	[−6.1; 34.8]	1.1%
Wu11a	Statement	Participants "had a good relationship and liked each other"	40	35.8%	(12.3)	40	47.0%	(27.4)	11.2	[1.9; 20.6]	1.9%
Dro15	Priming	Detection of cooperation−related words in a matrix, e.g., "teamwork" and "trust"	75	24.9%	(26.6)	75	35.9%	(33.9)	10.9	[1.2; 20.6]	1.9%
Lop12	Statement	"The earnings for the group are the highest when everybody contributes"	30	50.7%	(22.8)	30	60.7%	(30.1)	10.0	[−3.5; 23.5]	1.6%
Reg04b	Options	Options between "free ride on" or "cooperate with" the "community"	20	68.2%	(42.3)	20	77.3%	(38.4)	9.1	[−15.9; 34.1]	0.9%
Eri14b	Title	"Teamwork game"	84	49.9%	(29.5)	84	58.6%	(31.0)	8.7	[−0.5; 17.9]	1.9%
Eri14a	Title	"Teamwork game"	193	51.5%	(25.4)	47	56.9%	(26.4)	5.4	[−3.0; 13.7]	2.0%
Bou13	Statement	Simulation of "co−management of protected areas" like "the local wetland/forest"	16	40.6%	(27.4)	28	45.9%	(36.0)	5.3	[−13.6; 24.2]	1.2%
Ber14a	Title	"Community game"	15	52.4%	(29.9)	12	55.9%	(37.3)	3.5	[−22.4; 29.5]	0.8%
Eng14b	Embedding	Participants form a "team" to "successfully coordinate"	20	74.3%	(44.4)	50	76.8%	(43.1)	2.5	[−20.3; 25.4]	1.0%
Eng14a	Embedding	Participants form a "team" to "protect itself against damage"	20	74.3%	(44.4)	52	72.8%	(44.8)	−1.4	[−24.4; 21.5]	1.0%
Eng14d	Embedding	Participants form a "team" to "protect itself against damage"	20	53.5%	(51.0)	51	48.0%	(50.5)	−5.5	[−31.8; 20.8]	0.8%
Duf11a	Title	"Community experiment"	66	28.6%	(32.8)	51	22.6%	(31.9)	−6.0	[−17.8; 5.8]	1.7%
Tor11a	Embedding	Money could be put in an envelope labelled "free rider" or "community box"	50	62.8%	(34.0)	25	53.6%	(31.5)	−9.2	[−24.7; 6.3]	1.4%
Bra09b	Title	"Community Game"	16	42.5%	(13.7)	20	32.9%	(16.6)	−9.6	[−19.5; 0.3]	1.9%
Duf11b	Title	"Community experiment"	72	23.8%	(34.5)	66	13.9%	(22.2)	−10.0	[−19.6; −0.4]	1.9%
Bra09f	Title	"Community game"	16	28.8%	(23.8)	12	14.5%	(14.9)	−14.2	[−28.6; 0.1]	1.5%
Eng14e	Embedding	Participants form a "team" to "successfully coordinate"	20	53.5%	(51.0)	50	38.6%	(49.0)	−14.9	[−41.1; 11.2]	0.8%
Bra09h	Title	"Community Game"	6	55.1%	(36.0)	10	33.8%	(29.9)	−21.3	[−55.5; 13.0]	0.6%
Random effect			**1063**			**1010**			**7.5**	**[2.4; 12.6]**	**37.4%**
Heterogeneity: $I^2 = 70.8\%$, $\tau^2 = 114.3$, $p < 0.01$											
Generic vs. moral											
Ahm11a	Priming	Scrambled sentence task with religion−related words, e.g., "spiritual"	112	26.8%	(44.5)	112	43.8%	(49.8)	17.0	[4.6; 29.3]	1.7%
Ahm11b	Priming	Scrambled Sentence with religion−related words, e.g., "god", "belief", "salvation"	48	59.6%	(33.7)	48	73.5%	(23.6)	13.9	[2.3; 25.5]	1.7%
Sme03a	Priming	Scrambled sentence task with moral−related words, e.g., "honest", "fair"	17	57.2%	(25.0)	26	68.2%	(30.2)	11.0	[−5.6; 27.6]	1.4%
Ran14a	Priming	Participants read Bible passage on the importance of charity	44	28.9%	(46.2)	106	37.0%	(48.5)	8.1	[−8.4; 24.6]	1.4%
Bra09a	Statement	"For each token that you invest ... the earnings of the other person are reduced"	16	39.0%	(13.1)	44	44.1%	(18.3)	5.1	[−3.3; 13.5]	2.0%
Ran14c	Priming	Participants read text on a family's dedication to charitable giving	44	28.9%	(46.2)	112	33.5%	(47.6)	4.6	[−11.6; 20.9]	1.4%
Ben13a	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	266	55.9%	(41.6)	267	60.4%	(40.7)	4.4	[−2.6; 11.4]	2.1%
Hor11	Priming	Participants read Bible passage on the importance of charity	70	63.1%	(48.7)	99	67.3%	(47.0)	4.2	[−10.5; 18.9]	1.5%
Bra09e	Statement	"For each token that you invest ... the earnings of the other person are reduced"	22	39.6%	(37.8)	14	37.4%	(38.5)	−2.2	[−27.8; 23.4]	0.9%
Ran14b	Priming	Participants read S'rimad Bhagavatam passage on the importance of charity	44	28.9%	(46.2)	115	26.3%	(44.1)	−2.6	[−18.4; 13.3]	1.4%
Bra09d	Statement	"For each token that you invest ... the earnings of the other person are reduced"	17	39.0%	(35.0)	16	28.2%	(26.1)	−10.8	[−31.7; 10.2]	1.1%
Random effect			**700**			**959**			**6.2**	**[2.5; 9.9]**	**16.5%**
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.49$											
Combined random effect			**3726**			**3893**			**9.2**	**[6.3; 12.2]**	**100.0%**
Heterogeneity: $I^2 = 65.7\%$, $\tau^2 = 95.46$, $p < 0.01$											

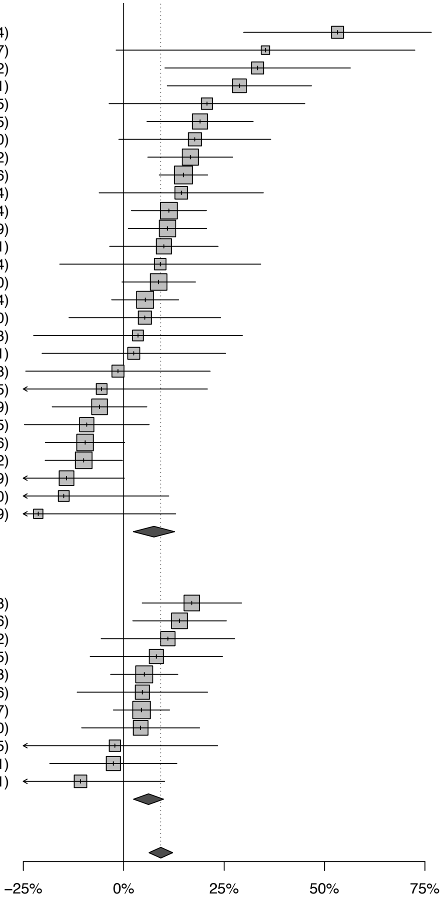


Figure 2.4. (continued)

Was the context framing effect due to a publication bias? Figure 2.5 plots the effect sizes observed for the first round against their standard errors after trim-and-fill adjustment. The funnel plots indicates that two contrasts (competitive vs. generic; competitive vs. cooperative) yielded systematically larger effects in studies with larger standard errors suggesting that context framing effects are overestimated. The opposite trend was observed for the generic vs. moral contrast. In other words, there are indicators of publication bias, but they point in opposite directions. Especially the unadjusted effect size of competitive frames vs. cooperative frames should be interpreted carefully: to adjust for publication bias, the trim-and-fill analysis added seven "unpublished" experiments whose integration lowered the framing effect from an unadjusted D = 11% to an adjusted D = 5% (Table 2.3). Trim-and-fill analysis may, however, have inappropriately adjusted for publication bias where none existed. Testing for funnel plot asymmetry yielded insignificant evidence to conclude publication bias, $t(8) = 0.48$, $p = .645$. The more conservative Copas selection model did not approximate any "unpublished" experiments for the contrasts. Instead, it estimated an adjusted the effect size from D = 11% to D = 8% (Table 2.3). Evidence for a publication bias in competitive vs. cooperative contrasts is thus inconclusive. Overall, the contrast of competitive vs. cooperative frames was the only one whose adjustment lowered the estimated framing effect by more than 1%. Table 2.3 provides an overview of all effect sizes, adjusted and unadjusted. Overall, adjustment lowered the combined random effect size across all contrasts from D = 9% to either D = 6% (trim-and-fill) or D = 8% (Copas). Neither adjustment suggested that the combined context framing effect was subject to random sampling error. In other words, context framing effects seem to be largely reliable phenomena, although the effect sizes are likely lower than the available literature suggests.

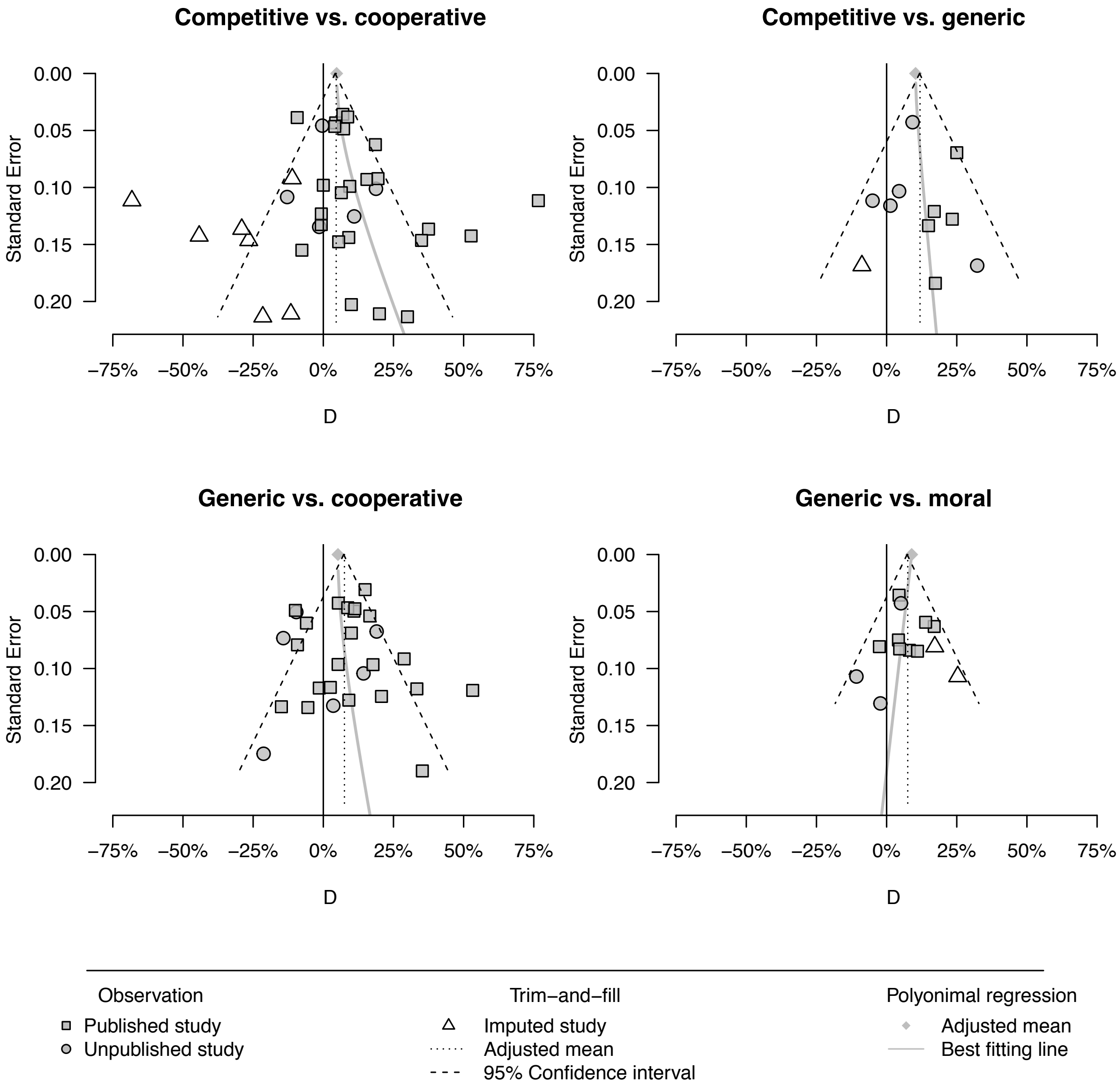


Figure 2.5. *Evidence for publication bias.* Funnel plots for the first round of social dilemma games after trim-and-fill adjustment. The dotted vertical lines indicate the means of the random effects models after the trim-and-fill correction. The dashed diagonal lines are the associated 95% confidence intervals. The solid gray lines represent the maximum likelihood estimates of observed experiments with local polynomial regression fitting. The combination of competitive vs. moral frames provided too few data points to meaningfully adjust the effect size.

Table 2.3.

Context Framing Effects in the First Round: Before and After Adjusting for Publication Bias

Contrast	*k*	*D (95% CI)*	*z*	*p*
Competitive vs. cooperative				
Unadjusted	29	10.8% (5.3%, 16.4%)	3.85	< .001
Trim-and-fill	36	4.6% (–1.6%, 10.8%)	1.45	.148
Copas selection	29	8.0% (3.0%, 13.0%)	—	.002
Competitive vs. generic				
Unadjusted	10	12.6% (6.3%, 19.0%)	3.90	< .001
Trim-and-fill	11	11.9% (5.4%, 18.4%)	3.61	< .001
Copas selection	10	12.6% (6.2%, 18.9%)	—	< .001
Competitive vs. moral				
Unadjusted	1	36.0% (22.6%, 49.4%)	5.26	< .001
Trim-and-fill	—	—	—	—
Copas selection	—	—	—	—
Generic vs. cooperative				
Unadjusted	28	7.5% (2.4%, 12.6%)	2.88	.004
Trim-and-fill	28	7.5% (2.4%, 12.6%)	2.88	.004
Copas selection	28	7.6% (2.4%, 12.6%)	—	.006
Generic vs. moral				
Unadjusted	11	6.2% (2.5%, 9.9%)	3.25	.001
Trim-and-fill	13	7.5% (3.5%, 11.6%)	3.63	< .001
Copas selection	11	6.2% (2.5%, 9.9%)	—	.001
Combined random effect				
Unadjusted	78	9.2% (6.1%, 12.1%)	6.13	< .001
Trim-and-fill	88	5.9% (2.5%, 9.2%)	3.45	< .001
Copas selection	82	8.0% (5.1%, 10.9%)	—	< .001

Note. k: number of studies per random effects model. The combination of competitive vs. moral frames provided too few data points to meaningfully adjust the effect size.

Persistence of context framing effects. Eleven of the social dilemma games analyzed were repeated games, allowing us to assess the persistence of context framing effects in later rounds. None of these games involved moral frames. For the remaining contrasts, the combined random effects across all rounds of repeated games indicated an increase in the cooperation rate of $D = 19\%$, from $C = 32\%$ to $C = 51\%$ (Table 2.4.; Figure 2.6). Even when we limited the analysis to the last round of repeated games, we still observed framing-induced cooperation of $D = 22\%$, from $C = 23\%$ to $C = 45\%$ (*95% CI of D* [8%, 35%], $z = 3.21$, $p = .001$). These results suggest that context framing can have a long-lasting influence on cooperation.

Table 2.4.

Average Rates of Cooperation in All Rounds of Repeated Social Dilemmas per Contrast

Contrast (Frame 1 vs. frame 2)	*k*	Cooperation rate: Frame 1	Cooperation rate: Frame 2	*D*	*z*	*p*
Competitive vs. cooperative	4	27.0%	54.6%	27.1%	3.07	.002
Competitive vs. generic	2	25.9%	32.8%	6.6%	2.02	.004
Generic vs. cooperative	5	38.8%	57.0 %	18.2%	3.08	.002
Combined random effect	11	32.3%	51.4%	19.1%	4.94	< .001

Note. For experiments with a larger standard error, smaller weights were applied in the calculation of the means.

Beliefs. We next examined the changes predicted by the three classes of theories. First, we assessed framing effects on beliefs. Overall, 25 experiments asked participants about their beliefs in the first round, before they learned about the actual decision(s) of their partner(s). The combined random effects indicated that framing increased the belief in cooperation by $D = 15\%$, from $C = 42\%$ to $C = 57\%$, but none of the single contrasts had a significant effect (Figure 2.7; Table 2.5). Nevertheless, framing effects on beliefs were strongly and positively correlated with framing effects on the decisions to cooperate ($\rho = 0.90$, $p < .001$; Spearman's rank correlation based on the means per experiment). In sum, the results suggest that, overall, context frames affect beliefs, although this conclusion is invalid for each single contrast alone.

Frame 1 vs. frame 2	Frame type	Framed content	n	Frame 1 C	Frame 1 (SD)	n	Frame 2 C	Frame 2 (SD)	D in %	95% CI	Weight
Competitive vs. cooperative											
Deu60d vs. Deu60c	Embedding	"Win as much money as you can" vs. "consider yourself to be partners"	170	25.48%	(13.90)	170	70.88%	(35.50)	45.4	[39.7; 51.1]	8.3%
Lib04a	Title	"Wall Street game" vs. "community game"	168	31.55%	(55.20)	168	66.07%	(37.77)	34.5	[24.4; 44.6]	7.5%
Lib04b	Title	"Market game" vs. "community game"	100	25.00%	(44.60)	100	45.00%	(61.60)	20.0	[5.1; 34.9]	6.5%
Hri13b vs. Hri13a	Statement	Participants "play a sequence of games based on team−work" vs. "based on competition"	9	27.30%	(36.89)	9	41.40%	(49.90)	14.1	[−26.4; 54.6]	2.5%
Wu11b vs. Wu11a	Statement	Participants "had a good relationship" vs. were the "winner" of the round	300	25.80%	(23.30)	300	39.70%	(29.75)	13.9	[9.6; 18.2]	8.4%
Random effect			**747**			**747**			**27.1**	**[9.8; 44.4]**	**33.2%**
Heterogeneity: $I^2 = 94.9\%$, $\tau^2 = 324.6$, $p < 0.01$											
Competitive vs. generic											
Deu60d	Embedding	"Win as much money as you can" and "do better than the other"	170	25.48%	(13.90)	170	35.59%	(16.00)	10.1	[6.9; 13.3]	8.5%
Hri13b	Statement	"You will play a sequence of games based on competition"	9	27.30%	(36.89)	9	36.40%	(42.00)	9.1	[−27.4; 45.6]	2.9%
Wu11b	Statement	After each round, participants were informed "you are (not) the winner"	300	25.80%	(23.30)	300	28.80%	(17.40)	3.0	[−0.3; 6.3]	8.5%
Random effect			**479**			**479**			**6.6**	**[0.2; 13.1]**	**19.9%**
Heterogeneity: $I^2 = 78.4\%$, $\tau^2 = 19.6$, $p < 0.01$											
Generic vs. cooperative											
Ell98	Priming	Statements stressing the benefits of cooperative businesses strategies	144	38.89%	(28.52)	144	75.00%	(32.60)	36.1	[29.0; 43.2]	8.1%
Deu60c	Embedding	"Consider yourself to be partners" who are "interested" in each others' "welfare"	170	35.59%	(16.00)	170	70.88%	(35.50)	35.3	[29.4; 41.1]	8.2%
Gri12	Statement	"Each of you will receive the amount of money you have earned together"	20	31.00%	(46.48)	20	56.00%	(49.76)	25.0	[−4.8; 54.8]	3.7%
Wu11a	Statement	Participants "had a good relationship and liked each other"	300	28.80%	(17.40)	300	39.70%	(29.75)	10.9	[7.0; 14.8]	8.5%
Lop12	Statement	"The earnings for the group are the highest when everybody contributes"	450	58.40%	(25.20)	450	64.80%	(32.00)	6.4	[2.6; 10.2]	8.5%
Hri13a	Statement	"You will play a sequence of games based on team−work"	9	36.40%	(42.00)	9	41.40%	(49.90)	5.0	[−37.6; 47.6]	2.3%
Bou13	Statement	Simulation of "co−management of protected areas" like "the local wetland/forest"	80	35.81%	(32.59)	140	38.54%	(36.81)	2.7	[−6.7; 12.1]	7.6%
Random effect			**1173**			**1233**			**18.2**	**[6.6; 29.7]**	**46.9%**
Heterogeneity: $I^2 = 94.6\%$, $\tau^2 = 190.5$, $p < 0.01$											
Combined random effect			**2399**			**2459**			**19.1**	**[11.5; 26.6]**	**100.0%**
Heterogeneity: $I^2 = 95.2\%$, $\tau^2 = 171.8$, $p < 0.01$											

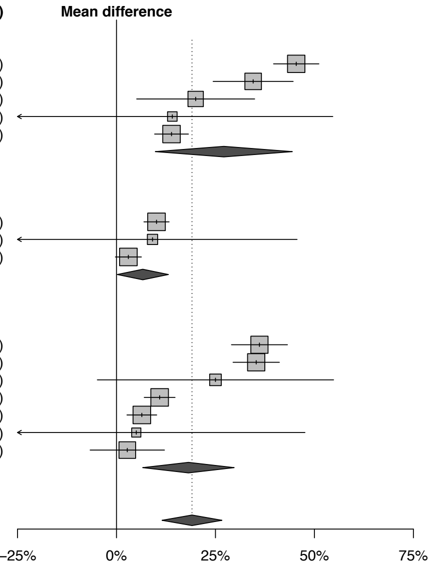


Figure 2.6. *Context framing effects persisted across all rounds of repeated social dilemmas.*

Frame 1 vs. frame 2	Frame type	Framed content	Frame 1 C	Frame 1 SD	n	Frame 2 C	Frame 2 SD	n	D in %	95% CI	Weight
Competitive vs. cooperative											
Deu60b vs. Deu60a	Embedding	"Win as much money as you can" vs. "consider yourself to be partners"	15.6%	(36.9)	32	91.3%	(28.5)	46	75.7	[60.5; 90.9]	4.3%
Deu60d vs. Deu60c	Embedding	"Win as much money as you can" vs. "consider yourself to be partners"	26.2%	(44.5)	42	79.4%	(41.0)	34	53.2	[33.9; 72.5]	4.0%
Lib04b	Title	"Market game" vs. "community game"	24.1%	(44.4)	20	45.8%	(51.0)	20	21.7	[−8.0; 51.4]	3.2%
Pil99	Embedding	"Investing" in an "investment corporation" vs. "donating" to "a social event"	28.6%	(47.0)	23	41.9%	(50.7)	23	13.2	[−15.0; 41.5]	3.3%
Ber14e vs. Ber14d	Title	"Wall Street game" vs. "Community game"	55.9%	(50.1)	48	55.1%	(50.4)	48	−0.8	[−20.9; 19.3]	4.0%
Ber14b vs. Ber14a	Title	"Wall Street game" vs. "Community game"	58.3%	(49.8)	48	52.8%	(50.5)	48	−5.5	[−25.6; 14.5]	4.0%
Random effect					**213**			**219**	**26.8**	**[−3.4; 56.9]**	**22.7%**
Heterogeneity: I^2 = 91.9%, τ^2 = 1286 , $p < 0.01$											
Competitive vs. generic											
Deu60b	Embedding	"Win as much money as you can" and "do better than the other"	15.6%	(36.9)	32	59.0%	(49.8)	39	43.3	[23.1; 63.5]	3.9%
Sme03b	Priming	Scrambled sentence task with might−related words, e.g., "assertive", "dominant"	37.8%	(49.0)	50	56.1%	(50.2)	52	18.4	[−0.9; 37.6]	4.0%
Deu60d	Embedding	"Win as much money as you can" and "do better than the other"	26.2%	(44.5)	42	44.1%	(50.4)	34	17.9	[−3.7; 39.6]	3.8%
Ber14e	Title	"Wall Street game"	55.9%	(50.1)	48	53.2%	(50.5)	44	−2.8	[−23.4; 17.8]	3.9%
Ber14b	Title	"Wall Street game"	58.3%	(49.8)	48	54.5%	(50.4)	44	−3.8	[−24.3; 16.7]	3.9%
Random effect					**220**			**213**	**14.7**	**[−2.3; 31.7]**	**19.6%**
Heterogeneity: I^2 = 71.1%, τ^2 = 266.4, $p < 0.01$											
Competitive vs. moral											
Sme03b vs. Sme03a	Priming	Scrambled sentence task with might−related words vs. moral−related words	37.8%	(49.0)	50	71.6%	(45.5)	53	33.8	[15.5; 52.1]	4.1%
Random effect					**50**			**53**	**33.8**	**[15.5; 52.1]**	**4.1%**
Heterogeneity: Not applicable											
Generic vs. cooperative											
Deu60c	Embedding	"Consider yourself to be partners" who are "interested" in each others' "welfare"	44.1%	(50.4)	34	79.4%	(41.0)	34	35.3	[13.5; 57.1]	3.8%
Deu60a	Embedding	"Consider yourself to be partners" who are "interested" in each others' "welfare"	59.0%	(49.8)	39	91.3%	(28.5)	46	32.3	[14.7; 50.0]	4.1%
LaB14b	Priming	Text emphasising common goals of EU−citizens	54.8%	(50.1)	63	67.6%	(47.3)	61	12.8	[−4.4; 29.9]	4.2%
Duf11c	Title	"Community experiment"	44.7%	(50.9)	24	53.1%	(50.9)	24	8.4	[−20.4; 37.2]	3.2%
Dro15	Priming	Detection of cooperation−related words in a matrix, e.g., "teamwork" and "trust"	33.5%	(47.5)	75	39.1%	(49.0)	75	5.6	[−9.8; 21.0]	4.3%
Eri14a	Title	"Teamwork game"	41.0%	(49.3)	193	44.7%	(50.3)	47	3.7	[−12.2; 19.7]	4.3%
Eri14b	Title	"Teamwork game"	42.0%	(49.6)	84	44.8%	(50.1)	84	2.7	[−12.3; 17.8]	4.4%
Ber14d	Title	"Community game"	53.2%	(50.5)	44	55.1%	(50.4)	48	1.9	[−18.7; 22.6]	3.9%
Ber14a	Title	"Community game"	54.5%	(50.4)	44	52.8%	(50.5)	48	−1.7	[−22.4; 18.9]	3.9%
Duf11b	Title	"Community experiment"	26.8%	(44.4)	72	20.8%	(41.2)	66	−6.0	[−20.2; 8.3]	4.4%
Duf11a	Title	"Community experiment"	40.2%	(49.5)	66	34.1%	(47.6)	51	−6.2	[−23.9; 11.5]	4.1%
Random effect					**738**			**584**	**7.4**	**[−0.5; 15.4]**	**44.7%**
Heterogeneity: I^2 = 53.4%, τ^2 = 94.4 , $p = 0.02$											
Generic vs. moral											
Sme03a	Priming	Scrambled sentence task with moral−related words, e.g., "honest", "fair"	56.1%	(50.2)	52	71.6%	(45.5)	53	15.4	[−2.9; 33.7]	4.1%
Ben13a	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	60.8%	(48.9)	266	62.4%	(48.5)	267	1.6	[−6.6; 9.9]	4.8%
Random effect					**318**			**320**	**6.0**	**[−6.6; 18.6]**	**8.9%**
Heterogeneity: I^2 = 44.7%, τ^2 = 42.5 , $p = 0.18$											
Combined random effect					**1539**			**1389**	**14.8**	**[6.1; 23.5]**	**100.0%**
Heterogeneity: I^2 = 82.9%, τ^2 = 394.5, $p < 0.01$											

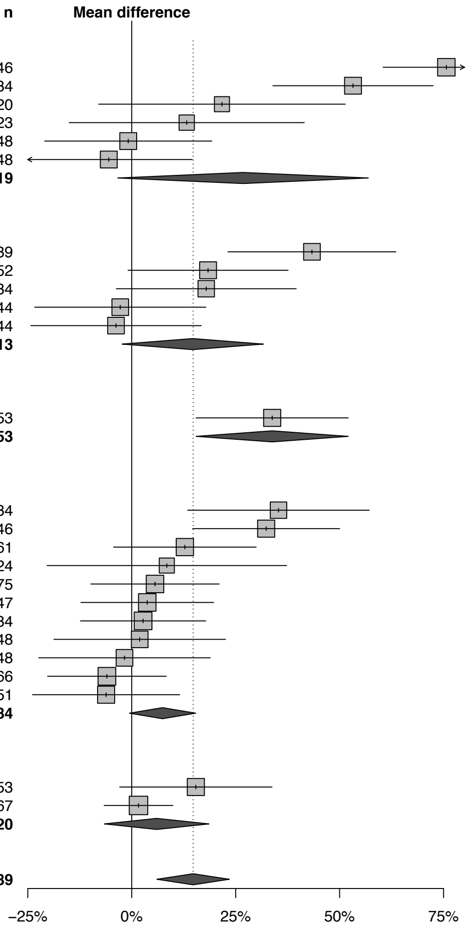


Figure 2.7. *Context framing effects on beliefs in cooperation.*

Table 2.5.

Average Rates of Beliefs in Partner's Cooperation per Contrast

Contrast (Frame 1 vs. frame 2)	*k*	Cooperation rate Frame 1	 Frame 2	*D*	*z*	*p*
Competitive vs. cooperative	6	35.0%	62.7%	26.8%	1.74	.082
Competitive vs. generic	5	38.8%	53.5%	14.7%	1.69	.090
Competitive vs. moral	1	37.8%	71.6%	33.8%	1.69	.090
Generic vs. cooperative	11	44.6%	52.5%	7.4%	1.83	.068
Generic vs. moral	2	58.6%	66.6%	6.0%	0.94	.350
Combined random effect	11	42.3%	57.0%	14.8%	3.32	< .001

Note. For experiments with a larger standard error, smaller weights were applied in the calculation of the means.

Dictator games. Were context framing effects also observable in dictator games? Overall, only 15 experiments compared context frames in dictator games. The combined random effect suggested an average increase in donations of $D = 10\%$, from $C = 28\%$ to $C = 38\%$ (Figure 2.8; Table 2.6). However, only one contrast (generic vs. moral) provided sufficient evidence to conclude a context framing effects.

Table 2.6.

Average Rates of Donations in Dictator Games per Contrast

Contrast (Frame 1 vs. frame 2)	*k*	Donation rate Frame 1	 Frame 2	*D*	*z*	*p*
Generic vs. cooperative	3	27.5%	33.6%	6.6%	1.90	.058
Generic vs. moral	13	28.0%	38.3%	10.6%	4.09	< .001
Combined random effect	15	27.9%	37.7%	9.8%	4.14	< .001

Note. For experiments with a larger standard error, smaller weights were applied in the calculation of the means.

Frame 1 vs. frame 2	Frame type	Framed content	n	Frame 1 C	Frame 1 (SD)	n	Frame 2 C	Frame 2 (SD)	D in %	95% CI	Weight
Generic vs. cooperative											
Ban14b	Statement/Options	Participants "constitute a two−member group" and recipients are the "group partner"	31	27.5%	(20.1)	64	35.4%	(19.1)	8.0	[−0.5; 16.4]	7.6%
Ban14a	Statement	"You and your paired partner constitute a two−member group"	31	27.5%	(20.1)	38	31.5%	(27.6)	4.1	[−7.2; 15.3]	6.4%
Random effect			**62**			**102**			**6.6**	**[−0.2; 13.3]**	**14.0%**
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.59$											
Generic vs. moral											
DeS13b	Embedding	"Customer" "can decide to pay" a "merchant" who "delivered all of the goods"	14	36.3%	(22.0)	36	69.1%	(41.0)	32.9	[15.2; 50.6]	4.1%
Sha07a	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	12	18.4%	(18.0)	25	42.2%	(26.5)	23.8	[9.3; 38.3]	5.1%
Sha07b	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	12	25.6%	(26.9)	25	45.6%	(30.3)	20.0	[0.7; 39.3]	3.7%
Bra07a	Statement	"Note that your recipient relies on you"	20	12.0%	(15.8)	20	31.0%	(14.1)	19.0	[9.7; 28.3]	7.2%
Sha07c	Priming	Scrambled sentence task with secular moral words, e.g., "civic", "jury" and "court"	25	25.6%	(26.9)	25	44.4%	(30.0)	18.8	[3.0; 34.6]	4.7%
DeS13a	Embedding	"Brother" and "sister" can decide to give" food to their "hungry" sibling	14	36.3%	(22.0)	40	55.0%	(19.1)	18.7	[5.7; 31.7]	5.7%
Bra10	Statement	"Note that your recipient relies on you"	26	23.5%	(16.0)	26	33.9%	(13.6)	10.4	[2.3; 18.5]	7.8%
Bra07b	Statement	"Note that your recipient relies on you"	27	24.8%	(16.3)	26	33.9%	(13.6)	9.0	[1.0; 17.1]	7.8%
DeS13c	Embedding	"Brother" and "sister" can decide to give" food to their "not too hungry" sibling	40	28.0%	(18.6)	14	36.3%	(22.0)	8.3	[−4.6; 21.2]	5.7%
Ahm11c	Priming	Scrambled sentence task with religion−related words, e.g., "spiritual"	112	27.4%	(17.5)	112	34.5%	(14.9)	7.1	[2.9; 11.4]	9.4%
Gom15b	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	108	44.9%	(34.9)	195	47.0%	(37.2)	2.1	[−6.3; 10.5]	7.6%
Gom15a	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	108	44.9%	(34.9)	239	42.8%	(36.7)	−2.1	[−10.2; 6.0]	7.8%
Ben13b	Priming	Scrambled sentence task with religion−related words, e.g., "spirit", "divine"	278	17.9%	(27.3)	276	14.6%	(23.2)	−3.3	[−7.5; 0.9]	9.4%
Random effect			**796**			**1059**			**10.6**	**[5.3; 16.0]**	**86.0%**
Heterogeneity: $I^2 = 78.5\%$, $\tau^2 = 65.4$, $p < 0.01$											
Combined random effect			**858**			**1161**			**9.8**	**[5.2; 14.5]**	**100.0%**
Heterogeneity: $I^2 = 75.1\%$, $\tau^2 = 55.3$, $p < 0.01$											

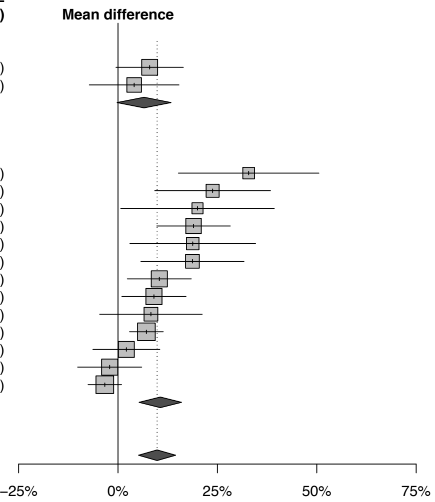


Figure 2.8. *Context framing increases donations in dictator games. n*: participants in the role of dictators only.

Moderator Analyses

What explains the great heterogeneity observed in context framing effects? And under which conditions are the effects more pronounced? To examine these questions, we ran a meta-regression model that predicted framing effects on cooperation in the first round of social dilemmas (Table 2.7).

Table 2.7.

Covariates of Context Framing Effects in the First Round of Social Dilemma Games

	b (95% CI)	*SE*	*z*	*p*
Intercept	25.5% (12.0%, 39.1%)	6.93	3.69	< .001
Cooperative frame	−2.2% (−11.3%, 6.9%)	4.64	−0.47	.637
Competitive frame	5.2% (−0.9%, 11.2%)	3.09	1.67	.094
Moral frame	7.9% (−18.8%, 3.1%)	5.58	−1.41	.159
Priming technique	9.2% (1.6%, 16.8%)	3.89	2.37	.018
Online setting	−7.2% (−1.4%, −13.0%)	2.98	2.42	.016
Public goods games	−8.9% (−15.6%, −2.1%)	3.44	−2.58	.010
Number of interaction partners	−0.1% (−1.8%, 1.7%)	0.88	−0.07	.941
Number of rounds	0.6% (−0.3%, 1.4%)	0.45	1.25	.210
Incentivization	−19.9% (−28.9%, −10.8%)	4.63	−4.29	< .001
Observations	$k = 79$, $n = 7{,}619$			
Residual heterogeneity	$I^2 = 46.6\%$, $\tau^2 = 48.42$			
Heterogeneity accounted for	$R^2 = 49.3\%$			

Cooperative frame, competitive frame, moral frame. To test whether one framed concept yielded bigger effects than another, we built dummy variables for reference to cooperation, competition, and morality. None of the dummies predicted the context framing effect, suggesting that the framed concepts were associated induced equivalent cooperation rates.

Priming technique. Context framings were implemented in several ways. To test whether the magnitude of the context framing effect depended on the "overtness" of the framing, we coded whether or not the framing was realized via priming. Priming techniques provide a more subtle association between the framed concept and the social dilemma than does a direct change in the description of the social dilemma's options, title, etc. On average, the context effects of priming experiments were $b = 9$ percentage points larger than those of experiments that changed the description of the experiment directly.

Online setting. Some experiments were conducted online, whereas others took place in laboratories. We tested whether the change in social proximity—among participants as well as between participants and experimenter—was associated with different magnitudes of framing effects. On average, the context effects of online experiments were $b = 7$ percentage points smaller than those of experiments in physical laboratories.

Public goods games. All social dilemma games examined were either prisoner's dilemmas or public goods games. The two differ on at least one key dimension: In prisoner's dilemmas, choices are dichotomous (cooperate vs. defect); in public goods games, choices entail gradual transfers to the common pool. We found that game type moderated the context framing effect: The average framing effect in public good games was $b = 9$ percentage points smaller than in prisoner's dilemmas.

Number of interaction partners. In prisoner's dilemmas, two participants typically interact; in public goods games, the number of participants can vary. In the experiments analyzed, all but two prisoner's dilemmas were played with two participants, whereas the number of participants in public goods games ranged from two to ten ($M = 4.13$, $SD = 2.27$). However, the results of the regression model suggest that the number of interaction partners did not predict the magnitude of the context framing effect.

Number of rounds. Anticipation of future interaction(s) with the same partner(s) may already have moderated the context framing effect in the first round. To test this possibility, we included the number of rounds as a predictor in the regression model. However, the number of rounds did not moderate the size of the context framing effect in the first round.

Incentivization. Only some experiments paid participants according to the outcome of the game. Others paid out a flat show-up fee that was independent of the interaction. Incentivization markedly lowered framing effects. On average, the effect sizes in incentivized experiments were $b = 20$ percentage points smaller than those in nonincentivized experiments. This result suggests that findings of context framing effects in social dilemmas were largely driven by studies that paid a flat fee. Does dropping these nonincentivized experiments eliminate the context framing effect? To test this hypothesis, we removed all experiments paying flat show-up fees from the analysis of cooperation in the first round. Even when we limited the analysis to incentivized social dilemmas, the combined random effects model estimated an increase in the cooperation rate of $D = 6\%$, from $C = 44\%$ to $C = 50\%$ (*95% CI* [3%, 9%], $z = 4.61$, $p < .001$). Overall, this finding suggests that incentivization reduces, but does not eliminate, context framing effects.

Summary

Using a variety of statistical models, we assessed the magnitude, robustness, and moderators of context framing effects in social dilemmas involving decisions about cooperation. We found framing effects on cooperation in the first round for all combinations of context frames: On average, there was an increase in cooperation of about $D = 9\%$, from $C = 42\%$ to $C = 51\%$. Such an effect corresponds to the magnitude of lifting the participants' mutual anonymity in a social dilemma game (Bohnet & Frey, 1999). Robustness analyses suggested that the evidence in favor of framing effects was partly due to publication bias. "Correcting" for publication bias decreased, but did not eliminate, the combined average framing effect. Framing effects were also present in all rounds of repeated social dilemmas, and they were even observed when the analysis was limited to the last round of interaction, suggesting that framing effects were relatively persistent. In social dilemmas, participants' beliefs were also subject to context framing, as were their donations in dictator games.

We tested several potential moderators of framing effects. The framing effect observed in the first round was not moderated by a specific framed concept (i.e., cooperation, competition, morality), the number of interaction partners, or the prospect of repetition. However, the context framing effects observed were larger in experiments that used priming techniques than in those that changed the description directly; in experiments conducted in laboratories than in those conducted online; and in prisoner's dilemmas than in public goods games. The strongest moderation effect was due to incentivizing: context framing effects in social dilemmas were much weaker when participants were paid according to their decisions than when flat fees were paid. Nonetheless, excluding the nonincentivized experiments from the analysis did not eliminate the context framing effects.

Discussion

In what follows, we discuss the limitations and theoretical implications of our results and we map out the next steps for future research.

Limitations

There are several limitations to our analyses. Most importantly, there was a substantial amount of variation between the context framing effects that we could not explain. It is possible that this was due to the theory-based categorization of frames along the two dimensions of the grid-group analysis. The categorization may not have fully reflected individuals' conceptual associations with the social situation. These associations ultimately depend on the participants' own *interpretation* of the social situation—and not on some external criteria for categorizing the context frames.

Furthermore, all analyses were based on effect sizes that used the mean cooperation rate as the underlying dependent variable. On the one hand, this was a conservative measure that likely underestimated framing effects: It is conceivable that context frames evoke opposing effects at the subgroup level that, when aggregated, cancel each other out. For example, framing moral concepts via religious appeals may have opposite effects for people of different religions (e.g., Benjamin, Choi, & Fisher, 2013; Rand, Dreber, et al., 2014). On the other hand, using mean cooperation rates may have obfuscated potential boundary conditions for the context framing effects. In public goods games, for example, changes in the mean cooperation rate may be due to some individuals being more sensitive to context frames or all individuals being subject to some context framing. As we did not have access to individual-specific information, we could not discriminate between these two effects.

Moreover, the finding that framing effects persisted into the last round needs to be interpreted with care. First, there were relatively few repeated experiments to begin with. Second, context frames may have affected choices in only the first round of the repeated games, with choices in all other rounds—including the last round—being mere reactions to the context framing effects in the first round. Alternatively, context frames may have elicited completely different strategies (Kölle, Gächter, & Quercia, 2014). Given that our analysis was based on mean cooperation rates and mean transfers, we cannot shed light on the dynamics of framing effects.

The findings of the meta-analysis suggest that context framing can induce cooperation and beliefs in social dilemma games and donations in dictator games. Neither social preference theories nor group identity theories sufficiently accounted for all the framing effects. Rather, social norms theories provided the most plausible explanation for context framing effects. However it does not mean that social norms theories are the only valid explanation. The distinct mechanisms of the three classes of theories are not mutually exclusive. It is conceivable that the mechanisms suggested by all three classes of theories hold, but for different context frames and/or individuals. For example, a combination of the mechanisms proposed by social preference theories and group identity theories could also explain the observed context framing effects. More fine-grained analyses are needed to test the precise mechanisms underlying the observed effects.

Furthermore, social norms theories argue that context frames can invoke different beliefs (in the wider sense) and different perceptions of normative standards, which in turn affect the individual's choice to cooperate in social dilemmas. However, they remain vague about the underlying mechanisms. In particular, the reference group of frame-induced changes is theoretically underdefined: An individual's beliefs about what other people would do and/or perceptions of the normative standard could relate to (1) the interaction partner(s), (2) the experimenter, (3) people in general, or (4) some other arbitrary but relevant reference group. The interpretation that context framing affects the experimenter's beliefs and/or perceptions suggests that participants deduce the experimenter's expected behavior and/or value judgment from his framing. If so, context frames "leak" choice-relevant information (see Sher & McKenzie, 2006, 2008). Such frame-induced change in perceptions of normative standards would constitute experimenter demand effects (Bardsley, 2008; Ortmann, 2005; Zizzo, 2010). Again, more fine-grained testing of the precise mechanisms is required.

Directions for Future Research

The existence of context framing effects has several implications for research practice and applications. It is common practice to use only a single frame per experiment, usually one that is perceived to be generic. This practice is reasonable given that generic frames are an established benchmark against which context frames can be tested (e.g., Camerer & Fehr, 2002). Context frames may invite associations with real-world situations, whereas the goal of many experiments is to create small-scale abstractions that eliminate hard-to-control sources of variance tied to natural settings (Hertwig & Ortmann, 2001; Ortmann, 2005). Many experimenters thus advise researchers to give instructions on social dilemmas in plain, abstract, and generic language (Binmore, 1999; Camerer & Fehr, 2002, p. 5; Friedman & Sunder, 1994, p. 17). Some even directly advise against including any real-world references (Camerer & Fehr, 2002; Friedman & Sunder, 1994, p. 17). Such advice suggests that only generic frames are "neutral" and permit "clean assessment" (Engel & Rand, 2014, p. 387). Decontextualization, as the hallmark of experimental control, seems justified if context framing induces potentially unwarranted information leakage from the experimenter (as suggested by social norms theories).

Our results support this view to the extent that context frames yield different cooperation rates than generic, and thus relatively decontextualized, frames. Our regression analysis suggested that incentivization and increasing social distance (by conducting experiments online) attenuated such framing effects. Both aspects have been recommended to reduce experimenter demand effects (Bardsley, 2008; Ortmann, 2005; Zizzo, 2010). Hence, if the goal is to reduce potentially unwanted context framing effects, we suggest that researchers incentivize experimental outcomes and conduct experiments online. The degree to which this is necessary may depend on the specific experimental paradigm. Some social dilemmas (prisoner's dilemmas) seem more prone to context framing effects than others (public goods games).

Conversely, generic frames do not necessarily render more generalizable inferences than any context frame does. Despite strong conventions in the use of language in social dilemma experiments, there are no set standards for generic frames. Consequently, generic frames include phrases from diverging domains and therefore potentially elicit different conceptual associations with the social dilemma. For example, participants may be addressed as "you and the other," options framed as "A" and "B," and the dilemma itself referred to as "the situation." Alternative frames may be taken from the domain of games (e.g., "the players," "the game"), finance (e.g., "the buyer," "the seller," "the exchange"), or experimental practice (e.g., "the participants," "the experiment").

One way to deal with potentially varying associations is, of course, to completely avoid words that link the game to concepts beyond the experiment itself. However, fully decontextualized frames may prompt individuals to develop their own interpretation of the experiment—akin to idiosyncratic interpretations of Rorschach ink blots. In such case participants may use cues from outside the experiment, even from totally unrelated sources, as in the case of priming. In other words, participants would go "beyond the information given" (Bruner, 1957) by associating more aspects with the situation than actually presented. This would undermine the goal of controlling for confounds via experiments (Alexander & Weil, 1969; Eiser & Bhavnani, 1974; Engel & Rand, 2014; Hagen & Hammerstein, 2006; Levitt & List, 2007). Postexperimental interviews have revealed large heterogeneity in the situational interpretation of "neutrally" framed games (e.g., Butler, Burbank, & Chisholm, 2011; Gerkey, 2013). Hagen and Hammerstein (2006), for example, argued that "most people do not, and cannot, see these games as abstract structures that can be logically analyzed ... The lack of explicit framing or contextualization merely allows participants to interpret the games in idiosyncratic ways that are often opaque to the experimenter" (p. 346).

Rather than using completely decontextualized frames, it may thus be preferable to use explicit frames and to systematically explore their effects on cooperation. Doing so might be particularly important, first, when providing context is unavoidable and, second, when the particular interpretation is relevant to participants' decisions. In some situations, context framing may be unavoidable—for instance, when testing atypical subject pools. For example, illiterate populations require more references to concepts beyond social dilemma games than do the populations frequently used in Western, educated, industrialized, rich, democratic (WEIRD) countries (see Henrich et al., 2001, 2005). Varying context frames can thus serve as a robustness check for the interpretation under which the tested theories hold—and under which interpretations particular effects vary or even disappear. Deliberate context framing seems preferable to the current practice in non-WEIRD populations, which is to use idiosyncratic frames that may vary across participant groups and even sessions (Ortmann, 2005). Context framing also makes it possible to test the degree to which choices depend on the interpretation and therefore on motivation, cognition, and other forms of situational appraisal. For example, it is possible that particular fairness norms are more likely to be invoked by specific context frames (Fiddick & Cummins, 2007; Hoffman, McCabe, Shachat, & Smith, 1994; Hoffman, McCabe, & Smith, 1996b; Hoffman & Spitzer, 1985). Whether a transaction is interpreted as a bribe or a gift may depend on the context frame (Lambsdorff & Frank, 2010). Context framing can induce such interpretations in the first place and thereby facilitate the systematic exploration of the boundary conditions of cooperation.

However, the insights provided by the present meta-analysis go beyond experimental practice. Social dilemmas represent a recurring problem in everyday life. Interventions that aim at increasing cooperation may benefit from harnessing the effects of context framing. Such interventions may be cheaper and easier to implement than conventional instruments aimed at boosting cooperation, such as sanctioning systems. As a first step, however, it will be necessary to test whether context framing can be applied to real-world cooperation problems in field experiments. We hope that the results of our analysis will serve as a catalyst for such investigations.

Conclusion

To conclude, context frames can have nontrivial effects on cooperation in social dilemmas. These effects are relatively robust and survive the experience of repeated play. Social norms theories provide a parsimonious explanation for this context framing effect. However, given the predominant focus on what are interpreted as generic and abstract frames in research practice, the potential effects of context framing are relatively little studied and understood.

Chapter 3

The Games Economists Play — Why Economics Students Behave More Selfishly than Other Students[3]

Economists seem to have never enjoyed a good reputation among their peers. In 1849, historian Thomas Carlyle described economics as "the dismal science" (Marglin, 2008, pp. 28–30). Thirty years later, economist Francis Walker felt compelled to explain why economists "tend to be in bad odor amongst real people" (de Waal, 2005, p. 243). And in 1945, psychoanalyst Donald Winnicott denounced economics as the "science of Greed" (Winnicott, 1945, p. 170). More recently, rather than working in the realm of speculation, researchers have sought to determine whether there is a sound empirical basis for economists' bad reputation. In particular, the results of experiments showing links between economic training and more selfish choices have lent firm support to the critics of the discipline: Economics students have been found to behave more selfishly than other populations across various situations involving monetary allocations (Charles Bram Cadsby & Maynes, 1998; Cappelen, Nygaard, Sørensen, & Tungodden, 2015; Carter & Irons, 1991; Childs, 2012a; Frank, Gilovich, & Regan, 1993, 1996; Grimm, Utikal, & Valmasoni, 2015; Haucap & Müller, 2014; López-Pérez & Spiegelman, 2009; Selten & Ockenfels, 1998; Van Lange, Schippers, & Balliet, 2011; Wang, Malhotra, & Murnighan, 2011).

Yet selfish behavior—like all behavior—is not free from context and may be driven by various motives. This chapter aims to shed a light on the potential links between studying economics and selfish behavior. More precisely, it investigates *why* economics students behave more selfishly than other people do. To this end, the chapter distinguishes three theoretical mechanisms that may account for more selfish behavior in economics students: economics students are less concerned with fairness[4] when making allocation decisions; they are equally concerned with fairness but have a different

[3] This chapter is based on: Gerlach, P. (2017). The games economists play. Why economics students behave more selfishly than other students. *PLOS ONE*. 12(9). e0183814.

[4] Throughout this chapter, fairness is defined in terms of what the people themselves perceive as fair.

notion of what is fair; they expect others to behave more selfishly and therefore feel less obliged to behave fairly themselves. These three mechanisms are empirically tested by comparing the decisions of students from various disciplines in a third-party punishment game. The results suggest that, relative to their fellow students, economics students are about equally likely to be concerned with fairness when making decisions; they have a similar notion of what is fair; but their greater skepticism about others' behavior mediates their more selfish behavior.

This chapter is organized as follows: The subsequent section outlines findings that suggest economics students behave more selfishly than others. Then, the three theoretical mechanisms potentially underlying this pattern of results are presented. Next, the experimental design is described, followed by the experimental results. Finally, the design, the results and the potential reasons for economics students' greater skepticism are discussed.

Experiments and Economics Students

A central finding of experiments involving monetary allocation decisions is that substantial numbers of people do not behave according to the predictions of game theory (Camerer, 2003; Kelley & Thibaut, 1978). Participants in such experimental games are frequently willing to contribute to the other participant's welfare in a non-trivial fashion—even at their own expense (Engel, 2011; Henrich et al., 2001, 2004; Kahneman et al., 1986). In public goods games, for instance, participants can choose how much of their private savings to contribute to a common pot, which is then multiplied and evenly distributed among all participants (Olson, 1965). The configuration of public goods games is such that participants are tempted to save everything privately and to contribute nothing to the common pot. This "free riding" is what game theory predicts for participants whose goal is to maximize payment. Yet, the less participants free ride, the less everybody earns—the "tragedy of the commons" (Hardin, 1968). Despite this bleak prediction findings from Western countries show that participants regularly contribute 40% to 60% of their initial stocks (Ledyard, 1995; Zelmer, 2003). There is, however, at least one systematic exception to this finding.

In a series of experiments, Marwell and Ames (1981) discovered that first semester economics graduate students contributed on average only 20% of their private savings to the common pot. In other words, the choices of the economics students were inclined towards free riding and thus consistent with the predictions of game theory. Despite some limitations in the experiments—for example, the samples were not strictly comparable—and a failed replication attempt (Isaac, McClue, & Plott, 1985), the work by Marwell and Ames stimulated a number of follow-up studies on whether economics students behave more selfishly than their peers. Overall, the investigations largely supported this claim: Economics students offered and accepted smaller amounts in ultimatum games (Carter & Irons, 1991); defected more in prisoner's dilemmas (Frank et al., 1993, 1996); deceived more in cheap talk games (Childs, 2012a; López-Pérez & Spiegelman, 2009); contributed less in threshold public goods games (Charles Bram Cadsby & Maynes, 1998); gave less in "solidarity games" (Selten & Ockenfels, 1998); shared less in dictator games (Cappelen et al., 2015; Grimm et al., 2015; Wang et al., 2011); trusted and reciprocated less in trust games (Haucap & Müller, 2014); and were less prosocial and more competitive in decomposed games (Van Lange et al., 2011).

Three theoretical mechanisms have been proposed to explain the more selfish behavior of economics students. The first argues that—in contrast to other students, who often indicate fairness as a motive driving their choices (Butler et al., 2011; Mellers, Haselhuhn, Tetlock, Silva, & Isen, 2010)—economics students are less concerned about fairness when making their decisions. The seminal investigation of Marwell and Ames (1981), for example, found that "the economics graduate students were about half as likely as other subjects to indicate that they were 'concerned with fairness' in making their investment decision" (Marwell & Ames, 1981, p. 308; Wang et al., 2011). This would suggest that economics students' behaviors are driven by motives other than fairness.

Marwell and Ames also speculated about an alternative theoretical account for their observation. Overall, they found "surprising unanimity of thought regarding what was considered fair" among the participants (Marwell & Ames, 1981, p. 308). Yet comparing economics students with other participants proved difficult because

> [m]ore than one-third of the economists either refused to answer the question regarding what is fair, or gave very complex, uncodable responses. It seems that the meaning of "fairness" in this context was somewhat alien for this group. Those who did respond were much more likely to say that little or no contribution was "fair" (Marwell & Ames, 1981, p. 308).

This explanation differs from the first: it suggests that economics students may have been as concerned about fairness as other students, but that they had a different notion of what was fair (Wang et al., 2011, p. 11).

A third theoretical mechanism suggests that economics students behave more selfishly due to their greater skepticism about the fair behavior of other people (Lanteri, 2008a, 2008b; López-Pérez & Spiegelman, 2009). This mechanism is derived from more general theorizing on social norms (Bicchieri, 1990, 2006; Messick, 1999; Weber et al., 2004). According to social norms theories, people first define a social situation as an exemplar of a class of social situations with which they are familiar (e.g., this situation resembles situations of class A). They then associate behavioral rules with that class of social situations (e.g., in situations of class A the rule is to split the endowment about equally). The underlying assumption is that people prefer to comply with the associated rule if two conditions are fulfilled: The individual expects (1) that rule following is what is the normative standard (e.g., what is fair) and (2) other people follow the rule as well (Bicchieri & Xiao, 2009; Kallgren et al., 2000). The hallmark of social norms is that people are willing to sanction rule deviant behavior of others—even that of third parties—at a cost to themselves (Bernhard, Fischbacher, & Fehr, 2006). Extrapolated to selfish behavior, economics students may thus share the motives and the notion of fairness, yet they simply do not expect others to behave fair (Lanteri, 2008a, 2008b). This skepticism can give the impression that selfishness is justified or even desirable (Ferraro, Pfeffer, & Sutton, 2005; D. T. Miller, 1999; Ratner & Miller, 2001). As a consequence, the economics students' greater skepticism would make them feel less obliged to behave fairly themselves and less willing to sanction the norm deviant behavior of others. Overall, the theoretical mechanism of the social norms hypothesis thus assumes that economics students have similar normative standards. However, in contrast to other students economics students are more skeptical. This skepticism is reflected in more selfish choices and a decreased willingness to sanction the unfair behavior of others. It is worth noticing that the social norms hypothesis assumes that economics students other students are equally motivated to comply with social norms. I return to this point in the discussion section.

To the best of my knowledge, no study dissected and measured the relative effect of the three mechanisms to explain why economics students behave more selfishly than other students do. The aim of this study is to fill that gap by means of an experimental game that was likely unfamiliar to all participants at the time the experiment was conducted. The following hypotheses are tested:

Hypothesis 1. Relative to other students, economics students are less concerned with fairness when making decisions and therefore behave more selfishly.

Hypothesis 2. Relative to other students, economics students have a different notion of fairness and therefore behave more selfishly.

Hypothesis 3. Relative to other students, economics students expect other participants to make more selfish decisions and therefore behave more selfishly. The skepticism also makes them less willing to sanction the norm deviant behavior of others.

METHODS

Participants

The study was conducted at a major British university at the end of the academic year. Undergraduate students from various semesters were recruited via the university's weekly bulletins. Altogether, 176 students participated in the online study. Eleven participants had to be excluded from the analysis: five participants did not understand the experiment, six did not reveal their field of study. Of the 165 remaining participants, 42 studied economics, 60 studied an art major, 63 studied a science major. The median age was 20 years (M_{age} = 20.81, SD_{age} = 3.41) and 104 participants were female. Because the proportion of women studying arts (M_{women} = 75%) was greater than the proportion of women studying economics (M_{women} = 55%) and sciences (M_{women} = 57%) analyses include gender as a predictor. The Research Ethics Committee of the Department of Psychology of the University of Cambridge approved the experiment and its consent procedure. To participate in this study all participants provided their informed consent. The database and the source code can be found online at https://osf.io/q9tjg/.

Materials and Procedure

The study involved a third-party punishment game (Fehr & Fischbacher, 2004; Fehr, Fischbacher, & Gächter, 2002). The game consisted of two stages and three roles, which were labeled A, B, and C to avoid evoking specific behaviors (cf. Chapter 2; Gerlach & Jaeger, 2016; see Appendix 3 for full instructions). For simplicity's sake, in the following, the roles are referred to as proposer, receiver, and judge, respectively. All participants rotated through all three roles in a fixed order. Decisions were made sequentially and the situation did not repeat itself. About one month after the experiment, the decisions of seven randomly picked groups (21 participants) were matched and the participants were paid in accordance with the outcome of the experiment. The time delay was necessary to complete data collection before matching. All participants were anonymous and fully informed about all aspects of the experiment. There was no deception involved. The average payment was £6.33 among the disbursed participants (about $9.50).

Figure 3.1 illustrates the two stages of the third-party punishment game. In stage one, proposers were the sole decision-makers. Receivers and judges remained passive. Proposers were endowed with £12 and could offer the receiver any amount between £0 and £12 in whole pounds sterling while themselves keeping the remainder. If, for example, a proposer offered a receiver £5, the proposer's income at the end of stage one was £7 and the receiver's was £5. If the proposer offered £0, the proposer's income at the end of stage one was £12 and the receiver's was £0. Stage one is thus similar to a dictator game (Forsythe et al., 1994), the most common measure of non-selfish motives. Studies suggest that offers in third-party punishment games positively and strongly correlate with donations in dictator games (Henrich et al., 2006).

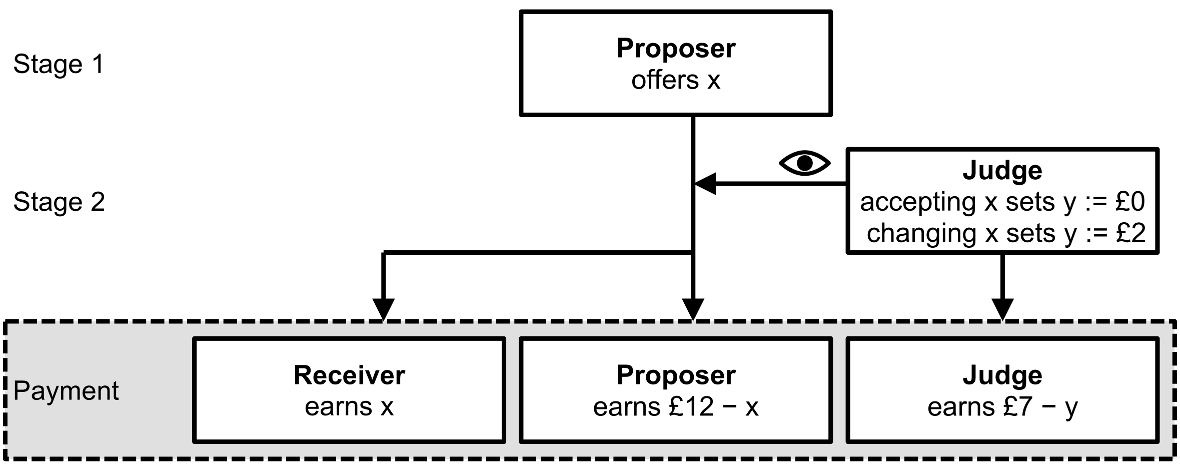


Figure 3.1. *Configuration of the third-party punishment game.*

In stage two, judges were the sole decision-makers. Proposers and receivers were passive. Judges could either accept or veto the proposer's offer. If judges accepted the offer, proposers and receivers were paid in accordance with stage one. The judges earned £7 and the transaction was complete. If, however, judges vetoed the offer, their earnings were reduced to £5, and they proceeded to re-allocate the £12 between proposer and receiver. For example, in stage one, a proposer might have offered £0. In stage two, the judge vetoed the offer and changed it to £9. In this case, the proposer's income was £3 (= £12 – £9), the receiver's £9, and the judge's £5 (= £7 – £2). If the judge instead accepted the offer, the proposer's income would be £12 (= £12 – £0), the receiver's £0, and the judge's £7 (= £7 – £0). The latter outcome is consistent with game theory: Judges are predicted to accept any offers proposed because vetoing does not benefit them personally; on the contrary, it costs them money. Accordingly, proposers are expected to anticipate that judges will accept any offer and consequently to offer receivers £0.

The third-party punishment game involves two incentivized decisions being made by the two active roles: proposers and judges. In each role, selfish motives compete with other potential motives. Judges may decide to veto a proposer's offer and reallocate the £12. In so doing, they can establish an equal distribution of money or punish the proposer by allocating a larger share to the receiver. Yet this redistribution comes at a financial cost of £2. A judge may therefore be tempted to permit an unfair offer and keep the full £7. Vetoing can never be in a judge's selfish interests. Nevertheless, judges may find it worth forsaking £2 to re-distribute the £12 and establish a fair outcome. Vetoing thus reflects judges' willingness to punish others for violating behavioral rules—such as, fairness norms—at a cost to themselves (Bernhard et al., 2006; Fehr et al., 2002).

The motives of proposers are less evident. Proposers may be motivated to maximize their own income, to maximize the receiver's income, to maximize the group's income, to establish a fair split, and so on. If proposers choose selfishly and strive to maximize their own payments, they must act strategically and be aware of the judges' power of veto: At what threshold are judges willing to step in? If proposers are primarily interested in a fair distribution, these considerations are less salient; proposers may assume that judges have no reason to veto an offer that is perceived as fair. To provide insights into the proposers' motives, all participants in this role were asked to comment on their choice: "In two to four sentences, please explain the reason behind your decision." This question served primarily to identify the aspect of the situation to which the participant paid most attention. It was assumed that an open-ended question was a less leading way of eliciting participants' motives than a question directly asking whether fairness concerns were involved (as, for example, in Marwell & Ames, 1981).

Two graduate students in psychology categorized comments into two categories. All comments that mentioned fairness as relevant to the decision were classified as reflecting concerns for fairness. Comments that did not mention fairness were classified as *not* reflecting concerns for fairness. The two coders were blind to the participants' majors and the hypotheses of the study. In cases of disagreement, the coders met to discuss and reconcile discrepancies. The classifications were used to operationalize fairness concerns (Hypothesis 1). After the transaction was completed, another question assessed participants' notion of fairness: "What would be a fair allocation of the £12 to the receiver?" Participants could reply to the question by stating a number between £0 and £12 or by replying "don't know" (Hypothesis 2). To operationalize expectations about the other participants' choices, receivers were asked—prior to making any payoff relevant decision—how much they expected the proposer would offer them. These expectations and the judge's vetoing behavior served as independent measures for the social norms hypothesis (Hypothesis 3).

The order in which participants rotated through the game was as follows: Participants were first asked about their expectations. Then, on one screen, they decided and explained their offer. Consequently, participants were confronted with a randomly assigned offer between £0 and £6, which they could veto or accept. Finally, participants were asked about their perception of a fair offer. The fixed order was chosen for three reasons. First, the question "What would be a fair allocation" was asked *after* all decisions had been made to avoid priming fairness. Second, participants were randomly assigned to offers *after* formulating their expectation and *after* deciding upon the offer to avoid anchoring specific values.[5] Third, participants had to formulate expectations before making an offer to properly test the mediation effect of expectations on decisions, as suggested by the social norms hypothesis.

[5] I also tested for potential spill over effect between the assigned offer and the subsequent fairness estimate. The assigned offer correlated with neither what was perceived as fair (ρ = –0.01, p = .916; Spearman's rank correlation) nor with the likelihood of responding "don't know" (t = –0.98, p = .337; Welsh test).

Results

Relative to Other Students, Economics Students Made More Selfish Decisions

A prerequisite for further analyses was that the third-party punishment game would replicate that economics students behave more selfishly. The offers made by economics students as proposers were indeed the least generous of all student groups. Their average offer (M = £2.83, SD = 2.56) was about £1.94 smaller than the average offers made by arts majors (M = £4.75, SD = 2.55; Z = 2.47, p < .001, r = 0.31) and science majors (M = £4.79, SD = 1.72; Z = 4.03, p < .001, r = 0.31; one-sided Wilcoxon rank-sum tests; arts vs. sciences: Z = 0.44, p > .21, two-sided Wilcoxon rank-sum tests). Figure 3.2 illustrates the offers made by study major.

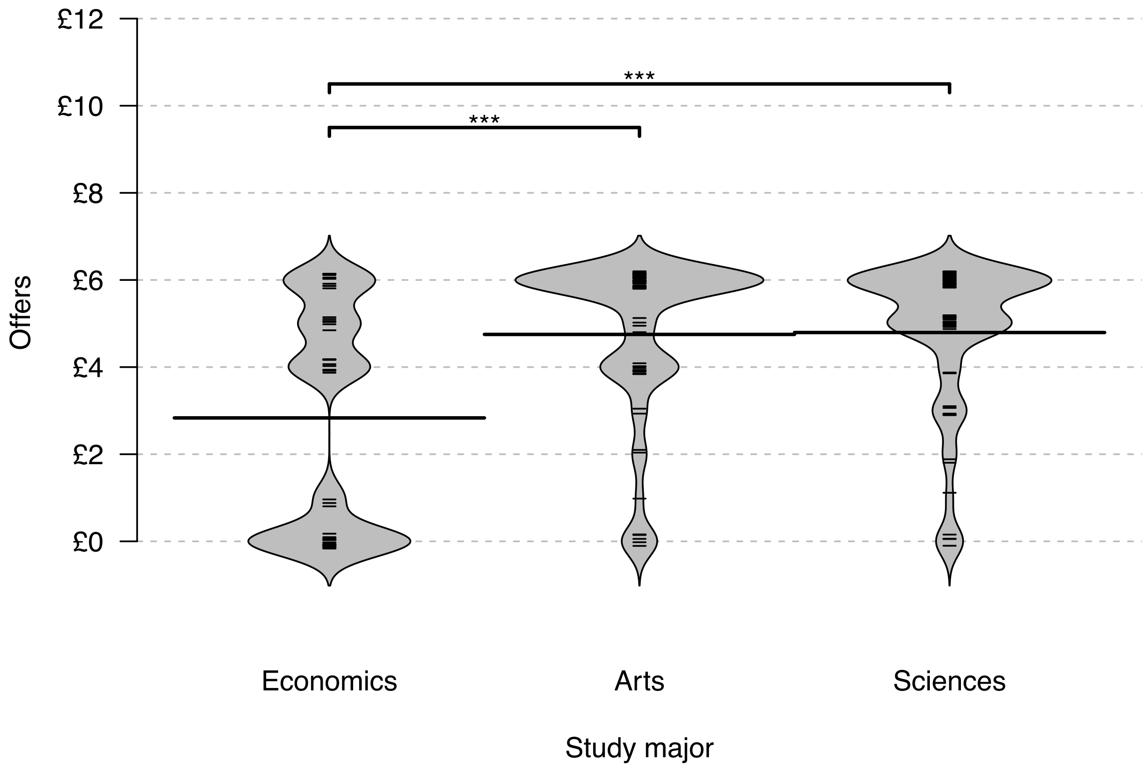


Figure 3.2. *Offers made by study major*. The bean plots shows the probability density of the offers per study major, the wider the area the more observation per offer. Each offer is also visualized through a dash. Dashes are vertically jittered for the sake of visualization only. The lower bars indicate the mean offer per major. The upper bars summarize the results of one-sided Wilcoxon rank-sum tests with *** p < .001.

There was also a main effect of gender: male students (M = £3.75, SD = 2.25) on average made £0.84 smaller offers than female students did (M = £4.59, SD = 2.09; Z = 2.98, p = .003, r = 0.23, two-sided Wilcoxon rank-sum test). Because gender was unequally distributed across the majors, Table 3.1 reports the results of a Tobit regression with economics major and gender as predictors. The negative effect of studying economics on the offer made persisted even when gender was controlled for.

Table 3.1.

Economics Students Made Lower Offers

Predictor	*b*	*SE*	*z*	*p*
(Intercept)	4.23	0.32	13.17	< .001
Economics	2.25	0.42	5.41	< .001
Female	0.73	0.37	1.95	.051

Note. Tobit regression with offers as the dependent variable. N = 165, log-likelihood: –344.73, df = 326, iterations = 7.

The pattern of economics students behaving more selfishly was thus reproduced in the third-party punishment game. I now turn to the analysis of the three mechanisms potentially underlying this pattern of results.

Relative to Other Students, Economics Students Were About as Often Concerned with Fairness when Making their Decisions.

Of the 165 participants, 97 (59%) were classified as mentioning fairness in their comments. For example, one participant offered £6 and commented: "£6 is an equal amount split between the two participants—it seems *fair* that I share the money equally" (emphasis added). Participants mentioning fairness did not necessarily make higher offers, however, as illustrated by this comment on a £3 offer: "This split is deliberately *unfair* in my favor, but not so massively *unfair* that I think C [the judge] would intervene to change the split at the cost of £2 of his own endowment. I would expect C's threshold to *unfairness* to be higher before he steps in to even" (emphasis added). The remaining 68 participants (41%) did not mention fairness. For example, one participant offered £5 and wrote "C [the judge] can change my decision anyway, and C would like to have 7 pounds instead of 5." Another participant offered £0, commenting that: "Assuming B [the receiver] is a complete stranger, I owe nothing to B. I have no reason to give any of the £12 to B."

To rule out that fairness was mentioned only in the context of strategic considerations (as suggested by the second comment), a Fisher's exact test assessed whether mentioning fairness and mentioning the judge were co-occurring. This was not the case ($p = .155$).

Overall, 48% of the economics students mentioned fairness. Compared to science majors (M = 59%, $p = .319$) the proportion of economics students who mentioned fairness was likely the result of sampling error. Compared to arts majors the proportion approached significance (M = 67%, $p = .067$; arts vs. sciences: $p = .456$; all Fisher's exact tests). However, the unequal distribution of gender cofounded the effect (although there was no main effect of gender: males: M = 52%; females: M = 63%; $p = .252$). To distinguish effects of gender and major, Table 3.2 presents the results of a binary logistic regression model with major and gender as covariates. Neither major nor gender predicted references to fairness concerns.

Table 3.2.

Neither Studying Economics Nor Gender Predicted References to Fairness.

						95% *CI*	
Predictor	*b*	*SE*	*z*	*p*	*OR*	2.5%	97.5%
(Intercept)	0.28	0.28	0.98	.325	1.32	0.28	0.83
Economics	0.57	0.36	1.58	.114	0.56	1.29	0.14
Female	0.37	0.33	1.11	.268	1.44	0.28	1.01

Note. Binary logistic regression with fairness references (yes = 1, no = 0) as the dependent variable. N = 165, adjusted R2 = 1%, Nagelkerke's R2 = 3%, iterations = 4.

Relative to Other Students, Economics Students Had a Similar Notion of Fairness in the Situation.

After the game ended, participants were asked "What would be a fair allocation of the £12 to role B [the receiver]?" They could state any figure between £0 and £12 or give the response "don't know." Figure 3.3 depicts the findings for those who responded. Students of economics ($n = 35$, M = £5.06, $SD = 1.78$), arts majors ($n = 60$, M = £5.27, $SD = 1.43$, $p = .728$) and sciences majors ($n = 58$, M = £5.32, $SD = 1.20$, $p = .147$) appeared to have similar notions of fairness (all one-sided Wilcoxon rank-sum tests; arts vs. sciences: $p = .677$, two-sided Wilcoxon rank-sum tests).

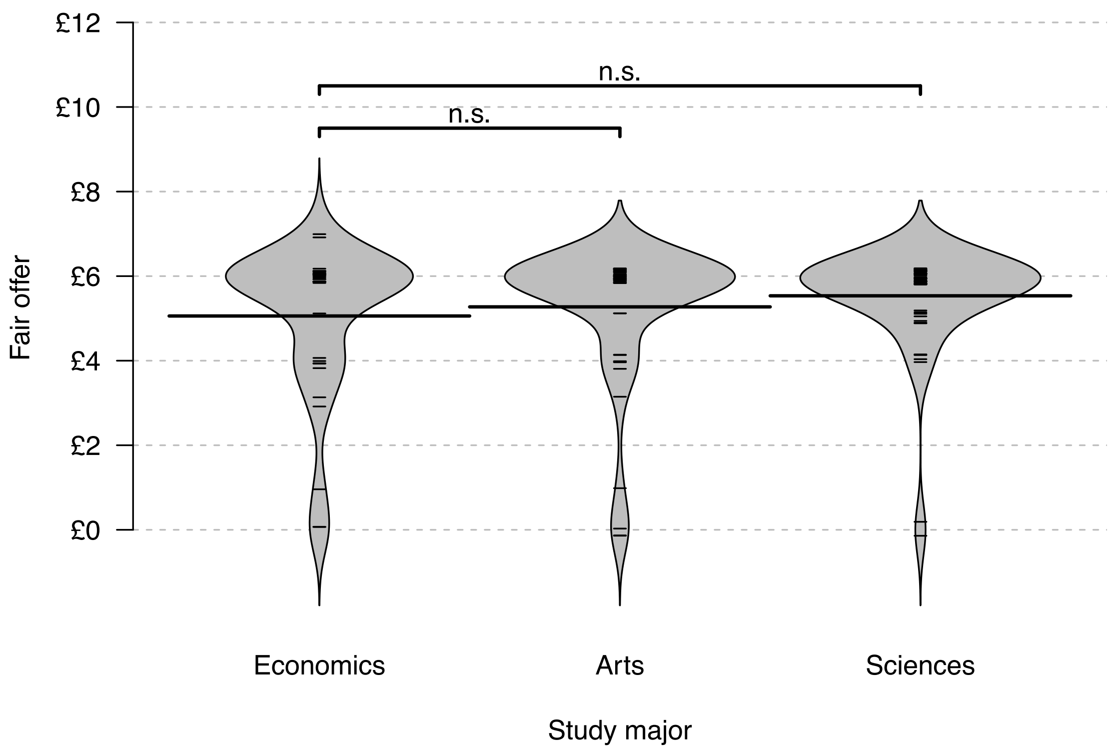


Figure 3.3. *Responses to "What would be a fair allocation" by major.* Bean plots and one-sided Wilcoxon rank-sum tests with n.s. = not significant

On average, men stated £0.65 smaller offers to be fair ($n = 53$, $M =$ £4.91, $SD = 1.89$) than women did ($n = 95$, $M =$ £5.56, $SD = 1.19$, $p = .008$). The Tobit regression model in Table 3.3 confirmed the gender effect. Yet study major did not predict the offer considered to be fair.

Moreover, there were no differences between economics students and other students in the percentage of participants answering "don't know" (economists: $M = 17\%$; arts: $M = 8\%$, $p = .225$; sciences: $M = 8\%$, $p = .215$; Fisher's exact tests). Neither was there a gender effect (males: $M = 13\%$; females: $M = 9\%$; $p = .429$; Table 3.4).

Table 3.3.

Studying Economics Did Not Predict the Notion of Fairness but Gender Did.

Predictor	*b*	*SE*	*z*	*p*
(Intercept)	4.95	0.22	22.23	< .001
Economics	0.35	0.30	1.17	.244
Female	0.68	0.26	2.59	.010

Note. Tobit regression with responses to "what would be a fair allocation" as the dependent variable. N = 148, log-likelihood: –272.55, df = 292, iterations = 5.

Table 3.4.

Neither Studying Economics Nor Gender Predicted the Response "Don't Know"

						95% *CI*	
Predictor	*b*	*SE*	*z*	*p*	*OR*	2.5%	97.5%
(Intercept)	2.18	0.45	4.86	< .001	0.11	3.06	1.30
Economics	0.77	0.53	1.45	.147	2.17	0.27	1.82
Female	0.39	0.52	0.75	.451	0.67	1.42	0.63

Note. Binary logistic regression with the reply "don't know" to "what would be a fair allocation". N = 17, adjusted R^2 = 1%, Nagelkerke's R^2 = 3%, iterations = 5.

Relative to Other Students, Economics Students Expected Other Participants to Make More Selfish Decisions.

In their role as receivers, all participants were asked how much they expected their assigned proposer to offer. Of all groups, economics students were the most skeptical, expecting on average £1.22 (M = £2.88, SD = 2.21) smaller offers than arts majors (M = £4.21, SD = 1.88; Z = 2.11, p = .001, r = 0.25) and science majors did (M = £3.98, SD = 1.84; Z = 2.57, p = .005, r = 0.32; one-sided Wilcoxon rank-sum tests; arts vs. sciences: p > .342, two-sided Wilcoxon rank-sum tests). Figure 3.4 illustrates the distributions of expectations across the majors.

There was also a main effect of gender: male students on average expected to be offered £0.77 less (M = £3.30, SD = 2.05) than female students did (M = £4.07, SD = 1.95; Z = 2.60, p = .009, r = 0.20, two-sided Wilcoxon rank-sum test). Table 3.5 presents the results of a Tobit regression model with economics major and gender as predictors. Both covariates predicted the offers expected, with economics major weighing heavier than gender.

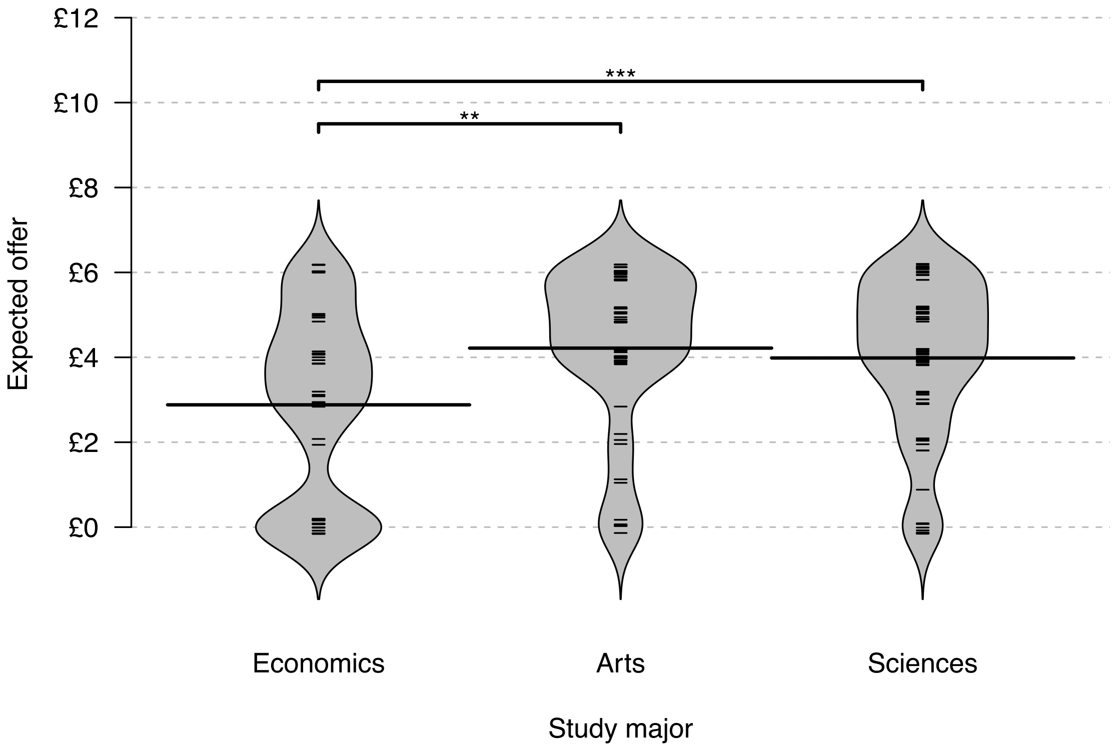


Figure 3.4. *Expected offers by major.* Bean plots and one-sided Wilcoxon rank-sum tests with ** $p < .01$, *** $p < .001$.

Table 3.5.

Economics Students Expected Lower Offers.

Predictor	*b*	*SE*	*z*	*p*
(Intercept)	3.52	0.31	11.30	< .001
Economics	1.39	0.40	3.46	< .001
Female	0.75	0.36	2.07	.038

Note. Tobit regression with offers as the dependent variable. $N = 165$, Log-likelihood: –340.72, $df = 326$, iterations = 6.

Are the smaller offers made by economics students attributable to their lower expectations? To answer this question I built a mediation model, in which a linear regression model for expectations served as a mediator for an "outcome model." The "outcome model" was similar to the regression model documented in Table 3.1 but it included expectations as an additional, independent predictor. All models included gender as a covariate. This setup allows to decompose the weights of the dissimilar offers of economics and non-economics students *(τ)* into a mediation effect of expectations *(δ)* and a direct effect as the non-explained remainder *(ζ)* by means of bootstrapping and Monte Carlo simulations with the mediator held constant (for details on causal mediation analysis, see Tingley, Yamamoto, Hirose, Keele, & Imai, 2014). The proportion mediated can be estimated as the quotient of δ/τ. On average, expectations (δ = 2.37, p < .001) mediated about 53% of the difference between the offers of economics and non-economics students (τ = 4.45, p < .001; ζ = 2.08, p < .001).

Social norms theories also hypothesized that, due to their skepticism of the fair behavior of other people, economics students would be less willing to sanction the rule disconfirm behavior of others. To test this prediction the decisions of the judges were analyzed. For each participant I calculated whether the observed offer was perceived as unfair (i.e., whether the observed offer was less than what was perceived as fair). Among the 119 judges who were presented with offers that they had perceived as unfair economics students (M = 16%) were about 3 to 4 times less likely to veto than arts majors (M = 48%, p = .009, OR = 0.21) and sciences majors were, respectively (M = 60%, p < .001, OR = 0.13; arts vs. sciences: p = .300, Fisher's exact tests). Figure 3.5 plots the likelihood of a veto by the judge's major as a function of the observed offer.

There was *no* main effect of gender on vetoing (males: M = 41%; females: M = 49%; p = .442, OR = 1.36). Table 3.6 presents the results of a binary logistic regression model with vetoing as the dependent variable. When gender was controlled for, economics students were about 6 (= 1 / 0.16) times less likely to veto unfair offers than other students were.

An overview of the judge's redistributions is provided in Table 3.7. The most popular choice (made by 27 of the 56 vetoing players) was an even split of £6 between proposer and receiver.

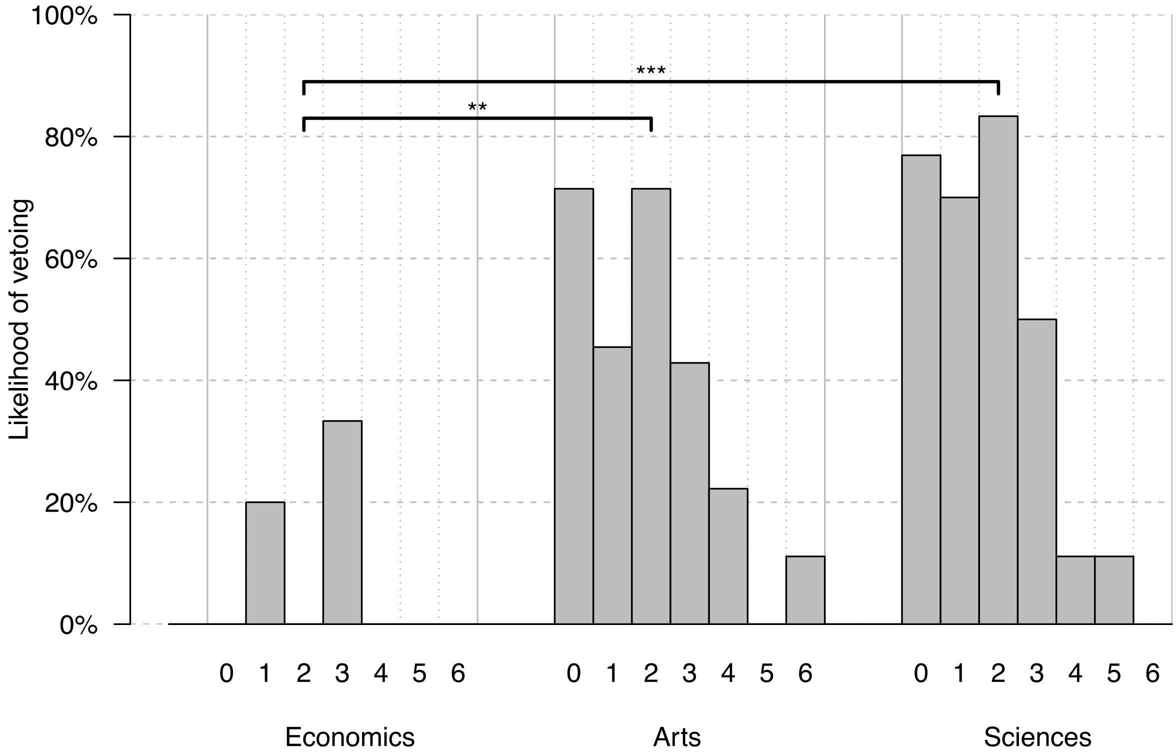


Figure 3.5. *Offers vetoed by study major.* Bar plots. The x-axis indicates the observed offer per study major. The y-axis indicates the probability of vetoing as a function of the observed offer (x-axis). Upper bars indicate the results of Fisher's exact tests for vetoes on offers that were perceived as unfair with ** $p < .01$, *** $p < .001$.

Table 3.6.

Economics Students Were Less Likely to Veto Offers that Were Perceived as Unfair

						95% CI	
Predictor	*b*	*SE*	*z*	*p*	*OR*	2.5%	97.5%
(Intercept)	1.05	0.47	2.23	.026	2.86	0.13	1.97
Economics	1.81	0.63	2.88	.004	0.16	3.05	0.58
Female	0.47	0.46	1.03	.304	1.60	0.42	1.36
Observed offer	0.53	0.13	4.07	< .001	0.59	0.74	0.28

Note. Binary logistic regression with vetoing an offer that perceived as unfair (1 = veto, 0 = accept) as the dependent variable, $N = 119$, adjusted $R^2 = 42\%$, Nagelkerke $R^2 = 33\%$, iterations = 4.

Table 3.7.

Redistribution of Vetoed Offers.

	Chosen redistribution to receiver								
Assigned offer	4	5	6	7	8	9	10	11	12
0	0	0	3	5	3	0	1	0	3
1	1	1	8[a]	1	0	0	0	2	0
2	1	2	9	1	1	0	1	0	0
3	0	0	4[a]	2[a]	0	1[a]	0	0	1
4	0	1	2	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0
6	0	0	1	0	0	0	0	0	0
Total	2	4	27	10	4	1	2	2	4

Note. [a] indicates that there was exactly one economics student in the group.

Discussion

This study demonstrated that, relative to their fellow students, economics students offered less in a third-party punishment game; they were about equally likely concerned with fairness, and they had a similar understanding of what was fair. However, economics students expected to receive smaller offers from others, which in turn mediated their own smaller offers. Moreover, economics students were less willing to veto unfair allocation of others. Taken together, the results suggest that economics students' more selfish behavior is not due different fairness standards but to social norms.

Several limitations to this investigation and the conclusions that may be drawn from it warrant consideration. Foremost, economics students' expectations only *partly* mediated the offers they made, which suggests that their greater skepticism is not the only explanation for their more selfish behavior. Further, participants may have decided how much to offer *before* forming expectations about others. If this were the case, causality would be reversed, with decisions informing expectations (in the sense "If I do B in the situation of class A, expect B to be the rule for situations of class A"). Reverse causality was addressed primarily through the study design: expectations were measured *before* participants decided on an offer. Nonetheless, it cannot be completely ruled out that participants made a hypothetical decision immediately after reading the instructions (Thoemmes, 2015).

Moreover, offers in the third-party punishment game cannot be unequivocally interpret as due to non-selfish motives only. As suggested by some of the participants' comments, high offers can be due to strategic motives, to avoid vetoing. Although studies suggest that offers in the third-party punishment positively and strongly correlate with established measures of selfishness (Henrich et al., 2006) and the proposer's offers primarily served as a replication of economics students' greater propensity to behave selfishly, a combination of experimental games would have been preferable to directly test the economics students' more selfish motives. Such combination of experimental games could also shed light on differences in the motives to veto. In the particular version of the third-party punishment game vetoing always caused judges to earn less than what one or both other participants would earn. However, third-party punishment in general can be motivated not only by norm enforcement but also by spiteful motives (Leibbrandt & López-pérez, 2012). Further experiments could not only assess the students' differences in their motives to veto but also test a central assumption of the social norms hypothesis, namely whether economics students and other students are similarly motivated to comply with social norms.

Another limitation is the cross-sectional design of the study. Why are economics students more skeptical in the first place? There are two possible explanations. One is that people who are more skeptical and behave more selfishly are drawn to study economics (self-selection hypothesis). The other explanation is that economics students learn to associate specific situations with more selfish behaviors (socialization hypothesis). It is also possible that both explanations hold. The two effects have been extensively discussed in the literature (Frey, Pommerehne, & Gygi, 1993; Scott & Rothman, 1975; see Kirchgässner, 2005 for an overview). In the specific context of experimental games, the findings of cross-sectional studies that correlate academic year with choices are frequently interpreted as pointing to socialization (Frank et al., 1993; Haucap & Müller, 2014; López-Pérez & Spiegelman, 2009) rather than self-selection (Carter & Irons, 1991), suggesting that economics students learn to behave more selfishly over the course of their studies. This interpretation, however, must not be valid. A positive correlation of selfish behaviors with study year could be as much due to learning (socialization) as to a systematic dropping out of less selfish economics students (self-selection) as to a third factor, such as the effect of aging in general. Cross-sectional data simply cannot be used for inferring changes within a sample over time. To assess the underlying mechanisms of what is driving the more selfish behavior of economics students longitudinal data sets that include a control group are required.

It is nonetheless possible to speculate that the body of theories on human behavior to which economics students are exposed during their studies can explain their greater skepticism and their more selfish behavior. Economic theories have traditionally been more concerned with the mathematical structure of the decision problem than with the psychology of the individual (Gigerenzer & McElreath, 2003). Rational choice theory, arguably one of the centerpieces of modern economic theory (Sugden, 1991) and the starting point of game theory (von Neumann & Morgenstern, 1944), frames decision-making as a calculative, emotionless weighing of the costs and the benefits of options. Although rational choice theory primarily serves to *describe* decision-making, its application to social situations may cause economics students to *interpret* a situation differently, ultimately leading to different choices (Blais & Young, 1999; Brunk, 1980; Stigler, 1981). For example, Fiske argued that in Western cultures taking calculative approaches to social relations is associated with greater psychological distance (A. P. Fiske, 1991a), less intense moral obligations (A. P. Fiske & Tetlock, 1997), and more selfishness (A. P. Fiske, 1991b). Economics students may thus not only learn rational choice theory but also learn to associate it with a specific behavior that they in turn expect from others. Rubinstein, for example, pointed out that students who come to the university "to 'study economics' instead become experts in mathematical manipulations" (Rubinstein, 2006, p. C1). As a consequence, economics students may be more committed to maximize profits rather than to sympathize with other individuals. Especially through being exposed to game theory in their studies, economics students who participate in game experiments find themselves in social situations for which they have learned the "correct" calculative approach (Frank et al., 1993). This does not necessarily mean that economics students are more skeptical and behave more selfishly outside the context of games. Research combining game experiments with field studies would be needed to test how well the choices of economics students—and other students—actually predict their behavior outside the laboratory.

The fact that economics students behave more selfishly than other students is rather critical for experimental practice. Experimental games aim at extrapolating findings from the laboratory to the world beyond (Levitt & List, 2007). Yet most experimental games are exclusively conducted among university students—especially economics students (Gerlach, Teodorescu, & Hertwig, 2017). Although research suggest that students behave rather similar to other population groups (Exadaktylos, Espín, & Brañas-Garza, 2013; Gerlach, Teodorescu, et al., 2017) variation within the student participant pool is frequently neglected, potentially constraining generalization (I return to this issue in Chapter 4).

Conclusion

To conclude, this study demonstrated that economics students behaved more selfishly than other students in a third-party punishment game. Analyses of three mechanisms potentially underlying this pattern of results suggest that the more selfish behavior is not due to differences in fairness concerns or notions of fairness, but to the greater skepticism among economics students. This finding sheds new light on the debate about potential links between studying economics and selfish behavior. Selfish behavior is not free from context and can have different motives. In some contexts, economics students behave more selfishly because they expect others to do so.

Chapter 4

The Truth About Lies — A Meta-Analysis on Dishonest Behavior[6]

The Enron accounting scandal, WorldCom's Ponzi scheme, Fifa's web of corruption, the Volkswagen Dieselgate emissions scandal, the Petrobras and Odebrecht bribery cases in Brazil—these are just a few examples of widespread dishonesty and fraud worldwide. According to Transparency International's annual global survey of corruption levels, over two-thirds of the 176 countries and territories in the Corruption Perceptions Index 2016 fell below the midpoint of the scale from 0 (highly corrupt) to 100 (very clean; Transparency International, 2017). The global average score is a measly 43, indicative of endemic corruption. It seems that dishonesty is a widespread phenomenon. Against this background, it is not surprising that dishonesty has become a research topic in its own right. How can it be that so many seemingly normal and well-adjusted people behave dishonestly to such an extent that their behavior gravely harms others? Will anybody succumb to dishonesty in the "right" situation, or is there a dishonest personality type? These and related questions have a long tradition in experimental research, which has sought to reveal the dynamics behind the dark sides of human nature—such as blind obedience to authority or gawking bystanders' failure to render help—by running highly controlled behavioral experiments (e.g., Darley & Latané, 1968; Milgram, 1974).

Within the last decade, experiments examining the prevalence and magnitude of dishonesty, as well as its enabling conditions, have generated a large body of empirical findings across the behavioral sciences, including behavioral economics (Fischbacher & Föllmi-Heusi, 2013; Gneezy, 2005), neuroscience (Greene & Paxton, 2009), and psychology (Mazar, Amir, & Ariely, 2008). Unlike past research using qualitative case studies and/or surveys to measure self-reported dishonesty, this new line of research harnesses the tool of experimentation to quantify manifestations of dishonest *behavior*. Despite marked differences in their experimental details (Rosenbaum, Billinger, &

---

[6] This chapter is based on: Gerlach, P. Teodorescu, K., & Hertwig, R. (2017). The truth about lies. A meta-analysis on dishonest behavior. *Manuscript under review at Psychological Bulletin.*

Stieglitz, 2014), the experiments all generate a basic conflict between the temptation to behave dishonestly and the capacity to resist that temptation.

Several reviews of when and why people engage in (dis)honest behavior have already been published (Gino, 2015; Gino & Ariely, 2016; Jacobsen, Fosgaard, & Pascual-Ezama, 2017; Rasmußen, 2015; Rosenbaum et al., 2014). Although valuable and informative, these reviews have relied on narrative summaries, synthesizing the empirical findings on a case-by-case basis. The aim of this meta-analysis is to complement the narrative reviews by providing a systematic and statistical synthesis of the experimental findings. Meta-analyses have several advantages over narrative summaries, as will be outlined below. We therefore applaud a recent endeavor to meta-analytically integrate empirical findings (Abeler, Nosenzo, & Raymond, 2016). In their quantitative review, Abeler, Nosenzo, and Raymond formalized a range of theories about when and why people engage in (dis)honest behavior and tested those theories against the data. Our meta-analysis extends their work by including further experimental paradigms, allowing us to pursue three broad goals.

Our first goal is to quantitatively synthesize the empirical findings on dishonest behavior yielded by four of the most popular experimental paradigms. Examining variations within and between the paradigms most commonly used to study dishonest behavior allows us to answer several open questions: Does the literature paint an adequate picture of the prevalence of dishonesty? Do studies with low power bias this picture (publication bias)? Do different experimental paradigms lead to different conclusions about the circumstances under which people behave dishonestly? Our quantitative answers are based on the combined empirical data of more than 95 experiments per experimental paradigm.

Our second goal is to examine personal and situational factors associated with dishonest behavior. Do greater rewards prompt more dishonest behavior? Do laboratory and online experiments yield similar conclusions about dishonest behavior? Are student samples representative for the degree of dishonesty in the population? Do men behave more dishonestly than women? Narrative reviews can provide useful summaries of the debates on situational and personal factors impacting dishonest behavior. Yet the quantitative nature of meta-analyses makes it possible to estimate the degree to which each factor promotes dishonest behavior. For example, by integrating even those articles that do not mention gender differences in our analyses, we can provide a more unbiased evaluation of potential gender differences than narrative reviews can.

The third goal of our meta-analysis is to examine the interactions of the experimental paradigms with the personal and situational factors. The conclusions emerging from the literature as to when people engage in (dis)honest behavior are to some extent unclear, or even contradictory. For example, whereas some experimental paradigms suggest that increasing the reward size leads to more dishonest behavior (Gneezy, 2005; Sutter, 2009), others suggest null effects (Fischbacher & Föllmi-Heusi, 2013; Gächter & Schulz, 2016; Hugh-Jones, 2016; Mazar et al., 2008). Meta-analytical techniques make it possible to assess the circumstances under which the size of the temptation affects dishonest behavior. Our results can thus inform future experimental research on the enabling and disabling conditions of dishonest behavior.

Before presenting our findings, we first introduce the four experimental paradigms included in this meta-analysis. We then summarize previous empirical findings, highlighting open questions and conflicting patterns of results. After outlining our data set, we then introduce our standardized measures that allow us to classify and compare different types of dishonest behavior.

How Is Dishonest Behavior Measured?

Experiments assessing dishonest behavior are relatively heterogeneous, thwarting easy comparison and replication (Rosenbaum et al., 2014). In order to be able to compile and analyze comparably designed experiments, we focused on four of the most widely employed experimental paradigms that assess dishonest behavior: sender–receiver games, coin-flip tasks, die-roll tasks, and matrix tasks.

In *sender–receiver games* (Gneezy, 2005), two participants interact. One, the *sender*, learns about the payoffs of two or more options. The sender then decides which message he wants the other participant, the *receiver*, to read. For example, the sender can send a false message stating “Option A will earn you more money” or a true message stating “Option B will earn you more money.” After reading the message, the receiver chooses between the options, not knowing the actual payoffs. However, it is the decision of the receiver that determines the one-off payoffs for both players. In sender–receiver games, the total earnings of the two players are typically a constant sum. That is, regardless of the sender’s message and the receiver’s choice, one party’s gain is the other party’s loss.

Senders thus face a dilemma between sending a truthful and sending a false message. The truthful message, if believed and acted upon by the receiver, will result in a lower payoff for the sender.

In *coin-flip tasks* (Bucciol & Piovesan, 2011) participants are asked to report private information on a randomly and self-generated dichotomous outcome, typically the result of a coin toss. Reporting one outcome (e.g., heads) wins participants some reward; reporting the other outcome (e.g., tails) leaves them empty handed. The rate of dishonest reporting can only be estimated on the aggregate level. If data from a sufficiently large number of participants are combined, experimenters can compare the proportion of reported wins (e.g., 75% heads) to the theoretical baseline of randomly generated wins (e.g., 50% heads). Thus, in contrast to sender–receiver games, coin-flip tasks do not allow *individual* dishonest behavior to be directly observed or spotted. Moreover, participants typically interact with the experimenter rather than with another participant. What happens to the experimental money that is not paid out to participants is usually not specified: the more people misreport, the more money is paid out (positive sum game). The negative consequences of dishonest behavior are thus less clear in coin-flip tasks than they are in sender–receiver games.

In *die-roll tasks* (Fischbacher & Föllmi-Heusi, 2013; Fischbacher & Heusi, 2008) participants are asked to report private information on a randomly generated continuous outcome, such as the roll of a die. Whereas the choice outcomes of sender–receiver games and coin-flip tasks are binary, die-roll tasks can have three or more possible outcomes per choice. Participants know in advance that each reported outcome is rewarded by a corresponding amount. For example, reporting a 1 pays $1, a 2 pays $2, a 3 pays $3, and so forth. Like coin-flip tasks, die-roll tasks do not detect dishonest behavior at the individual level. Instead, experimenters can estimate the degree of dishonest behavior by comparing the mean outcome reported at the aggregate level (e.g., on average, people reported a score of 4.2) with a hypothetical, randomly generated distribution of outcomes (e.g., on average, a fair die would yield a score of 3.5). Similar to coin-flip tasks, die-roll tasks are typically positive sum games, in which the victim of any dishonest behavior is not clearly defined and in which the more people cheat, the more money is paid out across the sample of participants.

In *matrix tasks* (Mazar et al., 2008), participants are typically presented with several matrices, each containing 12 three-digit numbers (e.g., 4.56). The task is to find the number pair in each matrix that adds up to exactly 10.00 (e.g., 4.56 + 5.44). However, most participants do not manage to solve all matrices in the time allotted. When time has run out, participants are asked count how many matrices they have solved and then to pay themselves accordingly (e.g., $1 per solved matrix). Dishonest behavior in matrix tasks can be measured at either the aggregate level (as in coin-flip and die-roll tasks) or the individual level (as in sender–receiver games). Matrix tasks that measure dishonest behavior at the aggregate level randomly assign participants to two groups: an experimental group and a control group. The answers of the experimental group are self-graded by the participants, whereas the answers of the control group are verified by the experimenter. The experimental group can thus cheat by inflating the number of allegedly solved matrices; the control group cannot. The amount of dishonest behavior can be estimated by comparing the total reported number of solved matrices from the experimental group with the actual number from the control group. Other matrix tasks allow dishonesty to be measured at the individual level. In this design, participants are given a collection slip in addition to the matrix sheet. After solving the matrices, they are asked to report the number of solved matrices on the collection slip. Unbeknownst to the participants, a unique identification code allows the matrix sheet to be matched to the collection slip. The experimenter thus has data on both the actual performance and the reported performance at the individual participant level. All matrix tasks, regardless of their design, use a continuous outcome measure that allows the degree of dishonest behavior to be identified—similar to die-roll tasks. Matrix tasks—like coin-flip and die-roll tasks—are typically positive sum games, in which cheating has no identifiable victim.

In sum, all four experimental paradigms involve information asymmetry and temptation: participants know more than the people who determined their payoffs. False information, if believed, is rewarded. Table 4.1 provides an overview of the four experimental paradigms, highlighting some of their key differences.

Table 4.1.

Typical Key Properties of the Four Experimental Paradigms

	Sender–receiver	Coin-flip	Die-roll	Matrix
Dishonest behavior	Sending a false message to another participant	Misreporting a randomly generated outcome	Misreporting a randomly generated outcome	Misreporting one's performance in a task
Measuring scale	Dichotomous	Dichotomous	Continuous	Continuous
Measuring level	Individual	Aggregate	Aggregate	Individual/ aggregate
Total payoffs	Constant	Positive	Positive	Positive
Identifiable victim	Yes	No	No	No

Note. Total payoffs refer to the typical sum of all participants' payoffs. In constant sum games, one participant's gain is the other's loss. Hence, the other participant is the identifiable victim of any act of dishonesty. In positive sum games, participants can earn more from being dishonest without inflicting a loss on another participant. Hence, there is no identifiable victim.

Studies directly comparing two or more of the four experimental paradigms are rare. To the best of our knowledge, such comparisons are limited to comparing die-roll tasks and matrix tasks (Gino, Krupka, & Weber, 2013; Gravert, 2013). Interestingly, the two paradigms regularly result in rather different estimates of dishonest behavior: whereas die-roll tasks typically find that at least some proportion of people improperly claim the maximum amount (e.g., Fischbacher & Föllmi-Heusi, 2013), it is often the case that few or none of the participants in matrix tasks lie to such full extent (Gino, Ayal, & Ariely, 2009; Mazar et al., 2008). The first goal of this meta-analysis is therefore to assess whether the experimental paradigms come to the same conclusions about dishonest behavior.

ACTING DISHONESTLY: THEORY AND DATA

The conventional economic model assumes that people are willing to misreport private information if the material incentives of acting dishonestly outweigh those of acting honestly (Becker, 1968). In theory, the prototypical *homo economicus* will engage in dishonesty whenever this behavior pays off. Yet experiments on dishonest behavior suggest that, in practice, people often behave otherwise: First, people acting like *homo* *economicus* only represent a fraction of all observations. A substantial proportion of individuals behaves completely honestly despite material incentives (Abeler et al., 2016; Fischbacher & Föllmi-Heusi, 2013). Second, the large majority of those who behave dishonestly do so only to the extent that they can appear honest (to oneself, in the form of internalized norms, or to others, in the form of social norms; Abeler et al., 2016; moral hypocrisy, see Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997; Dana, Weber, & Kuang, 2007; Schweitzer & Hsee, 2002; Tenbrunsel & Messick, 2004). That is, they often shy away from claiming the maximum potential payoff. Third, the degree to which people engage in dishonest behavior largely depends on situational and personal factors.

Personal Factors

Empirical investigations of dishonest behavior have looked at populations ranging from Franciscan nuns (Utikal & Fischbacher, 2013) to maximum security prisoners (Cohn, Maréchal, & Noll, 2015). Some of the personal factors most frequently assessed include gender, age, student status, and study major.

Gender. There has been a substantial debate on gender differences in dishonest behavior. Initially, Dreber and Johannesson (2008) found that 55% of men but only 38% of women sent false messages in sender–receiver games. Yet, two replication attempts failed to find similar results (Childs, 2012b; Gylfason, Arnardottir, & Kristinsson, 2013). On a broader scale, empirical evidence on gender effects seems far from clear cut. Whereas some studies have concluded that men behave more dishonestly than women (e.g., Cappelen, Sørensen, & Tungodden, 2013; Conrads, Irlenbusch, Rilke, & Walkowitz, 2013; Friesen & Gangadharan, 2012; Holm & Kawagoe, 2010; Houser, Vetter, & Winter, 2012; Ruffle & Tobol, 2014), others have found no gender differences (e.g., Abeler, Becker, & Falk, 2014; Aoki, Akai, & Onoshiro, 2013; Arbel, Bar-El, Siniver, & Tobol, 2014; Erat & Gneezy, 2012; Holm & Kawagoe, 2010; Lundquist, Ellingsen, Gribbe, & Johannesson, 2009), and one study even indicated that women behave more dishonestly than men (Ruffle & Tobol, 2014).

Age. An inconsistent picture has also emerged for age effects. Whereas some studies have found that younger participants behave more dishonestly than older participants (Conrads et al., 2013; Glätzle-Rützler & Lergetporer, 2015), others failed to find age effects (Abeler et al., 2014; Bucciol & Piovesan, 2011; Conrads & Lotz, 2015; Gino &

Margolis, 2011), and at least one study indicated that older people are more dishonest (e.g., Friesen & Gangadharan, 2012).

Students vs. non students. Most experiments have relied on student samples; relatively few have assessed the degree to which results obtained from student samples generalize to other population groups. Findings are conflicting. Some studies suggest that students behave more dishonestly than a more representative sample of the population (Abeler et al., 2014; Aoki et al., 2013; Fosgaard, 2016); another study found no systematic differences between students and more representative participant groups (Gunia, Barnes, & Sah, 2014).

Economics/business major. Numerous studies have compared students majoring in economics and business with other students. Some found that economics and business majors behaved more dishonestly than other students (Childs, 2012a, 2013; Lewis et al., 2012; Lundquist et al., 2009); others reported interaction effects with experimental factors or null effects (Gino, Krupka, et al., 2013; Muñoz-Izquierdo, Liaño, Rin-Sánchez, & Pascual-Ezama, 2014).

Situational Factors

Situational factors examined to date include the influence of the investigative setting, externalities, and the magnitude of the potential reward on dishonest behavior.

Investigative setting. Physical distance to the person one is lying to could be an enabling condition for dishonest behavior. Indeed more dishonesty has been observed in online studies than in laboratory studies (Abeler et al., 2014). In addition, payoff-maximizing responses are evidently more prevalent in online studies than in laboratory studies, whereas partial dishonesty (i.e., slightly overstating one's outcomes) is less prevalent in online settings (Conrads & Lotz, 2015).

Externalities. Dishonest behavior may be sensitive to the degree to which other people are harmed by or benefit from it (Erat & Gneezy, 2012; Faravelli, Friesen, & Gangadharan, 2015; Gino, Ayal, & Ariely, 2013; Gneezy, 2005; Hurkens & Kartik, 2009; Muñoz-Izquierdo et al., 2014; Rigdon & D'Esterre, 2014; Wiltermuth, 2011). In sender–receiver games, for example, one participant's gain is typically the other's loss (constant sum and identifiable victim). In coin-flip tasks, die-roll tasks, and matrix tasks, the externalities are commonly less clear (positive sum and no identifiable victim). Following Gneezy (2005), a number of sender–receiver games have manipulated the extent to which the other participant is harmed by trusting a false message, thereby changing the typical constant sum aspect of the game (Erat & Gneezy, 2012; Hurkens & Kartik, 2009; Sutter, 2009; Wang & Murnighan, 2016). Relatedly, payoffs in some coin-flip tasks (e.g., Muñoz-Izquierdo et al., 2014), die-roll tasks (e.g., Gino, Ayal, et al., 2013), and matrix tasks (e.g., Faravelli et al., 2015; Rigdon & D'Esterre, 2014) have been modified to constant sum games by imposing externalities on other participants (identifiable victim). Whereas findings on sender–receiver games suggest that greater externalities result in less dishonest behavior (Gneezy, 2005; Hurkens & Kartik, 2009), the behavioral consequences of externalities in other experimental paradigms are less clear. For example, introducing externalities in die-roll tasks does not seem to affect (dis)honest behavior (Abeler et al., 2016; Fischbacher & Föllmi-Heusi, 2013).

Reward size. The reward for acting dishonestly is a central element of experiments on dishonesty. Somewhat counterintuitively, it has been argued that greater rewards might lead to less dishonest behavior because the psychological costs of cheating increase (Mazar et al., 2008). However, most empirical findings suggest otherwise, showing either that dishonesty is relatively independent of reward size (Abeler et al., 2016; Fischbacher & Föllmi-Heusi, 2013; Hugh-Jones, 2016) or that greater incentives increase dishonesty (Conrads, Irlenbusch, Rilke, Schielke, & Walkowitz, 2014; Gneezy, 2005).

Overall, the empirical findings on personal and situational factors impacting dishonest behavior are mixed and, to some extent, contradictory. A systematic analysis of the factors associated with dishonest behavior thus seems desirable. Moreover, there is a possibility of publication bias—that is, the selective reporting of experiments with significant effects only—for both situational and personal factors. Most experiments have relatively small sample sizes, raising the question of how robust the identified effects are. This chapter aims to address these and related issues systematically and quantitatively using meta-analytical techniques.

Methods

Search

In September 2016, we searched the *Google Scholar* databases (scholar.google.com) for all scientific manuscripts that cited the seminal investigations introducing the four experimental paradigms: Gneezy (2005) for sender–receiver games, Bucciol and

Piovesan (2011) for coin-flip tasks, Fischbacher and Heusi (2008; or the later publication Fischbacher & Föllmi-Heusi, 2013) for die-roll tasks, and Mazar, Amir, and Ariely (2008) for matrix tasks. The search covered all journal articles, book chapters, working papers, discussion papers, and scientific theses in order to minimize potential bias arising from the publication of only significant results. In the following, we refer to each token as an article. Only one-shot, fully anonymous, and incentivized experiments, in which dishonest behavior could not be sanctioned, were included in our analysis. A detailed description of the inclusion criteria is given in Appendix 4.1. Figure 4.1 provides an overview of the selection process and the number of articles identified per step.

For all articles that fulfilled the inclusion criteria, we contacted the authors to obtain the primary data. If authors were unable or unwilling to provide us with the primary data, we inferred the necessary information from the test statistics and/or figures provided. Altogether, our inclusion criteria resulted in a pool of 102 articles (with a total of N = 30,043 observations) covering 470 experiments: 150 sender–receiver games (n = 7,463 observations), 126 coin-flip tasks (n = 8,512), 98 die-roll tasks (n = 8,359), and 96 matrix tasks (n = 5,709). A complete list of the studies included can be found in Appendix 4.1.

Coding

Standardized report. One way of comparing dishonest behavior within and between the experimental paradigms is the standardized report (Abeler et al., 2016), M_r:

$$M_r = \frac{m - t}{t - t_{\min}} \quad \text{if } m < t \text{, and} \tag{4.1}$$

$$M_r = \frac{m - t}{t_{\max} - t} \quad \text{if } m \geq t, \tag{4.2}$$

where m is the actual report per study, t is the expected report if participants were honest, t_{min} is the minimum possible report, and t_{max} is the maximum possible report. The standardized report can range from 100%, indicating that everybody cheated to the maximal degree, over 0%, indicating that participants reported honestly, to –100%, indicating that all participants (oddly) claimed the lowest possible reward.

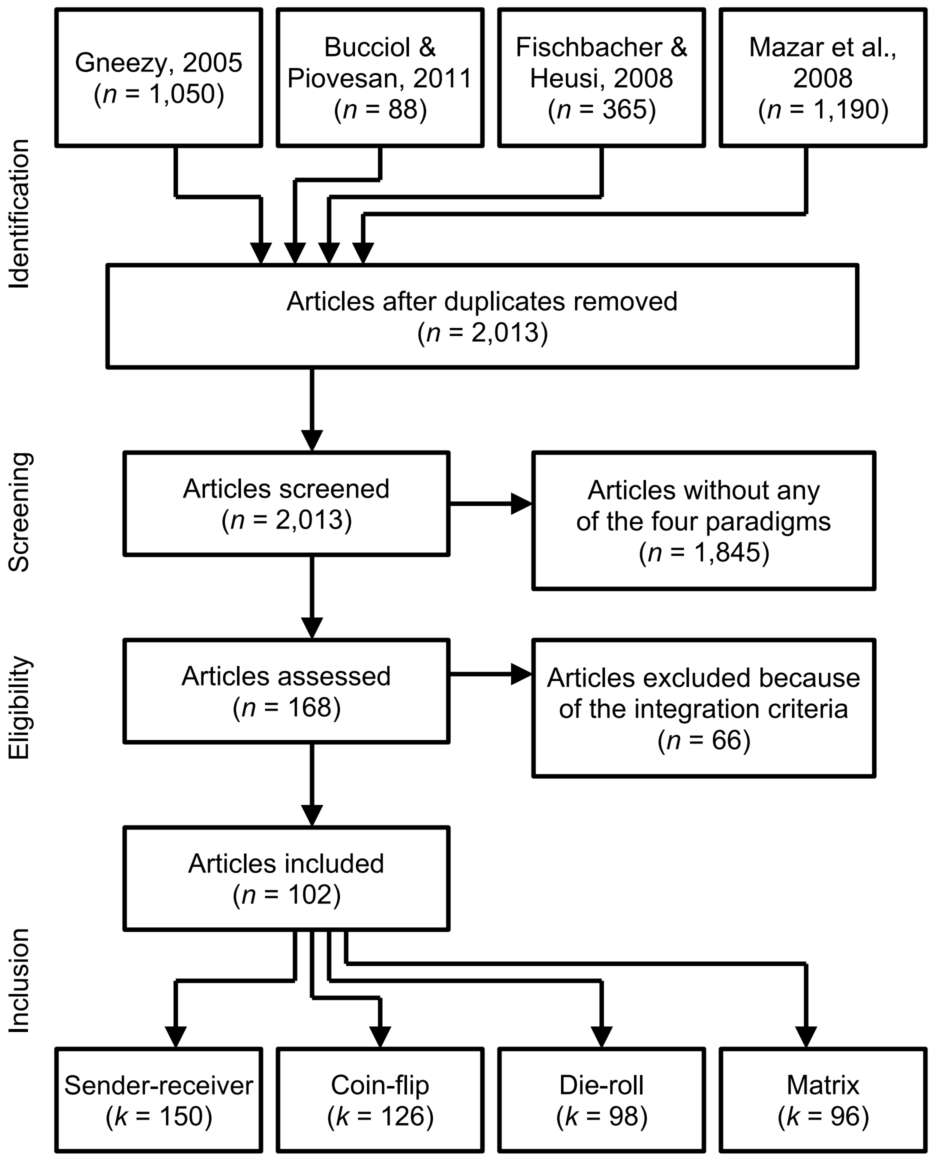


Figure 4.1. *PRISMA flow diagram describing the selection of relevant articles.* All publications until September 2016, with n = number of identified articles; k = number of identified experiments. Each experimental condition are counted as one experiment. In each experiment, actual responses were compared with a hypothetical distribution of honest answers (e.g., a hypothetical distribution of fair die rolls) or, in the case of aggregate-level matrix tasks, with the actual performance level of a control group in which dishonesty was impossible.

For example, in a die-roll task with a six-sided die where each score point (pip) translates into $1, the minimum report is $t_{min} = 1$, the maximum report is $t_{max} = 6$, and honestly reporting would, on average, result in $t = 3.5$. A die-roll task in which all participants report rolling $m = 6$ would thus convert to $M_r = 100\%$, indicating that everybody who could, cheated to the maximal degree; an average report of $m = 3.5$ would equal the expected report of a fair die roll and would thus convert to M_r = 0%, indicating that participants reported honestly (m = t); if all participants reported rolling the lowest possible score, m = 1, the standardized report would be M_r = −100%. Negative values of the standardized report indicate that participants claimed less than would be expected from honest reporting of a representative outcome distribution. For example, if the winning side in a coin-flip task with a single coin was reported in "only" m = 44% of tosses (expected: t = 50%), then the standardized report would be M_r = −12%; that is, (44% – 50%) / (50% – 0%).[7]

The standardized report thus quantifies the percentage of people who behaved dishonestly (the rate of liars) and the level of their dishonest behavior in a single measure. We focus on this measure in our analyses, as it allows us to conduct comparisons within and across the four paradigms without limiting the dataset (see Appendix 4.2 for details). Note, however, that sender–receiver games and coin-flip tasks do not allow for differentiation in the degree of dishonest behavior. Here, the decision to behave dishonestly is an all-or-nothing one (dichotomous measurement scale). Hence, for sender–receiver games and most[8] coin-flip tasks, the standardized report directly corresponds to the percentage of people acting dishonestly. For example, a standardized report of M_r = 23% in sender–receiver games means that 23% of senders chose to convey the false message. Die-roll tasks and matrix tasks, in contrast, have a continuous scale. Here, the standardized reports do not distinguish between the percentage of people who behaved dishonestly (the rate of liars) and the degree of their dishonest behavior. For example, in a matrix task a standardized report with M_r = 30% indicates that 30% of unsolved matrices were claimed as solved. This could result from a few people misreporting to a high degree (low rate of liars with high level of dishonesty) or from many people misreporting to a lower degree (high rate of liars with low level of

[7] It can be argued that negative standardized reports do not reflect dishonest behavior because there is no apparent reason to claim less than one truthfully observed. To avoid distortion in the distribution of the standardized report, we allowed for negative standardized reports rather than excluding these values or making them zero. This is because negative and positive standardized reports could be, to some degree, the result of random sampling errors. For example, in aggregate-level matrix tasks, the control group may solve more matrices than the experimental group. Consequently, the standardized report takes a negative value although the experimental group did not cheat. In the opposite case—i.e., when the control group solves fewer matrices than the experimental group—the standardized report takes a positive value although the experimental group did not cheat. If the standardized report is allowed to take both positive and negative values, such random fluctuations cancel each other out with sufficient observations.

[8] Some coin-flip tasks used randomly generated outcomes with a chance of winning other than 50:50. The standardized report and the rate of liars can therefore diverge for some coin-flip tasks (see Appendix 4.2).

dishonesty; see Appendix 4.3 for more detailed analyses in tasks with continuous outcome measures). To disentangle the percentage of dishonest people from the degree of dishonesty, we examined an additional measure of dishonest behavior.

Rates of liars. The rate of liars, M_{liars}, indicates the percentage of participants who acted dishonestly, irrespective of the degree of their dishonest behavior. In the two paradigms with dichotomous scales (sender–receiver and coin-flip tasks), the rate of liars is practically equivalent to the standardized report. In the two paradigms with continuous scales (die-roll and matrix tasks), the rate of liars indicates the proportion of participants who claimed more than they were eligible for—regardless of how much more they claimed. To estimate the rate of liars in coin-flip tasks and in die-roll tasks, we used only a fraction of all observations. For die-roll tasks, the rate of liars was calculated using only reports of the lowest possible outcome. We assumed that the lowest possible outcome was reported only by participants who truly observed it. Because participants who observed the lowest outcome were maximally tempted to lie the estimated rate of liars is expected to be an upper bound to the "true" rate of liars in die-roll tasks (see Appendix 4.2 for further details). For matrix tasks, the rate of liars was calculated only for experiments that measured individual-level behavior. For both calculations, we needed access to the primary data. These methodological constraints reduced the number of observations and the number of experiments eligible for calculating the rate of liars (see Appendix 4.2 for detailed calculations). In the main analyses, we therefore focus on the standardized report, which uses all observations.

To combine the measures of dishonest behavior, we used random effects models that account for variations in the effect size distributions between experiments. To quantify this variation, we provide the I^2 statistic, which is the between-study variance independent of the number of experiments, and the τ^2 statistic, which is the estimated variance of underlying effects across studies (Huedo-Medina et al., 2006). To identify and counteract the risk of selective reporting of only significant results (publication bias), we used the trim and fill method. The iterative algorithm adds hypothetical experiments to the analysis until the observations are symmetrically distributed around the average effect size (Duval & Tweedie, 2000). For all analyses, we used the statistical software *R* (R Development Core Team, 2008) and the default sensitivity parameters of the packages *meta* (Schwarzer, 2007) and *lme4* (Bates, Mächler, Bolker, & Walker, 2015). The database is available online at https://osf.io/d9jzv/.

Results

We begin by investigating whether and how our measures varied across the four experimental paradigms. We then present our findings on personal and situational factors impacting (dis)honest behavior. Finally, we examine the interactions of the experimental paradigms with the personal and situational factors by means of regression analysis.

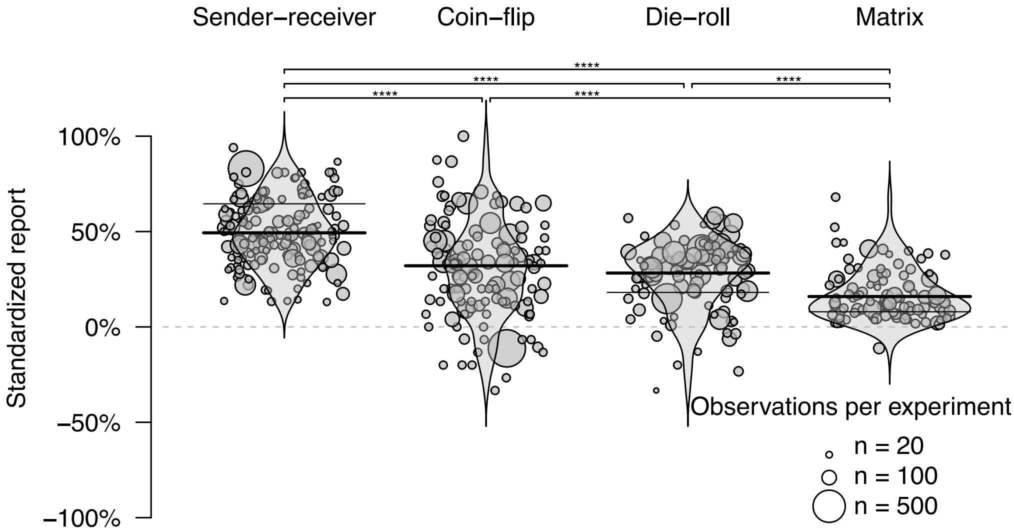
Variation Across the Experimental Paradigms

Standardized report. Standardized reports varied strongly between and within the four experimental paradigms. As depicted in Figure 4.2, the standardized report ranged from M_r = 100%, indicating that all participants cheated to the maximal degree, to M_r = –33%, indicating that participants claimed about 1/3 less than entirely honest participants (who observed a representative distribution of outcomes) would have done. Most experiments, however, fell somewhere between these extremes. In sender–receiver games, an average of M_r = 49% of messages were false. In coin-flip tasks, wins were reported M_r = 32% more often than would be expected from honest reporting. This value translates into an average of 66% reported wins when 50% of participants actually tossed the winning side. [9] In die-roll tasks, the reported outcomes averaged to M_r = 28%. For a six-sided die in which each reported score point (pip) paid \$1, this value converts to a mean claim of \$4.21 where \$3.50 would be expected from honest reporting.[10] Matrix tasks yielded the smallest average standardized report of all the paradigms, indicating that M_r = 16% of all unsolved matrices were reported as solved. Overall, the standardized reports for coin-flip tasks and die-roll tasks were relatively similar, $Q(1) = 1.45$, $p = .228$. Relative to those two experimental paradigms, sender–receiver games yielded higher standardized reports,

[9] For coin-flip tasks, the standardized report refers to the estimated percentage of falsely claimed wins. After a single coin toss, an honest sample would report about 50% wins, in which case M_r = 0. A standardized report of M_r = 32% thus translates for one-shot coin-flip tasks with a 50% honest chance of winning to 66% reported wins = 50% [estimated actually observed wins] + (100% – 50%) [estimated actually observed losses] × 32% [standardized report].

[10] For die-roll tasks, the standardized report indicates the percentage of claimed score points (pips) that were over and above the expected average claim that would result from honest reporting. For a six-sided die in which each reported pip paid \$1, an honest sample would claim, on average, \$3.50 = (\$1 + \$2 + \$3 + \$4 + \$5 + \$6)/6, in which case M_r = 0. The standardized report of M_r = 28% thus translates to a mean claim of \$4.21 = \$3.50 [estimated average eligible claim] + (6 – \$3.50) [estimated average maximal-possible over-claim] × 28% [standardized report].

$Q(1) = 67.25$, $p < .001$, whereas matrix tasks yielded lower standardized reports, $Q(1) = 42.65$, $p < .001$ (both comparisons combined coin-flip tasks and die-roll tasks).

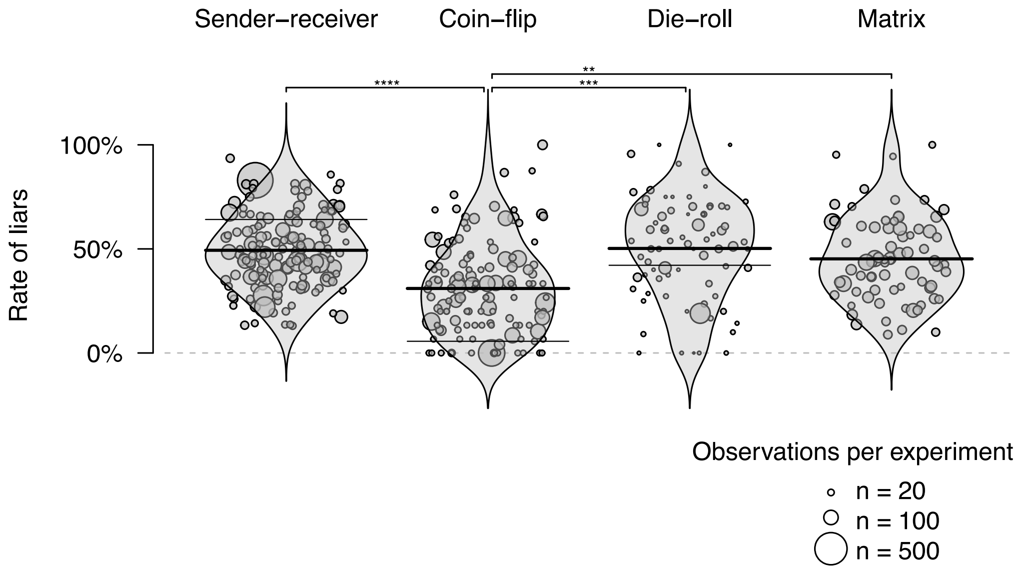


	Sender–receiver	Coin–flip	Die–roll	Matrix
Random effect [95% CI]	M_r = 49% [46%, 53%] z = 28.05 p < .001	M_r = 32% [27%, 37%] z = 13.25 p < .001	M_r = 28% [24%, 32%] z = 13.93 p < .001	M_r = 16% [13%, 19%] z = 10.56 p < .001
Observations	k = 150 n = 7463	k = 126 n = 8512	k = 98 n = 8365	k = 96 n = 4517
Heterogenity	I^2 = 90% τ^2 = 0.20	I^2 = 71% τ^2 = 0.20	I^2 = 81% τ^2 = 0.17	I^2 = 82% τ^2 = 0.12
Trim and fill [95% CI]	M_r = 65% [61%, 68%] z = 32.03 p < .001	M_r = 32% [27%, 37%] z = 13.25 p < .001	M_r = 18% [14%, 22%] z = 8.57 p < .001	M_r = 8% [5%, 11%] z = 4.67 p < .001
Added studies	k_{add} = +62	k_{add} = +0	k_{add} = +31	k_{add} = +35

Figure 4.2. *Violin plots showing the distribution of standardized reports by experimental paradigm.* Each dot represents an experiment. The dot size indicates the number of observations. All dots are horizontally jittered for the sake of visualization. The thick, solid bars indicate the estimated mean of the standardized report. The thin, solid lines indicate the estimated mean of the standardized report after trim and fill adjustments (publication bias correction). The bars above the plots summarize the results of tests for subgroup differences, with **** $p < .0001$. The table below the figure presents the summary statistics of the random effects models: M_r is the standardized report [with 95% confidence interval]; k is the number of experiments; n is the total number of observations; I^2 is the study variance independent of the number of experiments; and τ^2 is the between-study variance.

Heterogeneity (in terms of I^2 and τ^2) in all experimental paradigms was large, and the standardized report did not converge to a specific value. One reason for this could be publications bias: imprecise experiments yielded particularly extreme values of the standardized report and thus biased the grant mean of all estimates. Trim and fill adjustments indeed found evidence of publication bias in three of the four paradigms—although the adjustments did not all point in the same direction. For sender–receiver games, trim and fill increased the standardized report to M_r = 65% false messages (as reported in Figure 4.2). For coin-flip tasks, no adjustment was suggested. For die-roll tasks, trim and fill decreased the standardized report to M_r = 18%. For matrix tasks, it halved the standardized report to M_r = 8%. Given these substantial indications of publication bias, our subsequent analyses need to be interpreted with care: "True" rates of standardized reports are potentially higher in sender–receiver games and lower in die-roll and matrix tasks than suggested by the literature synthesized.

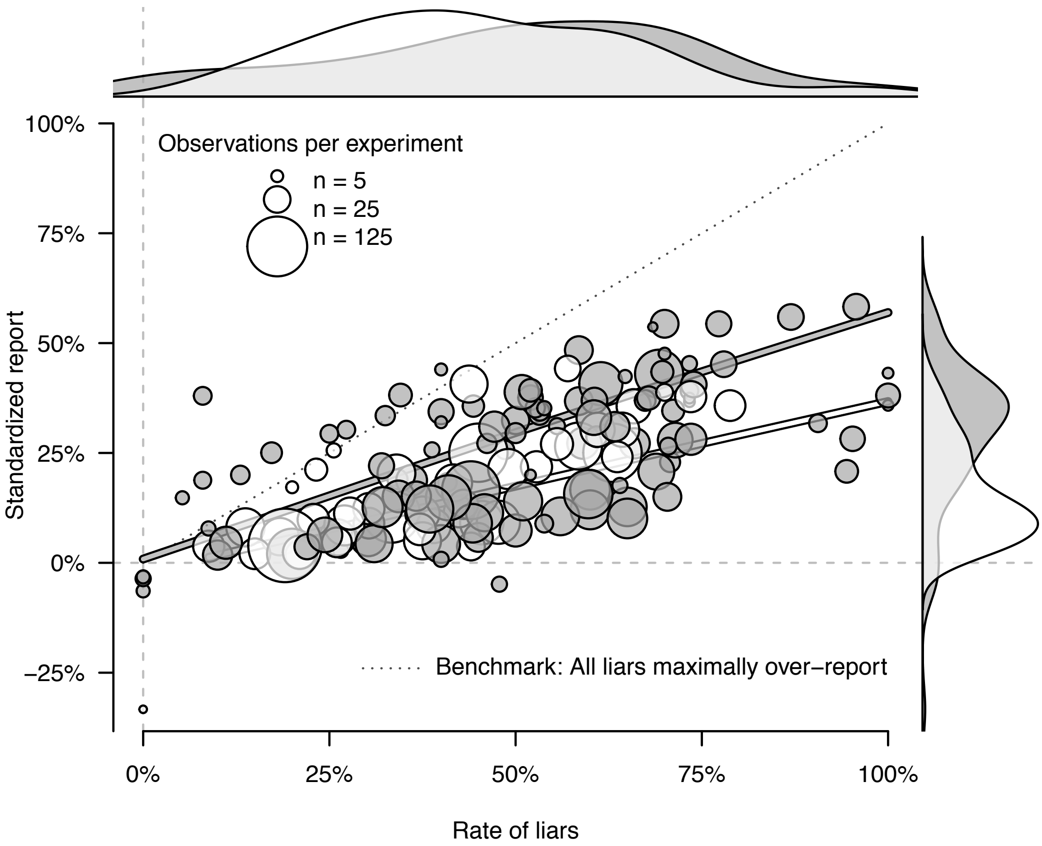
Rate of liars. We now turn to the percentage of people who behaved dishonestly, irrespective of the degree of cheating. These analyses are based on a subset of experiments and of observations. On average, sender–receiver games, die-roll tasks, and matrix tasks yielded relatively similar estimates of the liar rate, suggesting that 49%, 50%, and 45%, respectively, of participants lied (all pairwise comparisons: $Q[1] < 1.75$, $p > .185$; se Figure 4.3). By contrast, "only" about 31% of participants lied in coin-flip tasks (all pairwise comparisons: $Q[1] > 7.58$, $p < .005$). Overall, heterogeneity was large (see I^2 and τ^2 in Figure 4.3). Trim and fill analyses suggested substantial publication bias in all paradigms except matrix tasks: the rates of liars in coin-flip tasks, in particular, were substantially lowered, further increasing the gap in the liar rates between coin-flip tasks and the other paradigms.



	Sender-receiver	Coin-flip	Die-roll	Matrix
Random effect [95% CI]	M_{liars} = 49% [46%, 53%] z = 28.42 p < .001	M_{liars} = 31% [22%, 40%] z = 6.94 p < .001	M_{liars} = 50% [44%, 56%] z = 16.05 p < .001	M_{liars} = 45% [40%, 50%] z = 17.49 p < .001
Observations	k = 150 n = 7463	k = 126 n = 4869	k = 77 n = 1428	k = 73 n = 3377
Heterogenity	I^2 = 90% τ^2 = 0.20	I^2 = 99% τ^2 = 0.49	I^2 = 83% τ^2 = 0.24	I^2 = 91% τ^2 = 0.21
Trim and fill [95% CI]	M_{liars} = 64% [60%, 68%] z = 32.37 p < .001	M_{liars} = 6% [1%, 11%] z = 2.15 p = .031	M_{liars} = 42% [36%, 49%] z = 12.75 p < .001	M_{liars} = 45% [40%, 50%] z = 17.49 p < .001
Added studies	k_{add} = +61	k_{add} = +62	k_{add} = +15	k_{add} = +0

Figure 4.3. *Violin plots showing rate of liars by experimental paradigm.* The bars above the plots summarize the results of tests for subgroup differences, with ** $p < .01$, *** $p < .001$, **** $p < .0001$. See caption to Figure 4.2 for further explanation.

The rate of liars and its relation to the standardized report. Die-roll tasks and matrix tasks both have continuous outcome measures. Comparing the rate of liars with the standardized report in these two paradigms revealed a striking difference: Although a comparable rate of liars emerged for die-roll tasks (M_{liars} = 50%) and matrix tasks (M_{liars} = 45%; Q[1] = 1.44, p = .231, the standardized report in die-roll tasks (M_r = 28%) was almost twice that in matrix tasks (M_r = 16%; Q[1] = 23.74, p < .001). Taking additionally into account that M_{liars} in die-roll task is the upper bound (Appendix 4.2), this finding suggests that liars in die-roll tasks cheated to a substantially greater degree than liars in matrix tasks did (Figure 4.4).



	Die−roll	Matrix
Study Best fit		
Observations	k = 77 n = 1428	k = 73 n = 3377
Rate of liars Standardized report	M_{liars} = 49% M_r = 29%	M_{liars} = 45% M_r = 15%
Correlation	$\rho = 0.69^{***}$	$\rho = 0.76^{***}$

Figure 4.4. *Relationship of the rate of liars and the standardized report in experiments with continuous outcome measures.* The plot in the center shows the standardized report (y-axis) as a function of the rate of liars (x-axis). Each dot represents a single experiment. Larger dots represent experiments with more observations. The best-fitting regression lines summarize the relationship of the standardized report and the rate of liars in the die-roll task (gray) and the matrix task (white). The dotted line depicts the (hypothetical) relationship in the case that all liars claim the maximal payoff. Some die-roll tasks fell above this benchmark due to sampling errors in estimating the rate of liars (see Appendix 4.2 for details). The density plot on the right shows the distribution of the standardized report in die-roll (gray) and matrix tasks (white). The density plot on the top shows the distribution of the rate of liars in die-roll (gray) and matrix tasks (white). The table below the figure presents the model summary statistics. The Spearman rank correlation coefficient ρ indicates the relationship of the standardized report and the rate of liars on the level of the experiments, with *** $p < .001$.

There are several possible reasons for this difference. In die-roll tasks, even the most advantageous outcome is based on chance and is as (un)likely as the least advantageous one. In skill-based tasks like the matrix task, in contrast, top performances and, by extension, top outcomes are potentially much less likely and hence plausible than, say, medium performances. Consequently, the risk participants claiming large rewards being met with disbelief may seem smaller in matrix tasks. Relatedly, the social norms and/or psychological costs governing honesty and trickery may differ in domains of chance versus skill (e.g., shame or guilt). To test the possibility of different (objective) base rates of maximum outcomes, we calculated the chance of obtaining the highest reward through honest means for each experiment—that is, the probability of observing the highest score in die-roll tasks and the probability of solving all matrices in matrix tasks. Overall, 15 times more people were eligible for the highest reward in die-roll tasks than in matrix tasks ($W = 114$, $p < .001$; one-sided Wilcoxon test). In 83% of the matrix tasks, not a single participant solved all matrices (see Appendix 4.4 for further tests).

Personal and Situational Factors

Before assessing how personal factors are associated with dishonest behavior, let us inspect the participant composition of the data set. Figure 4.5 shows the total number of observations per country. By far most experiments were conducted in the United States and Germany, followed primarily by other WEIRD countries (WEIRD = Western, Educated, Industrialized, Rich, and Democratic; cf. Henrich, Heine, & Norenzayan, 2010). As shown in Figure 4.6, there were almost as many observations for women (49%) as for men (51%) and most participants were in their twenties ($Mdn_{age} = 22$ years, $SD_{age} = 11$). About 81% of the experiments were conducted in samples of university and college students. As a consequence, students accounted for 75% of all observations. Of the students, 41% were majoring in business and/or economics and only 3% in psychology. With the exception of gender, the overall participant composition was thus far from representative. Consequently, demographic analyses have to be interpreted with care.

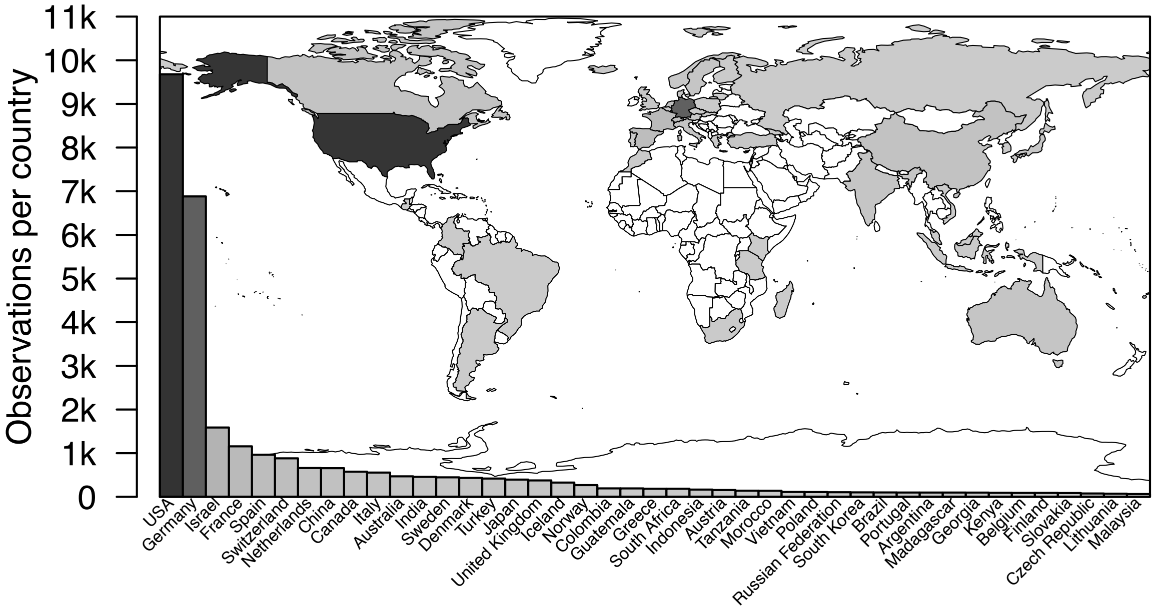


Figure 4.5. *Observations by country.* Most experiments were conducted in the United States, Germany, and other WEIRD countries. The shades of the bar plot represents the total number of observations. The shades are reflected in the world map and are not relative to the population sizes of the countries. Countries in white provided no observations.

Gender. We next compared the reporting behavior of men and women in the 331 experiments in which gender was elicited. Overall, men's standardized reports were 3% higher than women's, suggesting that men behaved slightly more dishonestly than women did (n = 22,956, $M_{male\ r} - M_{female\ r}$ = 34% – 31% = 3%, *95% CI* [2%; 5%], $z = 4.28$, $p < .001$). Differentiating by experimental paradigm, we found that men only behaved more dishonestly than women in sender–receiver games and coin-flip tasks (Figure 4.7). There were no gender differences in die-roll tasks. The gender effect in matrix tasks approached conventional levels of significance but was confounded by performance. On average, men solved 1.17 matrices more than women did. Controlling for performance reduced the gender differences in matrix tasks to insignificance (Appendix 4.3). Notably, however, low performers in matrix tasks had higher standardized reports and lied more than high performers did (Appendix 4.3). In total, 42% of men and 39% of women lied. This difference was statistically significant (k = 316, n = 12,696, $M_{male\ liars} - M_{female\ liars}$ = 42% – 39% = 3%, *95% CI* [1%; 5%], $z = 3.31$, $p < .001$).

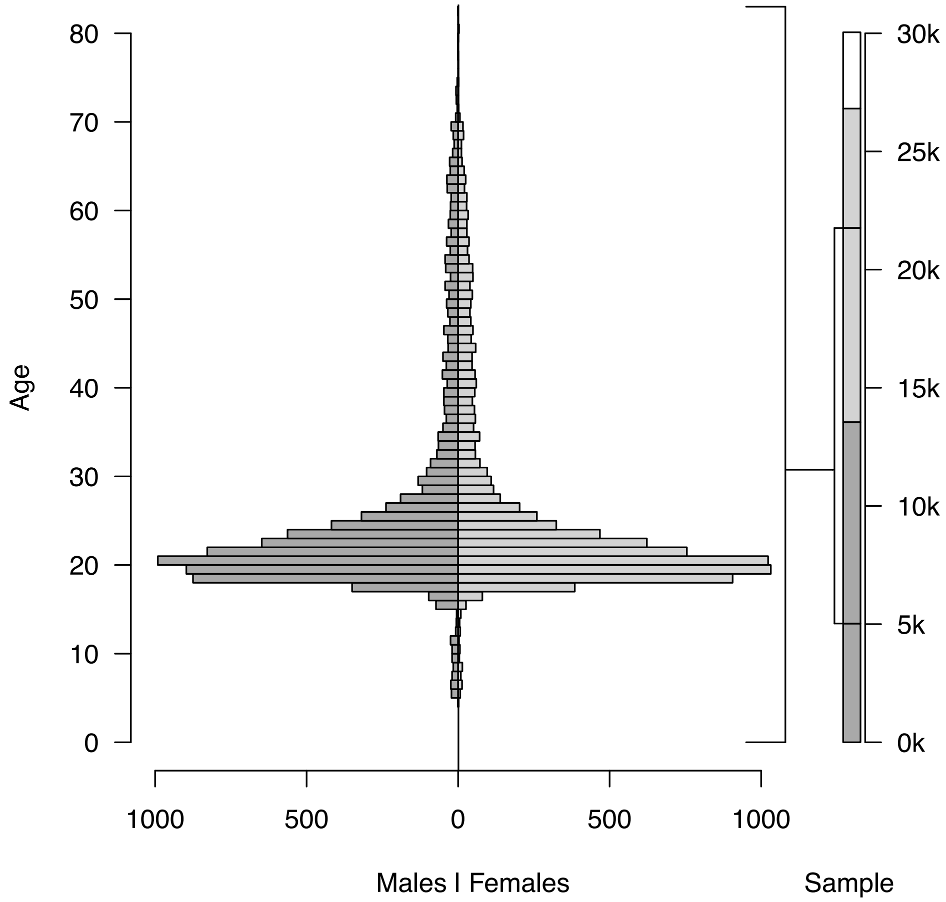
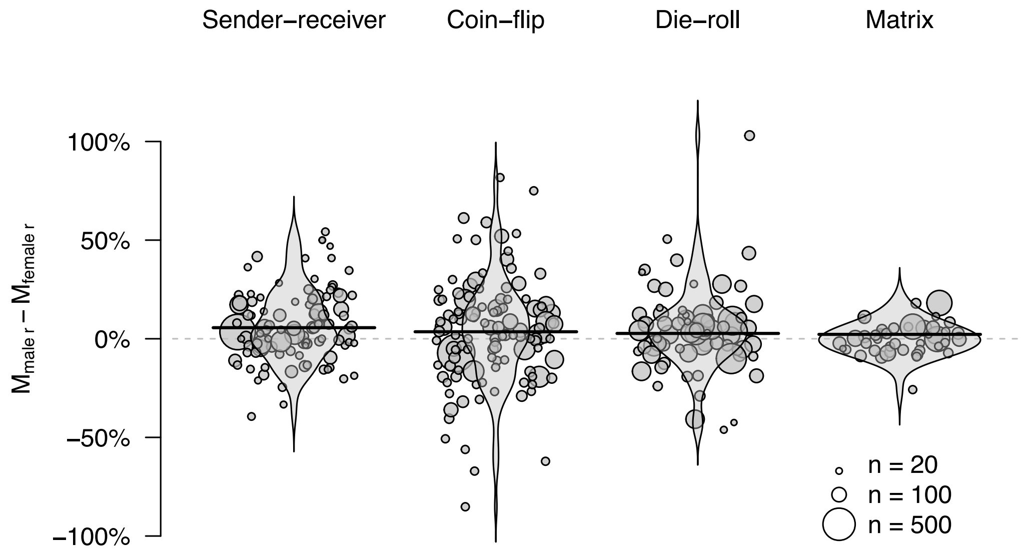


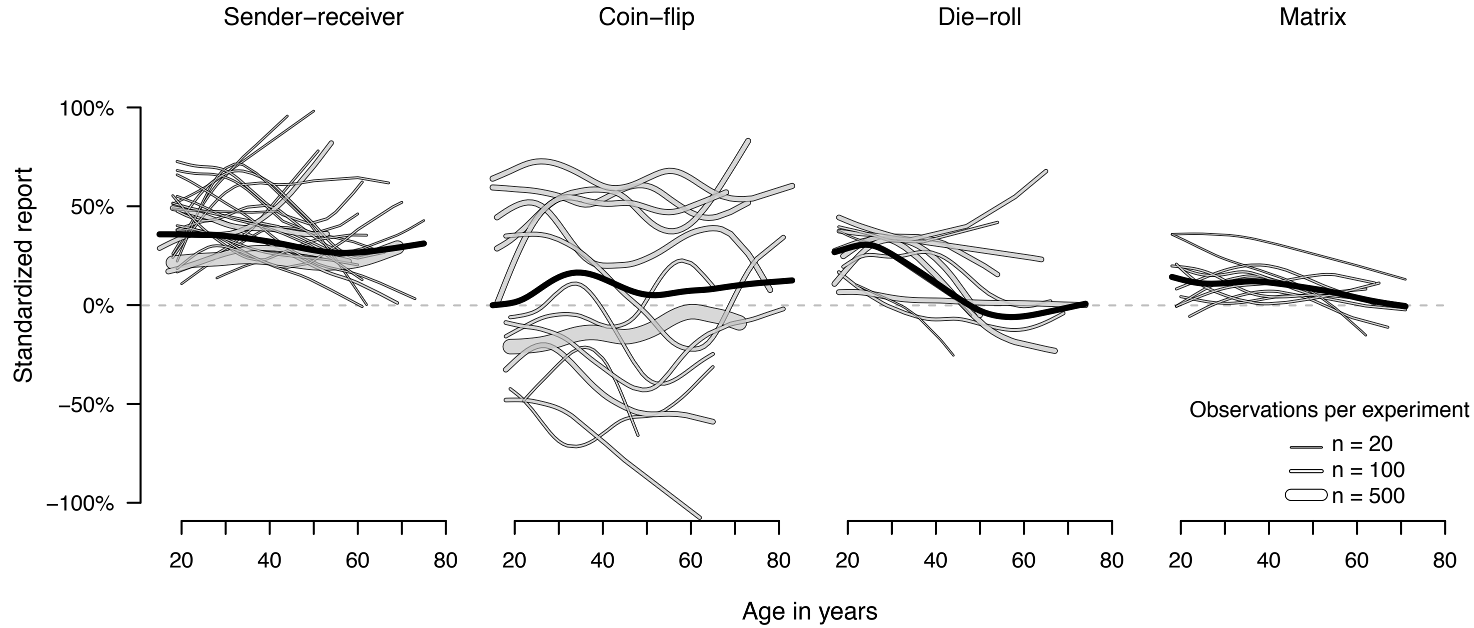
Figure 4.6. *Observations by gender and age*. There were about as many observations for men as for women; most participants were in their twenties. The demographic pyramid on the left shows the age distribution for the subsample of participants whose age and gender were known. The stacked bar plot on the right shows the total number of observations. The number of male participants is represented in dark gray. The number of female participants is shown in light gray. The number of participants whose gender was unknown is shown in white.

Age. In order to test whether age was related to dishonest behavior, we fitted a series of linear mixed effects models to the results of the 59 experiments in which age varied to at least some degree ($SD_{age} > 5$). Overall, we found a small negative effect of age on several measures of dishonest behavior, suggesting that younger participants behaved more dishonestly than older participants. Every year of life lowered the standardized report by 0.17 percentage points ($k = 59$, $n = 6{,}384$; $b_o = 28.67\%$ [$SE_o = 4.32$], $t = 6.64$, $p < .001$; $b_{age} = -0.17\%$ [$SE_{\mathrm{age}} = 0.07$], $t = -2.56$, $p = .002$; linear mixed model with random intercepts between experiments). The age effect in the standardized report was largely traceable to die-roll tasks; no other paradigm yielded an age effect (Figure 4.8). In matrix tasks, analyses that controlled for gender and the number of solved matrices revealed a small age effect, with younger participants exhibiting more dishonest behavior (Appendix 4.3). The rate of liars also decreased with age. Every additional year lowered the probability of lying by 0.24 percentage points (k = 36, n = 2,690; b_o = 52.09% [SE_o = 3.86], t = 13.49, p < .001; b_{age} = −0.24% [SE_{age} = 0.08], t = −3.09, p = .001; linear mixed model with random intercepts between experiments; limited to experiments with individual-level data).



	Sender−receiver	Coin−flip	Die−roll	Matrix
$M_{male\ r} - M_{female\ r}$ [95% CI]	$M_{r\ diff}$ = 6% [2%, 9%] z = 3.51 p < .001	$M_{r\ diff}$ = 4% [0%, 7%] z = 2.26 p = .024	$M_{r\ diff}$ = 3% [−1%, 6%] z = 1.45 p = .147	$M_{r\ diff}$ = 2% [0%, 5%] z = 1.90 p = .057
Observations	k = 98 n = 5230	k = 116 n = 7401	k = 72 n = 7233	k = 45 n = 3110
Heterogenity	I^2 = 9% τ^2 = 0.06	I^2 = 4% τ^2 = 0.03	I^2 = 9% τ^2 = 0.09	I^2 = 4% τ^2 = 0.02

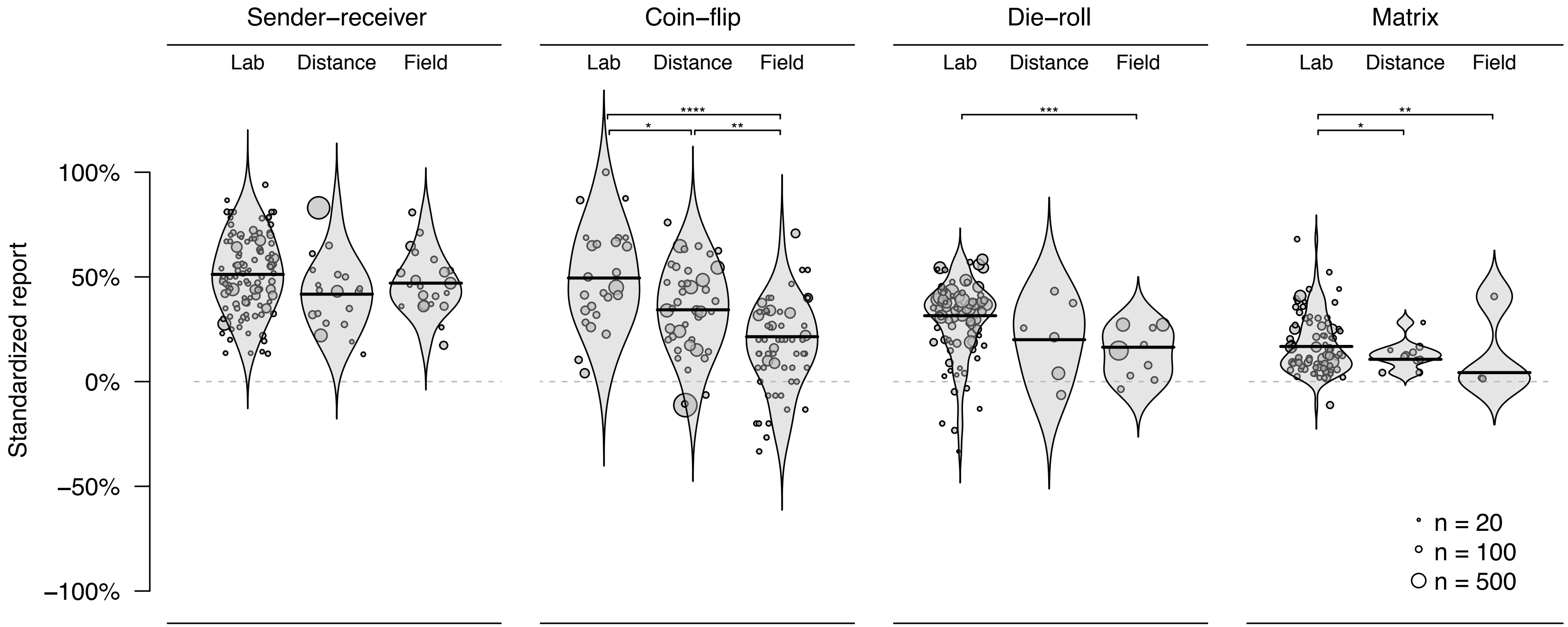
Figure 4.7. *Violin plots showing gender differences in the standardized report across the four paradigms.* See caption to Figure 4.2 for further explanation.



Intercept	b_0 = 0.45**** (0.04)	b_0 = 0.04 (0.12)	b_0 = 0.39**** (0.06)	b_0 = 0.18**** (0.04)
Age	b_{age} = −0.17% (0.10)	b_{age} = −0.04% (0.13)	b_{age}= −0.55%*** (0.15)	b_{age} = −0.19% (0.10)
Observations	k = 24 n = 1910	k = 14 n = 2589	k = 11 n = 1363	k = 10 n = 462

Figure 4.8. *Age effects in standardized reports across the four paradigms*. All analyses were limited to experiments with participants whose age varied by at least 5 standard deviations. The graphs depict the standardized report per experiment (y-axis) as a function of participants' age (x-axis). The lines were smoothed by local polynomial fitting and weighted by the number of observations. The line width indicates the total number of observations per experiment. The black lines are the estimated means of the standardized report per paradigm. The table below the figures reports the summary statistics of mixed linear regression models with random intercepts between experiments with standard errors in parentheses, and with *** $p < .001$ and **** $p < .0001$.

Investigative setting. To compare the effect of the investigative setting on dishonest behavior, we grouped all studies into three categories—laboratory experiments, online/telephone experiments, and field experiments—and compared standardized reports across these categories. Overall, we found no differences in standardized reports between laboratory experiments (lab: k = 300, n = 17,978, M_r = 36%) and telephone/online experiments (distance: k = 77, n = 6,549, M_r = 32%; lab vs. distance: $Q[1] = 0.87$, $p = .350$), or between telephone/online experiments and field experiments (field: k = 93, n = 5,516, M_r = 27%; distance vs. field: $Q[1] = 0.96$, $p = .326$). However, standardized reports in laboratory experiments were systematically higher than in field experiments (lab vs. field: $Q[1] = 6.89$, $p = .008$). Similar results were observed for the rate of liars (lab: k = 258, n = 9,311, M_{liars} = 49%; distance: k = 77, n = 4,512, M_{liars} = 40%; field: k = 91, n = 3,314, M_{liars} = 29%; lab vs. distance: $Q[1] = 1.31$, $p = .252$; distance vs. field: $Q[1] = 1.87$, $p = .171$; lab vs. field: $Q[1] = 51.57$, $p < .001$). It should be noted that simply comparing investigative settings is problematic because of the unequal distribution of experimental tasks per investigative setting. For example, coin-flip tasks were more frequently conducted in the field than sender–receiver games were. Coin-flip tasks thus contributed relatively more observations to the averaged standardized report of field experiments than sender–receiver games did. A more fine-grained analysis of investigative setting is summarized in Figure 4.9, which separates the average standardized reports by experimental paradigm. The results suggest that the standardized reports emerging from sender–receiver games did not differ between any of the investigative settings. By contrast, the standardized reports emerging from coin-flip tasks differed between all three investigative settings. In both die-roll tasks and matrix tasks, laboratory studies yielded greater standardized reports than field studies. To reduce the potential effect of confounds on investigative setting, we next ran regression analyses that also considered the effect of investigative setting controlling for other covariates.



Standardized report [+/– 95% CI]	M_r	51% [4%]	42% [14%]	47% [7%]	49% [12%]	34% [7%]	21% [5%]	31% [3%]	20% [16%]	16% [7%]	17% [3%]	11% [4%]	4% [7%]
	z	26.91	6.05	13.94	8.04	9.41	8.45	18.65	2.44	4.50	9.87	5.90	1.23
	p	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.015	<.001	<.001	<.001	.218
Observations	k	109	19	22	25	42	59	83	6	9	83	10	3
	n	8736	3194	2996	3832	7878	5314	13356	1102	2272	7848	924	262
Heterogenity	I^2	86%	97%	86%	84%	70%	16%	58%	90%	37%	79%	14%	80%
	τ^2	0.18	0.29	0.14	0.27	0.17	0.07	0.11	0.18	0.06	0.13	0.02	0.05

Figure 4.9. *Violin plots showing the distribution of standardized reports by experimental paradigm and investigative setting.* The bars above the plots summarize the results of tests for subgroup differences, with * $p < .05$, ** $p < .01$, *** $p < .001$, and **** $p < .0001$. See caption to Figure 4.2 for further explanation.

Regression analyses

Regression analyses allowed us to control for potential cofounds in the analysis of investigative setting and to take all experiments into account, without restricting the data set to studies that provided sufficient variation within each experiment (e.g., experiments whose participants varied in age or that assessed gender). Table 4.2 summarizes two regression models that predict the measures of dishonest behavior (M_r, M_{liars}) by lumping together the experimental paradigms and adding dummy variables for the experimental paradigms. Table 4.3 presents additional regression models that predict the standardized reports in each of the four experimental paradigms.

Online/telephone experiments. Overall, there was as much (dis)honest behavior in lab experiments as in experiments conducted online or via telephone. The exception were coin-flip tasks, in which the standardized reports were 25 percentage points lower in online/telephone experiments than in lab studies (Table 4.3).

Field experiments. There was more dishonesty in lab experiments than in field experiments. The standardized reports were 15 percentage points lower in field settings and the rate of liars was 17 percentage points lower (Table 4.2). This effect was mostly due to coin-flip tasks; here, standardized reports were 57 percentage points lower in field experiments than in lab studies (Table 4.3).

Nonstudents. The results of experiments with nonstudent groups were largely similar to those of studies with students who majored in disciplines other than economics (Table 4.2). The only exception were die-roll tasks, in which the standardized reports for nonstudent samples were 14 percentage points higher than those for non-economic student samples (Table 4.3), suggesting that students who major in a discipline other than economics behaved more honestly than nonstudent population groups.

Economics students. Within the student population, we tested whether majoring in economics was associated with more or less dishonest behavior. Overall, in experiments with 100% economics majors, the liar rate was 10 percentage points higher than among students of other majors (Table 4.2). However, the effect of student major was by no means stable across the four paradigms. In fact, the only statistically significant result to emerge when separating by experimental paradigm was that economics students' standardized reports in coin-flip tasks were 41 percentage points lower than those of students with other majors (Table 4.3).

Table 4.2.

Predictors of Different Measures of Dishonest Behavior Across the Four Paradigms

Dependent variable (reference category)	M_r	M_{liars}
Intercept	8.04 (9.63)	11.10 (14.08)
Experimental paradigm (coin-flip task)		
Sender–receiver	10.22%*** (2.96)	10.14%** (3.59)
Die-roll	−10.33%** (3.21)	13.15%** (4.14)
Matrix	−24.61%**** (3.35)	7.54% (4.47)
Investigative setting (laboratory)		
Online/telephone	−3.33% (3.91)	−2.71% (5.26)
Field experiment	−15.17%**** (2.87)	−17.39%**** (3.44)
Participant sample (non-economics students)		
Nonstudents	3.29% (3.48)	0.31% (4.71)
Economics students	5.77% (2.97)	10.12%** (3.83)
Primary data shared	−3.44% (2.57)	−0.60% (3.78)
Year of publication	−0.37% (0.48)	−0.53% (0.70)
Maximal externality	−0.05% (0.10)	0.03% (0.28)
Maximal gain	−0.08% (0.07)	−0.39% (0.27)
Observations	$k = 470$, $n = 30{,}043$	$k = 426$, $n = 17{,}061$
Residual heterogeneity	$I^2 = 82\%$, $\tau^2 = 0.03$	$I^2 = 91\%$, $\tau^2 = 0.05$
Heterogeneity accounted for	$R^2 = 57\%$	$R^2 = 54\%$

Note. Linear regression models with random effects at the experiment level. Unless denoted otherwise, values refer to beta weights with standard errors in parentheses, with * $p < .05$, ** $p < .01$, *** $p < .001$, and **** $p < .0001$.

Table 4.3.

Predictors of the Standardized Report by Experimental Paradigm

Paradigm (reference category)	Sender–receiver	Coin-flip	Die-roll	Matrix
Intercept	10.85 (16.64)	198.98**** (42.65)	–8.86 (22.25)	–15.66 (14.31)
Investigative setting (laboratory)				
Online/telephone	–2.50% (7.36)	–25.22%** (7.72)	–5.63% (6.58)	–4.84% (13.82)
Field experiment	–7.32% (5.60)	–56.56%**** (6.47)	–7.51% (6.58)	–9.47% (7.67)
Participant characteristics (non-economics students)				
Nonstudents	5.31% (6.72)	–4.50% (6.54)	13.96%* (5.81)	–3.21% (13.10)
Economics students	1.88% (4.74)	–40.69%** (12.71)	–8.75% (6.06)	9.58% (7.47)
Primary data shared	0.24% (4.93)	49.51%**** (8.67)	1.04% (5.99)	–10.16%** (3.78)
Year of publication	–0.52% (0.83)	–9.86%**** (2.12)	0.45% (1.11)	0.79% (0.71)
Maximal externality	–1.39%* (0.57)	2.80%**** (0.59)	–0.06% (0.29)	0.08% (0.25)
Maximal gain	1.19%* (0.60)	–3.79%**** (0.60)	–0.02% (0.30)	–0.02% (0.06)
Observations	$k = 150$ $n = 7{,}463$	$k = 126$ $n = 8{,}406$	$k = 98$ $n = 8{,}359$	$k = 96$ $n = 5{,}709$
Residual heterogeneity	$I^2 = 88\%$ $\tau^2 = 0.03$	$I^2 = 36\%$ $\tau^2 = 0.02$	$I^2 = 60\%$ $\tau^2 = 0.01$	$I^2 = 76\%$ $\tau^2 = 0.01$
Heterogeneity accounted for	$R^2 = 14\%$	$R^2 = 71\%$	$R^2 = 56\%$	$R^2 = 20\%$

Note. Linear regression models with random effects at the experiment level. Unless denoted otherwise, values refer to beta weights with standard errors in parentheses, with * $p < .05$, ** $p < .01$, *** $p < .001$, and **** $p < .0001$.

Primary data shared. We wrote three emails to the corresponding author(s) of each article included in the analysis requesting their primary data. As detailed in Appendix 4.1, not all experimenters provided us with their primary data. We tested whether sharing of primary data was associated with systematic differences in the measures of dishonest behavior. In coin-flip tasks, standardized reports were 50 percentage points higher in experiments with shared data than in experiments without (Table 4.3). In matrix tasks, standardized reports were 10 percentage points lower in experiments with shared data than in experiments without (Table 4.3).

Year of publication. In coin-flip tasks, the standardized report decreased over time. Each year of publication was associated with a 10 percentage points decrease in the standardized report (Table 4.3). The meta-analysis by Abeler and colleagues (2016) found a comparable decrease in standardized reports over time. These findings suggest that the magnitude of experimentally studied dishonesty, like other experimentally studied psychological phenomena (e.g., the choice overload effect; Scheibehenne, Greifeneder, & Todd, 2010), declines with new replications, a pattern known as the "decline effect" (Schooler, 2011) or as the "Proteus phenomenon" (Trikalinos & Ioannidis, 2005). The idea is that the strongest (and usually most counterintuitive) findings are more attractive to editors and investigators, and thus more likely to be published first, whereas further examinations reveal less spectacular findings.

Maximal externality. To assess the effect of externalities, we calculated the upper limit of harm that participants could inflict on other participants by behaving dishonestly. In sender–receiver games, receivers typically gain less when they trust a false message, although the amount lost can vary (e.g., Gneezy, 2005). The other three paradigms do not typically involve such externalities. However, some experiments did change the usual setup, for example, by making one's participant's gain another participant's loss (constant sum; e.g., Gino, Krupka, et al., 2013). Having calculated the maximal externality that could be inflicted on other participants, we converted these costs from national currency units to US dollars using 2015 purchasing power parity exchange rates (see Appendix 4.2 for details). The results suggest an interaction effect of maximal externality and experimental paradigm (Table 4.3). On the one hand, standardized reports decreased with greater externalities in sender–receiver games. Here, every additional US$ kept from receivers was associated with a 1 percentage points decrease in false messages sent. By contrast, the opposite effect was observed for coin-flip tasks, where increasing externalities were associated with more dishonest behavior. Specifically, every additional

US$ kept from receivers was associated with a 4 percentage points increase in the standardized report. Notably, the effect of maximal externality was independent of what the participants themselves stood to gain, as discussed next.

Maximal gain. In order to test whether reward sizes affected dishonesty, we calculated the upper limit that participants could themselves gain from behaving dishonestly, again converting national currency units to 2015 US$ purchasing power parity. We observed another interaction effect with experimental paradigm (Table 4.3). In sender–receiver games, greater potential reward sizes were linked to higher standardized reports (Gneezy, 2005); specifically, every additional US$ increased the false messages sent by 1 percentage point. By contrast, in coin-flip tasks, every additional US$ in maximal gain decreased the standardized report by 4 percentage points, suggesting that greater rewards were associated with more honest reporting.

We further inspected the effect of reward size on dishonesty in a more fine-grained analysis. A total of 25 studies experimentally manipulated the incentive size and randomly assigned participants to either a high or a low gain condition while keeping all other experimental parameters constant. This setup allows isolating the effect of possible gains on dishonest behavior. In order to compare the effect of maximum gain (*MG*) across the experiments, we subtracted the standardized report of the highest paying condition in a given experiment ($M_{r\ for\ high\ MG}$) from the standardized report of the lowest paying condition ($M_{r\ for\ low\ MG}$). We thus obtained the mean differences, *D*, in the standardized report for every experiment that manipulated reward size, D (= $M_{r\ for\ high\ MG} - M_{r\ for\ low\ MG}$). We then combined all mean differences through a random effects model. In essence, the model estimates to what degree an increase in reward size causes the standardized report to change. Overall, the results confirm that greater reward sizes were associated with higher standardized reports in sender–receiver games. No such effect was observed for the other three paradigms, including coin-flip tasks (Figure 4.10).

Source	n	M_r	MG_{high} (SD)	MG_{high}	n	M_r	MG_{low} (SD)	MG_{low}	D	95%−CI	Weight
1. Sender−receiver games											
Gneezy, 2005	75	52%	(0.50)	$10.32	75	17%	(0.38)	$ 1.03	0.35	[0.20; 0.49]	6.7%
Sutter, 2009: Individuals	93	59%	(0.49)	$12.72	96	35%	(0.48)	$ 1.27	0.24	[0.10; 0.38]	6.8%
Erat & Gneezy, 2012: Study 1	102	65%	(0.48)	$10.00	101	49%	(0.50)	$ 1.00	0.16	[0.03; 0.30]	6.9%
Wang & Murnighan, 2016: Study 4	27	44%	(0.51)	$ 2.75	21	38%	(0.50)	$ 2.00	0.06	[−0.22; 0.35]	3.6%
Sutter, 2009: Teams	24	25%	(0.44)	$12.72	22	23%	(0.43)	$ 1.27	0.02	[−0.23; 0.27]	4.2%
Random effect	**321**				**315**				**0.20**	**[0.09; 0.31]**	**28.0%**
Heterogeneity: $I^2 = 45.72\%$, $\tau^2 = 0.0064$, $p = 0.12$											
2. Coin−flip tasks											
Hugh−Jones, 2016: Turkey	38	47%	(0.89)	$ 3.95	47	15%	(1.00)	$ 2.37	0.32	[−0.08; 0.73]	2.2%
Hugh−Jones, 2016: Greece	58	38%	(0.93)	$ 8.20	36	6%	(1.01)	$ 4.92	0.32	[−0.09; 0.73]	2.2%
Hugh−Jones, 2016: United Kingdom	42	14%	(1.00)	$ 7.22	47	−6%	(1.01)	$ 4.33	0.21	[−0.21; 0.62]	2.1%
Hugh−Jones, 2016: Denmark	33	33%	(0.96)	$ 0.67	56	21%	(0.99)	$ 0.40	0.12	[−0.30; 0.54]	2.1%
Hugh−Jones, 2016: China	50	76%	(0.66)	$ 1.44	51	65%	(0.77)	$ 0.87	0.11	[−0.17; 0.39]	3.7%
Hugh−Jones, 2016: Argentinia	40	35%	(0.95)	$ 1.88	58	31%	(0.96)	$ 1.13	0.04	[−0.34; 0.42]	2.4%
Hugh−Jones, 2016: Russia	49	43%	(0.91)	$ 0.22	57	44%	(0.91)	$ 0.13	−0.01	[−0.36; 0.34]	2.8%
Hugh−Jones, 2016: South Korea	62	55%	(0.84)	$ 0.01	41	61%	(0.80)	$ 0.00	−0.06	[−0.38; 0.26]	3.1%
Hugh−Jones, 2016: Japan	50	56%	(0.84)	$ 0.05	49	63%	(0.78)	$ 0.03	−0.07	[−0.39; 0.25]	3.1%
Hugh−Jones, 2016: Switzerland	40	20%	(0.99)	$ 3.92	33	27%	(0.98)	$ 2.35	−0.07	[−0.53; 0.38]	1.9%
Hugh−Jones, 2016: Brazil	52	38%	(0.93)	$ 2.70	49	47%	(0.89)	$ 1.62	−0.08	[−0.44; 0.27]	2.7%
Hugh−Jones, 2016: India	51	53%	(0.86)	$ 0.29	48	62%	(0.79)	$ 0.18	−0.10	[−0.42; 0.23]	3.0%
Hugh−Jones, 2016: Portugal	54	11%	(1.00)	$ 8.43	45	33%	(0.95)	$ 5.06	−0.22	[−0.61; 0.16]	2.4%
Hugh−Jones, 2016: South Africa	47	−11%	(1.01)	$ 0.91	45	20%	(0.99)	$ 0.54	−0.31	[−0.71; 0.10]	2.2%
Hugh−Jones, 2016: United States	36	11%	(1.01)	$ 5.00	45	47%	(0.89)	$ 3.00	−0.36	[−0.78; 0.06]	2.1%
Random effect	**702**				**707**				**−0.01**	**[−0.10; 0.08]**	**37.9%**
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.45$											
3. Die−roll tasks											
Conrads et al., 2014	160	54%	(0.60)	$ 6.36	159	37%	(0.67)	$ 1.27	0.18	[0.04; 0.32]	6.7%
Gachter & Schulz, 2016	138	41%	(0.57)	$ 1.90	99	43%	(0.54)	$ 0.50	−0.03	[−0.17; 0.11]	6.7%
Fischbacher & Follmi−Heusi, 2013	80	34%	(0.68)	$11.76	389	41%	(0.62)	$ 3.92	−0.07	[−0.23; 0.09]	6.2%
Random effect	**378**				**647**				**0.03**	**[−0.12; 0.18]**	**19.6%**
Heterogeneity: $I^2 = 68.85\%$, $\tau^2 = 0.0124$, $p = 0.04$											
4. Matrix tasks											
Gino, Ayal, & Ariely, 2013: Study 2	36	11%	(0.50)	$40.00	35	10%	(0.50)	$20.00	0.01	[−0.22; 0.25]	4.5%
Mazar, Amir, & Ariely, 2008: Study 2, honor code	38	−7%	(0.62)	$40.00	31	−8%	(0.64)	$10.00	0.00	[−0.30; 0.30]	3.4%
Mazar, Amir, & Ariely, 2008: Study 2, control	39	10%	(0.21)	$40.00	30	17%	(0.37)	$10.00	−0.06	[−0.21; 0.09]	6.5%
Random effect	**113**				**96**				**−0.03**	**[−0.15; 0.08]**	**14.4%**
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.83$											
Combined random effect	**1514**				**1765**				**0.05**	**[−0.02; 0.12]**	**100.0%**
Heterogeneity: $I^2 = 49.17\%$, $\tau^2 = 0.0133$, $p < 0.01$											

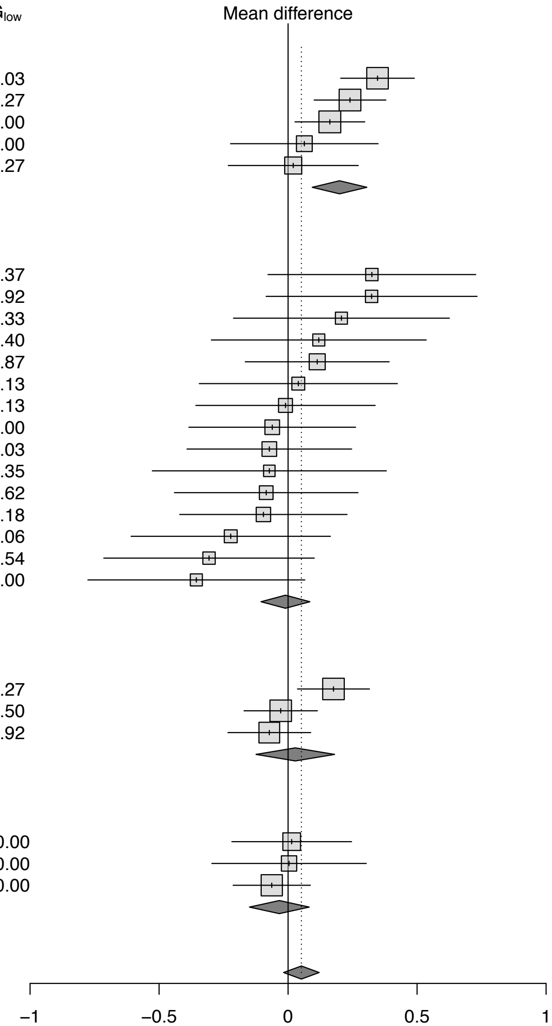
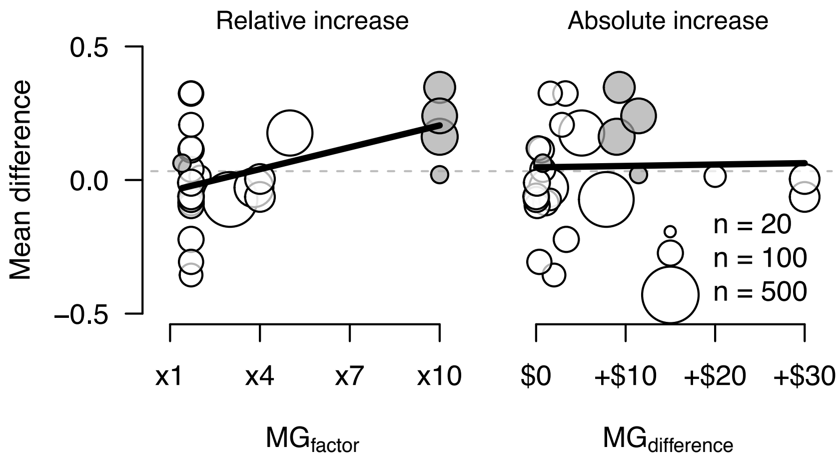


Figure 4.10. *Forest plot: Increasing the incentive was associated with higher standardized reports in sender–receiver games but not in the other experimental paradigms.*

The effect of maximal gains on (dis)honest behavior was potentially confounded by the reward size being increased to a greater extent in sender–receiver games than in the other paradigms. To further inspect the effect of maximal gain (*MG*), we calculated by how many times the maximal gain in the highest paying condition (MG_{high}) was increased relative to the maximal gain in the lowest paying condition (MG_{low}); that is, the relative increase, MG_{factor} (= MG_{high} / MG_{low}). In addition, we calculated the absolute difference in the maximum amount paid in the highest paying condition relative to the lowest paying condition, $MG_{difference}$ (= MG_{high} – MG_{low}). We then independently regressed MG_{factor} and $MG_{difference}$ to predict the mean difference in the standardized report, *D* (Figure 4.11). The greatest relative increase was observed for sender–receiver games (cf. Figure 4.10). Moreover, a relative increase in maximal gains, MG_{factor}, predicted an increase in the standardized report. By contrast, the absolute increase $MG_{difference}$ did not predict the standardized report. Consequently, the observed effect of increasing reward sizes on standardized reports in sender–receiver games may be attributable to the fact that sender–receiver games increased reward sizes by the greatest factor.

Discussion

Various experimental paradigms have been used to assess dishonest behavior. Our first goal was to synthesize four of the most widely used paradigms. Overall, the results suggest that the degree of (dis)honest behavior hinged on the experimental paradigm used. For example, liars in die-roll tasks cheated substantially more than liars in matrix tasks. There was substantial evidence for publication bias in most measures of dishonesty, with large differences between the paradigms. For example, correcting for publication bias increased the standardized report in sender–receiver games, but it decreased the standardized report in die-roll and matrix tasks.



	Relative increase	Absolute increase
Intercept	$b_0 = -0.08$ (0.04)	$b_0 = 0.05$ (0.05)
Maximal gain (MG)	$b_{MG_{factor}} = 3.02\%$**** (0.72)	$b_{MG_{difference}} = 0.06\%$ (0.41)
Observations	$k = 26$ $n = 3417$	$k = 26$ $n = 3417$
Residual heterogenity	$I^2 = 10\%$ $\tau^2 = 0.00$	$I^2 = 51\%$ $\tau^2 = 0.01$
Heterogenity accounted for	$R^2 = 88\%$	$R^2 = 0\%$

Figure 4.11. *A relative increase in maximal gains—but not an absolute increase in maximal gains—predicted change in the standardized report.* Dots in gray represent sender–receiver games; dots in white represent either coin-flip tasks, die-roll tasks, or matrix tasks. The left panel shows change in the standardized report, *D*, as a function of by how many times the reward size was increased, MG_{factor}. The right panel shows the change in the standardized report, *D*, as a function of the absolute difference in maximal gains, $MG_{difference}$. Best-fit regressions are shown by solid black lines. The table below the figures describes the summary statistics of linear regression models with beta weights and standard errors in parentheses, with **** $p < .0001$.

Substantial heterogeneity was also found *within* all four experimental paradigms, suggesting that the manifestation of dishonest behavior strongly depends on situational and personal factors. Assessing these factors and their interaction with experimental paradigm were our second and third goal, respectively. We identified several such situational and/or personal factors: Overall, men behaved slightly more dishonestly than women (but only in sender–receiver games and coin-flip tasks). Younger participants behaved more dishonestly than older participants (but only in die-roll tasks). Nonstudents behaved largely as (dis)honestly as non-economics students (except in die-roll tasks, where students behaved less dishonestly). Economics students lied more than non-economics students (except in die-roll tasks, where economics students behaved more honestly than their fellow students). Online/telephone experiments and, in particular, field experiments were associated with less dishonesty than laboratory experiments (this effect was largely due to coin-flip tasks). In addition, payoffs appear to matter. Decreasing externalities and increasing reward sizes were associated with more dishonesty in sender–receiver games; in coin-flip tasks, the opposite effect was observed. Overall, fine-grained analyses of reward sizes suggested that a reward-induced increase in dishonest behavior was observed only when reward sizes were at least quintupled, circumstances that have only been tested in sender–receiver games.

Limitations

Some limitations of our analysis warrant consideration. First, the quantitative estimates have to be interpreted against the background of potential biases. We found substantial indication of publication bias in almost all measures of dishonest behavior. Trim and fill analyses indicated that the magnitude of dishonest behavior was overestimated in our data base, except in sender–receiver games, where it was more likely underestimated (see also Appendix 4.3). Moreover, the participant composition was far from representative of the general population. For example, student samples contributed substantially more data points to the analyses than other population groups. However, the gender effect—as well as other effects—may have been solely due to student populations. To what degree these effects generalize beyond the typical student population therefore remains an open, empirical question (see, e.g., Abeler et al., 2014). Overall, more representative participant pools seem highly desirable for future research on dishonesty.

A second limitation is that our analyses of personal and situational factors were cross-sectional and correlational. As such, these analyses do not permit causal interpretation. For example, age effects assessed cross-sectionally may have been confounded by cohort effects. Only longitudinal analyses can disentangle how aging as opposed to shared experienced of events (e.g., collective experience of societal corruption) affects behavior (see also Gächter & Schulz, 2016). We are not aware of any such longitudinal analyses on dishonest behavior.

Third, some of our regression models accounted for relative little heterogeneity, suggesting a poor fit of the covariates to the data. In the regression models on sender–receiver games, matrix tasks (Table 4.3), in particular, variables making important contributions to the relatively large heterogeneity seemed to be lacking. For matrix tasks, one reason may be our central measure of dishonest behavior, the standardized report. Matrix tasks typically report dishonest behavior in terms of the absolute number of unsolved matrices claimed as solved—not as the percentage thereof (however, see Appendix 4.3). For sender–receiver games, specific idiosyncrasies in the experimental setup may explain part of the large variance in the amount of dishonest behavior observed (see Appendix 4.3).

Directions for Future Research

Much of the correlational evidence seems worth exploring through more rigorous experimental techniques. For example, most researchers maximally quadrupled the reward size. Yet dishonest behavior seemed relatively robust to such "petty" increases. Our findings indicate that, to causally infer the degree to which dishonest behavior responds to greater reward sizes, experimenters need to randomly assign participants to payoffs that multiply gains by more than a factor of 5. Such magnitudes have not yet been realized in research on experimental paradigms other than sender–receiver games.

Findings on the rates of liars and levels of (dis)honest behavior were mixed and to some extent contradictory between the experimental paradigms. Given that all four experimental paradigms are widely used, it seems worth exploring why they yielded such disparate results. For example, our findings suggest that liars over-report more in die-roll tasks than they do in matrix tasks. Inferring general insights from either experimental paradigm would thus imply different conclusions about the nature of dishonest behavior. The desire to appear honest may explain some of the differences in observations between the experimental paradigms (see Appendix 4.4). However, experimental manipulations would be required to infer causality about when and why people shy away from over-reporting. Such experiments could also help to determine whether participants want to appear honest to themselves (in the form of an internalized norm and thus guilt; as suggested by Mazar et al., 2008) or to the others (in the form of a social norm and thus shame; as implied by the results of Abeler et al., 2014; Conrads & Lotz, 2015; Yaniv & Siniver, 2016). For example, experiments could manipulate the degree to which experimenters are blind to the payment (Hoffmann, Diedenhofen, Verschuere, & Musch, 2015), thus permitting exogenous control for the (perceived) likelihood of appearing dishonest to the experimenter.

It also seems worth directly comparing different experimental paradigms—for example, the reporting behavior of similar participant groups in die-roll tasks versus matrix tasks (Gino, Krupka, et al., 2013; Gravert, 2013). Different internalized and/or social norms may exist for the reporting of outcomes that are randomly generated (die-roll tasks) than for outcomes indicative of one's own effort (matrix tasks). For instance, standards of academic integrity in colleges and universities address issues such as cheating, plagiarism, unfair advantages, and misrepresentation. It could well be the case that skill-based tasks represented as achievement tasks are more likely to remind participants of a code of academic integrity than are chance-based tasks (see Mazar et al., 2008). In this respect, it is also noteworthy that low performers tend to cheat more in matrix tasks than high performers. The reasons for this difference seem unclear. Do low performers cheat more to avoid being seen as incompetent by their peers, by the experimenter, or even by themselves? Do they feel ill treated and cheat more to compensate for participating in the experiment?

Another source of the differences observed in dishonest behavior between the experimental paradigms could be the outcome measure. Coin-flip tasks and die-roll tasks are relatively similar experimental paradigms, the crucial difference being their outcome measure, which is either dichotomous (coin-flip tasks) or continuous (die-roll tasks). Many reviews consequently treat them as two versions of the same experimental paradigm (e.g., Abeler et al., 2016; Jacobsen et al., 2017; Rosenbaum et al., 2014). Yet our results suggest that the two paradigms are associated with notable differences in the rate of liars. In addition, almost all of our analyses of situational and personal factors led to different conclusions for coin-flip tasks versus die-roll tasks (i.e., gender analysis, age analysis, and all but one covariate of the regression analyses presented in Table 4.3). Even the relatively similar standardized reports for coin-flip tasks and die-roll tasks (Figure 4.2) seemed confounded by differences in the experimental setup practices (Table 4.3). Two recent large-scale analyses used die-roll tasks and coin-flip tasks to compare the reporting behavior of citizens of several countries. Whereas the corruption index on the country level predicted reporting behavior in die-roll tasks (Gächter & Schulz, 2016), there was no such effect in coin-flip tasks (Pascual-Ezama et al., 2015). Why do the two experimental paradigms come to such different conclusions?

Generally, it seems worth exploring why the results of one experimental paradigm do not generalize to another. Such insights could have far-reaching consequences on which conclusions can be derived from experimental analyses. Overall, it is important to understand how findings from artificial experiments translate to other settings. Our finding of different results emerging across experimental designs relates to a methodological and theoretical concern that Brunswik (1943, 1944) described as the double standard in the practice of sampling in psychological research: Why is the logic of generalization from participant population to the general population not equally employed to the sampling of experimental stimuli (tasks)? Brunswik's proposed solution was representative design (as opposed to systematic design), which involves randomly sampling experimental stimuli from the environment or designing stimuli in which theoretically or practically important properties are preserved (Dhami, Hertwig, & Hoffrage, 2004). It seems a promising approach to put more systematic thought into what exactly differs between real-world situations that offer the opportunity to behave dishonestly and how those situations are represented in experiments.

Conclusion

In July 2017, Luiz Inácio Lula da Silva, the former president of Brazil, was sentenced to nearly 10 years in prison for corruption and money laundering ("Luiz Inácio Lula da Silva," 2017). Just a few days earlier, the former Israeli prime minister Ehud Olmert was granted parole on charges of bribery, fraud, obstruction of justice, and breach of trust ("Ex-PM Olmert released," 2017). These are just two recent examples from a long list of prominent figures, who have been found guilty of deceiving the public. Exploring the psychological foundations of unethical behavior in the personal, professional, and political spheres is clearly as important as ever. But this also means carefully scrutinizing and understanding the experimental tools used to this end. Our statistical synthesis of 470 experiments offers answers to many of the ongoing debates on who behaves dishonestly and under which circumstances. We showed that the degree and direction of (dis)honest behavior strongly depended on the experimental paradigm, on situational factors (e.g., investigative setting, reward size, externalities), and on personal factors (e.g., gender, age, study major). Yet many questions remain (e.g., why do low performers cheat more in matrix tasks?). We hope that our work will spark more research shedding light on the differences observed and on the factors underlying dishonest behavior.

CHAPTER 5

General Discussion

This dissertation integrated findings from experimental games and explored general theories from various disciplines—including anthropology, economics, psychology, and sociology. The overall aim was to combine these findings and theories to provide broad insights into when and why people behave (un)ethically. In this chapter, I recapitulate the main insights of this dissertation, including its criticisms of experimental games as a research tool.

Throughout this dissertation, I have suggested that when and why people engage in (un)ethical behavior—in the real world and in experimental games—depend on several layers of interpretation, including the perception of the social situation (e.g., the experimental paradigm, context framing), individual disposition (e.g., study major, economics), and the ethical issue itself (e.g., cooperation, honesty). As outlined in Chapter 1, these interpretative layers are not independent but constitute dynamic social processes: Social situations influence individual dispositions, which are the basis of individual decision-making, which leads to individual, concrete behaviors, which become aggregated and form new social situations. I have investigated various aspects of this social framework, contributing to a broader understanding of when and why people engage in (un)ethical behavior. That is, in Chapter 2 I assessed when and why people choose to cooperate by introducing social norm theories and contrasting them to other theories on prosocial behavior, such as social preference theories. By means of meta-analyses, I demonstrated that small changes in the context framing of social dilemma games can have long-lasting effects on participants' propensity to cooperate. Context framing also shaped beliefs about the cooperative behavior of interaction partner(s) as well as donations in dictator games that are known to correlate with cooperation. In combination, these results suggest that cooperation—and its sensitivity to context framing—can be captured by social norm theories. In Chapter 3 I extended the notion of cooperation by introducing explicit sanctions, in the form of costly third-party punishment. The goal was to understand why economics students behave more selfishly than other students. It was demonstrated that economics students and other students had a similar understanding of fairness in the particular experimental game. However, economics students were more skeptical about the cooperativeness of their interaction partners. This skepticism in turn partly mediated their lower willingness to comply with a shared fairness norm. In Chapter 4 I demonstrated that sanctions need not be explicit for ethical behavior to emerge. Instead, implicit and intrinsic sanctions—psychological costs, for example, shame and guilt—can help explain why people refrain from (dis)honest behavior. The level of the (dis)honest behavior thereby strongly depended on several elements, including the experimental paradigm, situational factors, such as the investigative setting and incentives, and personal factors, such as gender, age, and study major.

This dissertation not only built on experimental games, it was also a work of methodological criticism of experimental games. As explained in Chapter 1, experimental games are a promising approach for studying the social framework of (un)ethical behavior. Nonetheless, experimental games have several disadvantages that are worth noting. For example, in Chapter 2 I criticized the idea of "clean" experimental design in studies with human subjects. I argued that participants always bring their own interpretation into the experiment and that this interpretation should not be seen as a source of noise but as an essential element of human cognition. In Chapter 3 I criticized the conventional practice of sampling from highly unrepresentative participant pools. Economics student behave systematically differently compared to students of other majors. Conclusions drawn from experiments with primarily economics students therefore may be inadequate for generalization to other population groups, even to other students. In Chapter 4 I again addressed the question of generalizability, showing that most participant pools in studies on dishonest behavior are far from representative. For instance, widely discussed effects, such as gender differences in dishonesty, are largely based on student samples. I also showed that conclusions about when people behave (dis)honestly partly depend on the specific experimental game. Moreover, I criticized that some experimental games are ill-defined because they leave the unethical consequences of dishonest behavior ambiguous.

In sum, using experimental games to examine the (un)ethical behavior of people is a boon and a bane. On the one hand, experimental games prompt many new questions, such as: How do participants interpret the experimental game (Butler et al., 2011; Gerkey, 2013)? Do individuals behave differently when payoffs are larger than the usual "fistful of dollars" (Hoffman, McCabe, & Smith, 1996a)? How do experimenters influence the behavior of their participants (Bischoff & Frank, 2011)? Overall, it is worth emphasizing that experimental games are, by definition, highly artificial situations that feature aspects rarely observed in natural environments (e.g., “clean” frames; windfall money: Cherry, 2001). It is important to keep these limitations in mind when generalizing insights from experimental games.

On the other hand, experimental games are a highly rigorous investigative tool for studies that tap into the mechanisms of the social framework (cf. Chapter 1). Experimental games, for example, are repeatable encounters, allow straightforward quantification of social behaviors, and provide a taxonomy of social situations. Hence, experimental games can serve as a promising building block for (re)connecting insights from the social and behavioral sciences in an academic world of increasingly isolated disciplines. Formulating general theories and synthesizing empirical findings are essential to a holistic understanding of (un)ethical behavior. I hope to have made a modest contribution toward this ambition.

APPENDICES

Appendices to Chapter 2

APPENDIX 2.1

Conceptualizing the Three Classes of Theories

Social preference theories describe the willingness of individual i to cooperate in social dilemmas and to donate in dictator games as a person-specific function α that describes her own material gain m_i and the material gain of the other individual(s) m_j:

$$C_i = \alpha_i(m_i, m_j) \tag{A2.1.1}$$

In social dilemmas, but not in dictator games, beliefs (in the narrow sense) can affect an individual's choices to cooperate without changing her preferences. Formally speaking, this means that the willingness of individual i to cooperate C_i is the result of a function β_i that describes how her (stable) social preferences $\alpha_i(m_i,m_j)$ are linked to her (flexible) first-order beliefs b_i. Beliefs are defined here as individual i's estimation of the interacting partner(s) choice to cooperate under her interpretation of the context frame f_i:

$$C_i = \beta_i\left(\alpha_i\left(m_i, m_j\right), f_i\left(b_i\right)\right) \tag{A2.1.2}$$

The function β_i entails that individuals with the same social preferences may be differently sensitive to first-order beliefs (i.e., beliefs in the narrow sense), regardless of whether these beliefs are frame-induced or not. For example, whereas a skeptical conditional cooperator may need more evidence than a context frame to believe that her partner will cooperate, a more optimistic conditional cooperator may accept it as sufficient.

Group identity theories assume that the cooperation C_i of individual i is a function of her social preferences $\alpha_i(m_i,m_j)$, which are again a function of her interpretation of the frame f_i:

$$C_i = f_i\left(\alpha_i\left(m_i, m_j\right)\right) \tag{A2.1.3}$$

The function f_i entails that individuals with the same social preferences may be differently sensitive to the frame. That is, the context framing effect depends on their individual *i*'s interpretation of the frame. For example, a frame may induce ingroup preferences among some individuals but outgroup preferences among others.

Social norms theories argue that individual *i*'s cooperation C_i is a function γ_i of beliefs b_i (in the wider sense) and the perception of normative standards e_i. Both b_i and e_i are subject to the interpretation of the context framing f_i:

$$C_i = \gamma_i\left(f_i\left(b_i\right), f_i\left(e_i\right)\right) \tag{A2.1.4}$$

The function γ_i entails that individuals may differ in (1) the weighting they give to adhering to the social norm in the particular situation and/or (2) the weighting they give to beliefs as opposed to perceptions of normative standards. The function γ_i itself is not affected by framing. However, γ_i may be domain specific, and individuals may care more about adhering to a norm in one situation than in another.

Appendix 2.2

Additional Remarks on the Integration and Coding of Experiments

Screening

Our literature search identified context framing experiments involving not only linear public goods games (as described in Figure 2.2) but also threshold public goods games (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001, study 2; Galinsky, Gruenfeld, & Magee, 2003, study 3; Hertel & Fiedler, 1994; Liu & Li, 2009). In threshold public goods games, a shared resource is established only if the sum of individual transfers exceeds a certain provision point. This step-level function differs from the payment function of linear public goods games, in which contributions to the public good are simply multiplied. Meta-analytical comparisons of threshold public goods games are not straightforward when the provision points vary. We therefore excluded them from our investigation. Moreover, we identified one prisoner's dilemma (Chen, Li, Liu, & Shih, 2014, study 1) and one public goods game (Bernold, Gsottbauer, Ackermann, & Murphy, 2015) that used a strategy method to elicit cooperation. We excluded both of these experiments from the meta-analysis.

Eligibility

We now explain the two inclusion criteria in more detail. The first inclusion criterion—including only experiments that manipulated the context frame but whose other parameters were exactly the same—meant that we excluded all articles that used context frames in combination with other, noncontextual variations between treatments, such as alternating the payoff schemes (Handgraaf, van Dijk, Vermunt, Wilke, & de Dreu, 2008; Orwant & Orwant, 1970; Zhong, Loewenstein, & Murnighan, 2007) or assigning participants to different tasks before playing the game (Hoffman et al., 1994). Some experiments had participants recall different elements of previous experiences without precise control of the memories retrieved (Capraro, Smyth, Mylona, & Niblo, 2014; McClure, Bartz, & Lydon, 2013; D. G. Rand, Greene, & Nowak, 2012, study 8; Tao & Au, 2014; Uziel & Hefetz, 2014, study 1 and 2). Moreover, for a study to be included in the meta-analysis, participants of all treatment groups had to believe that they would be interacting with a comparable interaction partner. This condition rules out the possibility that game play is the result of different (but potentially stable) other-regarding preferences for specific interaction partners. We therefore excluded studies that revealed information about the interaction partners' performance in prior tests (Ramalingam, 2012), their surnames (Ahmed, 2010), or their political affiliation (D. G. Rand, Newman, & Wurzbacher, 2014, study 1).

The second inclusion criterion—including only experiments with context frames that referred to comparable concepts—specified that it should be possible to categorize all framed concepts on the two dimensions of the grid-group analysis. In one study, the two coders disagreed on the categorization (Hoffman et al., 1996). Six further studies were excluded because the frames could not be uniquely categorized according to the grid-group analysis (Bicchieri & Xiao, 2009; Bouma, Joy, Paranjape, & Ansink, 2014; Gerkey, 2013; Harrell, 2012; Lesorogol, 2007; Zaleskiewicz, Gasiorowska, & Kesebir, 2015)

If the corresponding authors were unwilling or unable to share the primary data we generated the necessary information from the figures. We could not analyze the last round of interaction in Lib04b or the first or last round of interaction in Gri12, Hri13a, and Hri13b because we did not have access to the primary data of these studies (Appendix 2.3 provides an overview of all integrated studies with their codes). It should also be mentioned that we integrated two experiments co-authored by Dirk Smeesters. The scientific integrity of Smeester's work has been questioned and some of his work has been retracted. The experiments we integrated have previously been tested for violation of scientific integrity (van der Heijden, Groenen, Zeelenberg, & te Lindert, 2014). No irregularities were found.

Coding

In our analysis of game play in all rounds and in the last round, we only integrated repeated dilemma games in which all participants were given full information about their partner's choices after every round and knew that they would interact with the same partner(s) in the next round(s). We allowed repeated social dilemma games in which participants did not receive feedback on their partner's choices between rounds to be integrated in the analysis of the first round by using the mean cooperation rate across all rounds. In our analysis of the last round, we integrated only repeated games in which participants were aware that the game would end immediately afterward (finite time horizon).

Some of the primary experiments compared three or more frames. For example, Brandts and Schwieren compared (1) a cooperative frame with (2) a competitive frame and (3) a generic frame (Brandts & Schwieren, 2009). Combining dependent data for summary effects is problematic (Borenstein, Hedges, Higgins, & Rothstein, 2009, Chapter 25; Higgins & Green, 2011, Chapter 16.5). Rather than dropping these experiments from the investigation, we allowed pair-wise comparisons by evenly dividing the sample size by the number of comparisons. This approach reduced the impact of the experiments with multiple comparisons on the overall effect size (Borenstein et al., 2009, Chapter 25).

Wu11b manipulated the context framing after the first round. This experiment was thus excluded from the analysis of first round effects but it was included in the analysis of repeated games and of the last round. The "paying taxes frame" in Eriksson & Strimling (2014) and the "common heritage frame" in La Barbera, Ferrara, and Boza (2014) did not match any of the four categories of framed concepts. Eriksson and Strimling (2014) did not find a context framing effect between a generic framing condition and the "paying taxes frame". La Barbera, Ferrara, and Boza (2014) did not find a context framing effect between a generic framing condition and the "common heritage frame". In follow-up investigations, compared the "paying taxes frame" with a context frame that did match our framed concepts (Eri14b). La Barbera, Ferrara, and Boza (2014) did the same for the "common heritage frame" (LaB14b). We thus assumed that the "paying taxes frame" and the "common heritage frame" were similar to generic frames.

Appendix 2.3

Complete Lists of the Studies Integrated

Table A2.3.1

Social Dilemma Games Integrated in the Meta-Analysis

Source: Experimental condition	Contrasted frames	Code	*G*	*C*	*R*	*P*	*I*
Deutsch, 1960: no communication condition in combination with simultaneous choice condition	"Cooperative" vs. "individualistic orientation"	Deu60a	PD	US	1	2	no
	"Competitive" vs. "individualistic orientation"	Deu60b	PD	US	1	2	no
	"Cooperative" vs. "individualistic orientation"	Deu60c	PD	US	10	2	no
	"Competitive" vs. "individualistic orientation"	Deu60d	PD	US	10	2	no
Elliott et al., 1998	"Cooperative" vs. "entrepreneur news brief"	Ell98	PD	US	6	4	yes
Batson & Moran, 1999: no communication	"Social exchange" vs. "business transaction"	Bat99a	PD	US	1	2	yes[L]
Batson & Moran, 1999: low empathy	"Social exchange" vs. "business transaction"	Bat99b	PD	US	1	2	yes[L]
Batson & Moran, 1999: high empathy	"Social exchange" vs. "business transaction"	Bat99c	PD	US	1	2	yes[L]
Pillutla & Chen, 1999	"Social task" vs. "investment task"	Pil99	PGG	HK	2	4	yes
Smeesters et al., 2003: study 4	"Neutral primes" vs. "morality primes"	Sme03a	PGG	BE	1	10	no
	"Neutral primes" vs. "might primes"	Sme03b	PGG	BE	1	10	no
Liberman et al., 2004: study 1	"Community game" vs. "Wall Street game"	Lib04a	PD	US	7	2	yes

Note. G = Game type (PD = prisoner's dilemma; PGG = public goods game); C = Country code, indicating where the study was conducted (international vehicle registration code, with o = online studies, – = study not restricted to a particular country); R = Number of rounds; P = Number of interacting partners; I = Incentivized interaction (P = only one selected pair was paid according to the game outcome; L = incentives were in the form of lottery tickets).

Table A2.3.1 (continued)

Source: Experimental condition	Contrasted frames	Code	*G*	*C*	*R*	*P*	*I*
Liberman et al., 2004: study 2	"Kommuna game" vs. "bursa game"	Lib04b	PD	IL	5	2	no
Rege & Telle, 2004: approval condition	"Non-associative frame" vs. "associative frame"	Reg04a	PGG	SE	1	10	yes
Rege & Telle, 2004: no approval condition	"Non-associative frame" vs. "associative frame"	Reg04b	PGG	SE	1	10	yes
Brandts & Schwieren, 2009: study 1	"Public bad frame" vs. "Andreoni frame"	Bra09a	PGG	E	1	3	yes
	"Public good frame" vs. "community game"	Bra09b	PGG	E	1	3	yes
	"Public good frame" vs. "stock exchange game"	Bra09c	PGG	E	1	3	yes
Brandts & Schwieren, 2009: study 2, Decision 4	"Public bad frame" vs. "Andreoni frame"	Bra09d	PGG	E	1	4	yes
Brandts & Schwieren, 2009: study 2, Decision 8	"Public bad frame" vs. "Andreoni frame"	Bra09e	PGG	E	1	2	yes
Brandts & Schwieren, 2009: study 2, Decision 4	"Public good frame" vs. "community game"	Bra09f	PGG	E	1	4	yes
Brandts & Schwieren, 2009: study 2, Decision 4	"Public good frame" vs. "stock exchange game"	Bra09g	PGG	E	1	4	yes
Brandts & Schwieren, 2009: study 2, Decision 8	"Public good frame" vs. "community game"	Bra09h	PGG	E	1	2	yes
Brandts & Schwieren, 2009: study 2, Decision 8	"Public good frame" vs. "stock exchange game"	Bra09i	PGG	E	1	2	yes
Ahmed & Salas, 2011: study 2	"Control" vs. "religious prime"	Ahm11a	PD	CL	1	2	yes[p]
Ahmed & Hammarstedt, 2011	"Control condition" vs. "prime condition"	Ahm11b	PGG	SE	1	3	yes[p]
Dufwenberg et al., 2011: give condition	"Neutral frame" vs. "community frame"	Duf11a	PGG	DE	1	3	yes
Dufwenberg et al., 2011: take condition	"Neutral frame" vs. "community frame"	Duf11b	PGG	DE	1	3	yes
Dufwenberg et al., 2011: Appendix B	"Neutral frame" vs. "community frame"	Duf11c	PGG	CH	1	3	yes
Horton, Rand, & Zeckhauser, 2011: study 2	"Neutral prime" vs. "religious prime"	Hor11	PD	US[o]	1	2	yes
Torsvik et al., 2011: no discussion condition	"No label" vs. "label"	Tor11a	PGG	NO	1	5	yes
Torsvik et al., 2011: discussion condition	"No label" vs. "label"	Tor11b	PGG	NO	1	5	yes

Table A2.3.1 (continued)

Source: Experimental condition	Contrasted frames	Code	*G*	*C*	*R*	*P*	*I*
Wu, Loch, & Ahmad, 2011	"Control" vs. "relationship prime"	Wu11a	PGG	FR	15	2	yes
	"Control" vs. "status prime"	Wu11b	PGG	FR	15	2	yes
Ellingsen et al., 2012: study 1	"Community game" vs. "stock market game"	Ell12	PD	SE	1	2	yes
Grinberg, Hristova, & Borisova, 2012	"Proportionality" vs. "unity"	Gri12	PD	BG	40	2	yes
Lopez, Murphy, Spraggon, & Stranlund, 2012	"Baseline" vs. "frame"	Lop12	PGG	CO	15	5	yes
Benjamin et al., 2013: study 1	"Unprimed" vs. "primed"	Ben13a	PGG	US	1	4	yes
Bouma & Ansink, 2013	"Non-framed" vs. "framed"	Bou13	PGG	CR	5	4	yes
Dreber et al., 2013: appendix	"Community game" vs. "profit game"	Dre13	PD	–^{o}	1	2	yes
Hristova et al., 2013	"Players condition" vs. "team condition"	Hri13a	PD	BG	40	2	yes
	"Opponents condition" vs. "players condition"	Hri13b	PD	BG	40	2	yes
Cone & Rand, 2014: time pressure	"Cooperative context" vs. "competitive context"	Con14a	PGG	US^{o}	1	4	yes
Cone & Rand, 2014: time delay	"Cooperative context" vs. "competitive context"	Con14b	PGG	US^{o}	1	4	yes
Engel & Rand, 2014: low temptation	"Baseline" vs. "protection frame"	Eng14a	PD	US^{o}	1	2	yes
	"Baseline" vs. "contribution frame"	Eng14b	PD	US^{o}	1	2	yes
	"Baseline" vs. "competition frame"	Eng14c	PD	US^{o}	1	2	yes
Engel & Rand, 2014: high temptation	"Baseline" vs. "protection frame"	Eng14d	PD	US^{o}	1	2	yes
	"Baseline" vs. "contribution frame"	Eng14e	PD	US^{o}	1	2	yes
	"Baseline" vs. "competition frame"	Eng14f	PD	US^{o}	1	2	yes
Eriksson & Strimling, 2014: studies 1–2	"No label" vs. "teamwork frame"	Eri14a	PGG	US^{o}	1	4	no
Eriksson & Strimling, 2014: study 3	"Paying taxes frame" vs. "teamwork frame"	Eri14b	PGG	US^{o}	1	4	yes

Table A2.3.1 (continued)

Source: Experimental condition	Contrasted frames	Code	*G*	*C*	*R*	*P*	*I*
La Barbera et al., 2014: study 1	"Control condition" vs. "common project condition"	LaB14a	PGG	IT	1	2	yes[P]
La Barbera et al., 2014: study 2	"Common heritage" vs. "common project condition"	LaB14b	PGG	IT	1	2	yes[P]
Rand, Dreber, et al., 2014: study 2	"Neutral" vs. "Christian prime"	Ran14a	PD	–[o]	1	2	yes
	"Neutral" vs. "Hindu prime"	Ran14b	PD	–[o]	1	2	yes
	"Neutral" vs. "secular prime"	Ran14c	PD	–[o]	1	2	yes
Rand, Newman, et al., 2014: study 2, time press.	"Collaboration context" vs. "competition context"	Ran14d	PGG	US[o]	1	4	yes
Rand, Newman, et al., 2014: study 2, time delay	"Collaboration context" vs. "competition context"	Ran14e	PGG	US[o]	1	4	yes
Bosch-Domènech & Silvestre, 2015	"Framed treatment" vs. "frameless treatment"	Bos15	PD	E	1	3	yes
Bernold et al., 2015: one-shot public goods game	"Neutral frame" vs. "community frame"	Ber15a	PGG	CH	1	4	yes
	"Neutral frame" vs. "Wall Street frame"	Ber15b	PGG	CH	1	4	yes
Bernold et al., 2015: repeated public goods game	"Neutral frame" vs. "community frame"	Ber15d	PGG	CH	1	4	yes
	"Neutral frame" vs. "Wall Street frame"	Ber15e	PGG	CH	1	4	yes
Drouvelis, Metcalfe, & Powdthavee, 2015	"Neutral" vs. "primed"	Dro15	PGG	GB	1	3	yes
de Haan & van Veldhuizen, 2015: study 1, contr.	"Community game" vs. "banker game"	Haa15a	PD	NL	1	2	yes
de Haan & van Veldhuizen, 2015: study 1, depl.	"Community game" vs. "banker game"	Haa15b	PD	NL	1	2	yes
de Haan & van Veldhuizen, 2015: study 2, contr.	"Community game" vs. "banker game"	Haa15c	PD	NL	1	2	yes
de Haan & van Veldhuizen, 2015: study 2, depl.	"Community game" vs. "banker game"	Haa15d	PD	NL	1	2	yes
de Haan & van Veldhuizen, 2015: study 3, contr.	"Community game" vs. "banker game"	Haa15e	PD	NL	1	2	no
de Haan & van Veldhuizen, 2015: study 3, depl.	"Community game" vs. "banker game"	Haa15f	PD	NL	1	2	no

Table A2.3.2

Dictator Games Integrated in the Meta-Analysis

Source: Experimental condition	Contrasted context frames	Code	*G*	*C*	*R*	*P*	*I*
Brañas-Garza, 2007	"T1" vs. "T2"	Bra07a	DG	E	1	2	yes[C]
	"R1" vs. "R2"	Bra07b	DG	E	1	2	yes
Shariff & Norenzayan, 2007: study 1	"No prime" vs. "God concepts prime"	Sha07a	DG	CA	1	2	yes
Shariff & Norenzayan, 2007: study 2	"Neutral prime" vs. "God concepts prime"	Sha07b	DG	CA	1	2	yes
	"Neutral prime" vs. "secular prime"	Sha07c	DG	CA	1	2	yes
Brañas-Garza et al., 2010	"Baseline" vs. "framing"	Bra10	DG	E	1	2	yes
Ahmed & Salas, 2011: study 1	"Control" vs. "prime"	Ahm11c	DG	CL	1	2	yes[P]
Benjamin et al., 2013: study 4	"Unprimed" vs. "primed"	Ben13b	DG	US	1	2	yes
DeScioli & Krishna, 2013: studies 1 & 2	"Baseline" vs. "high need"	DeS13a	DG	–[o]	1	2	yes
DeScioli & Krishna, 2013: studies 1 & 3	"Baseline" vs. "high debt"	DeS13b	DG	–[o]	1	2	yes
	"Baseline" vs. "high need"	DeS13c	DG	–[o]	1	2	yes
Banerjee & Chakravarty, 2014	"Frame I" vs. "frame GA"	Ban14a	DG	IN	1	2	yes
	"Frame I" vs. "frame GAO"	Ban14b	DG	IN	1	2	yes
Gomes & McCullough, 2015	"Control" vs. "standard religious prime"	Gom15a	DG	US	1	2	yes
	"Control" vs. "enhanced religious prime"	Gom15b	DG	US	1	2	yes

Note. *G* = Game type (DG = dictator game); *C* = Country code, indicating where the study was conducted (international vehicle registration code, with [o] = online studies, – = study not restricted to a particular country); *R* = Number of rounds; *P* = Number of interacting partners; *I* = Incentivized interaction (with [P] = only one selected pair was paid according to the game outcome; [C] = incentives were in the form of course credits).

Appendix to Chapter 3

Appendix 3

Instructions for the Third-Party Punishment Game

Thank you for participating. This questionnaire will take approximately 7 minutes of your time. Furthermore, you have the chance to obtain real money. Please read the following instructions carefully.

General Information

This study is the empirical foundation for a master's thesis at the Department of Social and Developmental Psychology, University of Cambridge. The purpose of this study is to analyse decisions and perceptions in transaction situations. Please note that none of the tasks test your personal intelligence or ability.

Consent

This study has received ethical approval from the Departmental Research Committee. Participation is voluntary and you may refuse to answer certain questions. You may withdraw from the study at any time without penalty. This does not waive your legal rights. Results may be presented at conferences and published in journals. Results are presented in terms of groups of individuals. Any data presented would be totally anonymous.

At the end of this study, you may state your e-mail address. This is the only confidential data recorded. It will not be forwarded, and no third party will have access. It is retained purely to inform you of the study's outcome. Your e-mail address and all other data are handled in accordance with the University's Data Protection Act, 1998 (for further information, please see http://www.admin.cam.ac.uk/univ/information/dpa/).

Procedure and Payment

The study is in two parts. In Part One, participants make decisions in a transaction situation. This involves real money transfers. Twenty-one participants will be randomly chosen and paid according to these transactions. In end-February, randomly selected participants are informed and paid. In Part Two, participants are asked about their habits and personal background. If you have any further questions, please contact the conducting researcher.

Transaction Situation

This study comprises three roles: A, B, and C. You are asked to make decisions in each one sequentially. At the end of the study, you are randomly assigned to one role. All roles are assigned to real people with whom you will interact. The interaction does not repeat itself.

Stage one. In Stage One, role A is the sole decision-maker, B is passive and C simply observes the interaction. Role A gets an endowment of £12. Passive role B gets no endowment. Active role A must decide how much of the £12 is assigned to passive role B. Role A can transfer any figure in pound sterling between 0 and 12. If, for example, role A grants role B £5, role A's income at the end of stage one will amount to £7, and role B's income will be £5. If role A grants role B £0, role A's income will be £12, and role B's will be £0.

Stage two. In Stage Two, role C is the sole decision-maker. Roles A and B are entirely passive. Role C has two options. They can either accept or alter role A's decision in Stage One. If role C decides to accept role A's decision, A and B are paid in accordance with Stage One. Role C receives £7 as an endowment and the transaction is complete. If role C decides to alter role A's decision, role C's endowment is limited to £5. However, role C may then proceed to reallocate a new figure in pound sterling between 0 and 12 to role B. If, for example, role A decided to allocate £0 to role B in Stage One and role C disagrees, role C might alter the allocation to £9. In consequence, role A's income would be £3.

You are asked to make the decisions in each of these roles.

This is stage one. You are role B. How much of the £12 do you expect role A will decide to allocate to you?

This is still stage one. You are now role A. How much of the £12 do you wish to allocate to role B? In two to four sentences, please explain the reason behind your decision.

This is stage two. You are role C. Role A decided to grant role B £[0–6]. I accept the allocation. Thus, I get £7; role B gets £[0–6]; and role A will receive £[6–0]. I alter the allocation. My endowment is limited to £5, and I reallocate the money to role B.

This is still stage two. You decided to alter role A's decision. Please state how much you want role B to obtain. Role A will receive the remainder.

The transaction has now ended. In the following, we ask you questions about your perception and personal background.

Perception. What would be a fair allocation of the £12 to role B?

Appendices to Chapter 4

Appendix 4.1

Inclusion Criteria and Integrated Experiments

This section details the article selection process. First, in order to exclude possible learning effects within experiments, we integrated only *one-shot* experiments in which participants had a single opportunity to behave dishonestly. Second, we integrated only experiments that guaranteed full *anonymity* to participants and any partner(s). Revealing the *participant's* identity may elicit reputational concerns, such as honor and shame. Revealing the *assigned partner's* identity may elicit distinct other-regarding preferences, such as different preferences for (mis)reporting to women and men (e.g., Van Zant & Kray, 2014). Third, to rule out fear of explicit punishment as a possible motive for honest reporting, we integrated only experiments in which dishonest actions could not be *sanctioned*. Fourth, the outcome that would be reported by a completely honest sample had to be *unambiguous*. For example, only die-roll tasks with fair dies were included. Fifth, we excluded experiments in which dishonest behavior was not directly *incentivized*. For example, experiments in which only third parties, but not the participants themselves, profited from dishonest behavior were excluded.

If the above criteria were fulfilled, we contacted the authors to request the primary data. We wrote at least three emails to the corresponding author. If authors did not reply, could not share the data (for technical reasons), or refused to share the data, we retrieved the necessary data points from the figures and summary statistics reported in the article. One article had to be excluded from further analyses because the authors did not share the primary data and figures or summary statistics were not provided in a codeable manner (Gino & Mogilner, 2014).

The following tables list all integrated experiments in chronological order: sender–receiver games (Table A4.1.1), coin-flip tasks (Table A4.1.2), die-roll tasks (Table A4.1.3), and matrix tasks (Table A4.1.4).

Table A4.1.1

Integrated Sender–Receiver Games

Study: Condition	n	M_r	Data
Gneezy, 2005: treatment 1	75	36%	no
Gneezy, 2005: treatment 2	75	17%	no
Gneezy, 2005: treatment 3	75	52%	no
Dreber & Johannesson, 2008	156	47%	yes
Cohen, Gunia, Kim-Jun, & Murnighan, 2009: individual, certain	46	48%	yes
Cohen, Gunia, Kim-Jun, & Murnighan, 2009: group, certain	38	71%	yes
Cohen, Gunia, Kim-Jun, & Murnighan, 2009: individual, unc.	37	32%	yes
Cohen, Gunia, Kim-Jun, & Murnighan, 2009: group, uncertain	31	19%	yes
Hurkens & Kartik, 2009: treatment 4	58	38%	yes
Hurkens & Kartik, 2009: treatment 5	32	47%	yes
Sutter, 2009: individual, treatment 1	96	44%	yes
Sutter, 2009: individual, treatment 2	96	35%	yes
Sutter, 2009: individual, treatment 3	93	59%	yes
Sutter, 2009: group, treatment 1	22	23%	yes
Sutter, 2009: group, treatment 2	22	23%	yes
Sutter, 2009: group, treatment 3	24	25%	yes
Rode, 2010: study 1, cooperative	32	81%	yes
Rode, 2010: study 1, competitive	32	75%	yes
Rode, 2010: study 2, cooperative	54	70%	yes
Rode, 2010: study 2, competitive	54	70%	yes
Rode, 2010: study 3, cooperative	27	81%	yes
Rode, 2010: study 4, competitive	27	78%	yes
Cohen, Wolf, Panter, & Insko, 2011: study 2	72	32%	yes
Zhong, 2011: study 1, deliberative	22	68%	yes
Zhong, 2011: study 1, intuitive	22	36%	yes
Zhong, 2011: study 2, deliberative	21	62%	yes
Zhong, 2011: study 2, intuitive	20	30%	yes
Burks & Krupka, 2012	27	26%	no
Childs, 2012b: gain frame	49	53%	yes
Childs, 2012b: loss frame	47	62%	Yes
Erat & Gneezy, 2012: study 1, T[1, 10]	101	49%	no

Note. n = number of participants, M_r = standardized report, Data = primary data shared.

Table A4.1.1 (continued)

Study: Condition	n	M_r	Data
Erat & Gneezy, 2012: study 1, T[10, 10]	102	65%	no
Erat & Gneezy, 2012: study 1, T[1, –5]	104	37%	no
Erat & Gneezy, 2012: study 1, T[10, 0]	109	52%	no
Gunia, Wang, Huang, Wang, & Murnighan, 2012: contemplat.	30	13%	yes
Gunia, Wang, Huang, Wang, & Murnighan, 2012: immediate	34	44%	yes
Gunia, Wang, Huang, Wang, & Murnighan, 2012: moral	25	20%	yes
Gunia, Wang, Huang, Wang, & Murnighan, 2012: control	29	31%	yes
Gunia, Wang, Huang, Wang, & Murnighan, 2012: self-interested	28	50%	yes
Hershfield, Cohen, & Thompson, 2012: study 3	27	56%	yes
Vetter, 2012: study 2, control	48	44%	no
Vetter, 2012: study 2, responsibility	46	48%	no
Angelova & Regner, 2013: study 1, obligatory 1€, with payment	32	56%	yes
Angelova & Regner, 2013: study 1, obligatory 1€, without paym.	32	75%	yes
Angelova & Regner, 2013: study 1, obligatory 2€, with payment	32	66%	yes
Angelova & Regner, 2013: study 1, obligatory 2€, without paym.	32	81%	yes
Angelova & Regner, 2013: study 1, voluntary 1€, with payment	31	81%	yes
Angelova & Regner, 2013: study 1, voluntary 1€, without payment	31	94%	yes
Angelova & Regner, 2013: study 1, voluntary 2€, with payment	32	59%	yes
Angelova & Regner, 2013: study 1, voluntary 2€, without paym.	32	78%	yes
Aoki, Akai, & Onoshiro, 2013: A100-S	78	41%	yes
Aoki, Akai, & Onoshiro, 2013: A100-NS	31	32%	yes
Aoki, Akai, & Onoshiro, 2013: A1000-S	70	49%	yes
Aoki, Akai, & Onoshiro, 2013: A1000-NS	27	26%	yes
Cappelen, Sørensen, & Tungodden, 2013: base	68	69%	yes
Cappelen, Sørensen, & Tungodden, 2013: market	65	72%	yes
Cappelen, Sørensen, & Tungodden, 2013: intuition	69	42%	yes
Cappelen, Sørensen, & Tungodden, 2013: personal	67	55%	yes
Childs, 2013: gain frame	48	58%	yes
Childs, 2013: loss frame	49	41%	yes
Erat, 2013: treatment T[–2]	132	67%	no
Erat, 2013: treatment T[–6]	131	64%	No
Gu, Zhong, & Page-Gould, 2013: study 2, normal	31	58%	yes
Gu, Zhong, & Page-Gould, 2013: study 2, fast	32	31%	yes

Table A4.1.1 (continued)

Study: Condition	n	M_r	Data
Gu, Zhong, & Page-Gould, 2013: study 3, normal	31	71%	yes
Gu, Zhong, & Page-Gould, 2013: study 3, fast	36	47%	yes
Gu, Zhong, & Page-Gould, 2013: study 4, normal, decision-mak.	45	56%	yes
Gu, Zhong, & Page-Gould, 2013: study 4, normal, intuitive	46	63%	yes
Gu, Zhong, & Page-Gould, 2013: study 4, fast, decision-making	37	62%	yes
Gu, Zhong, & Page-Gould, 2013: study 4, fast, intuitive	40	38%	yes
Gylfason, Arnardottir, & Kristinsson, 2013	184	44%	yes
Innes & Mitra, 2013: Arizona, control	97	41%	yes
Innes & Mitra, 2013: Arizona, 15% untruthful	25	36%	yes
Innes & Mitra, 2013: Arizona, 40% untruthful	26	46%	yes
Innes & Mitra, 2013: Arizona, 60% untruthful	33	46%	yes
Innes & Mitra, 2013: Arizona, 85% untruthful	52	81%	yes
Innes & Mitra, 2013: California, control	26	42%	yes
Innes & Mitra, 2013: California, 0–2 untruthful	27	37%	yes
Innes & Mitra, 2013: California, 3–5 untruthful	52	71%	yes
Innes & Mitra, 2013: India, control	54	56%	yes
Innes & Mitra, 2013: India, 15% untruthful	39	31%	yes
Innes & Mitra, 2013: India, 85% untruthful	38	71%	yes
Innes & Mitra, 2013: India, control	54	56%	yes
Innes & Mitra, 2013: India, 0–2 untruthful	46	37%	yes
Innes & Mitra, 2013: India, 3–5 untruthful	37	81%	yes
Kouchaki, Smith-Crowe, Brief, & Sousa, 2013: study 3, money	46	46%	yes
Kouchaki, Smith-Crowe, Brief, & Sousa, 2013: study 3, control	45	22%	yes
López-Pérez & Spiegelman, 2013	30	60%	yes
Kouchaki & Smith, 2014: study 3, afternoon	51	65%	yes
Kouchaki & Smith, 2014: study 3, morning	51	43%	yes
Ter Meer, 2014: no feedback, piece rate	21	52%	no
Ter Meer, 2014: no feedback, revenue sharing	21	68%	no
Ter Meer, 2014: no feedback, tournament incentive	21	87%	no
Ter Meer, 2014: feedback, piece rate	24	78%	no
Ter Meer, 2014: feedback, revenue sharing	24	67%	no
Ter Meer, 2014: feedback, tournament incentive	24	67%	no
Welsh et al., 2014: deprivation, influence, caffeine	29	52%	yes

Table A4.1.1 (continued)

Study: Condition	n	M_r	Data
Welsh et al., 2014: no deprivation, influence, caffeine	31	55%	yes
Welsh et al., 2014: deprivation, no influence, caffeine	31	52%	yes
Welsh et al., 2014: no deprivation, no influence, caffeine	25	48%	yes
Welsh et al., 2014: deprivation, influence, no caffeine	28	79%	yes
Welsh et al., 2014: no deprivation, influence, no caffeine	28	50%	yes
Welsh et al., 2014: deprivation, no influence, no caffeine	25	32%	yes
Welsh et al., 2014: no deprivation, no influence, no caffeine	30	47%	yes
Winterich, Mittal, & Morales, 2014: study 1b, neutral	25	35%	no
Winterich, Mittal, & Morales, 2014: study 1b, disgust	25	67%	no
Winterich, Mittal, & Morales, 2014: study 3, neutral, cleansing	32	43%	no
Winterich, Mittal, & Morales, 2014: study 3, neutral, no cleans.	32	30%	no
Winterich, Mittal, & Morales, 2014: study 3, disgust, no cleans.	32	35%	no
Winterich, Mittal, & Morales, 2014: study 3, disgust, no cleans.	32	53%	no
Biziou-van-Pol et al., 2015: Pareto white lie	598	83%	yes
Kilduff, Galinksy, Gallo, & Reade, 2015: study 2, Michigan	26	46%	yes
Kilduff, Galinksy, Gallo, & Reade, 2015: study 2, Berkeley	23	13%	yes
Kilduff, Galinksy, Gallo, & Reade, 2015: study 2, Virginia	21	19%	yes
Lee, Im, Parmar, & Gino, 2015: study 2	160	43%	yes
Mai, Ellis, & Welsh, 2015: creativity activation	93	44%	yes
Mai, Ellis, & Welsh, 2015: no creativity activation	85	41%	yes
Peeters, Vorsatz, & Walzl, 2015: treatment SR	192	28%	yes
Gylfason, Halldorsson, & Kristinsson, 2016	143	36%	yes
Roeser et al., 2016: message task	195	22%	yes
Wang & Murnighan, 2016: study 1, six-choice, baseline	20	50%	yes
Wang & Murnighan, 2016: study 1, six-choice, exhortation	20	40%	yes
Wang & Murnighan, 2016: study 1, six-choice, maximize profit	20	50%	yes
Wang & Murnighan, 2016: study 1, six-choice, money for hon.	22	14%	yes
Wang & Murnighan, 2016: study 1, six-choice, extra money only	20	40%	yes
Wang & Murnighan, 2016: study 1, two-choice, baseline	17	47%	yes
Wang & Murnighan, 2016: study 1, two-choice, exhortation	18	61%	yes
Wang & Murnighan, 2016: study 1, two-choice, maximize profit	15	53%	yes
Wang & Murnighan, 2016: study 1, two-choice, money for hon.	22	14%	yes
Wang & Murnighan, 2016: study 1, two-choice, extra money only	16	69%	yes

Table A4.1.1 (continued)

Study: Condition	n	M_r	Data
Wang & Murnighan, 2016: study 2, baseline no. 1	23	61%	yes
Wang & Murnighan, 2016: study 2, baseline no. 2	23	70%	yes
Wang & Murnighan, 2016: study 2, money for honesty	24	29%	Yes
Wang & Murnighan, 2016: study 2, exhortation no. 1	26	58%	yes
Wang & Murnighan, 2016: study 2, exhortation no. 2	24	54%	yes
Wang & Murnighan, 2016: study 3, baseline, $4	42	50%	yes
Wang & Murnighan, 2016: study 3, baseline, $8	36	61%	yes
Wang & Murnighan, 2016: study 3, baseline, $12	47	51%	yes
Wang & Murnighan, 2016: study 3, baseline, $16	30	53%	yes
Wang & Murnighan, 2016: study 3, baseline, $20	39	44%	yes
Wang & Murnighan, 2016: study 3, money for honesty, $4	43	28%	yes
Wang & Murnighan, 2016: study 3, money for honesty, $8	43	35%	yes
Wang & Murnighan, 2016: study 3, money for honesty, $12	40	33%	yes
Wang & Murnighan, 2016: study 3, money for honesty, $16	44	27%	yes
Wang & Murnighan, 2016: study 3, money for honesty, $20	36	44%	yes
Wang & Murnighan, 2016: study 4, baseline	23	57%	yes
Wang & Murnighan, 2016: study 4, money for honesty, $1.00	21	38%	yes
Wang & Murnighan, 2016: study 4, money for honesty, $0.75	21	14%	yes
Wang & Murnighan, 2016: study 4, money for honesty, $0.50	20	45%	yes
Wang & Murnighan, 2016: study 4, money for honesty, $0.25	27	44%	yes

Table A4.1.2

Integrated Coin-Flip Tasks

Study: Condition	n	M_r	Data
Bucciol & Piovesan, 2011: control	89	71%	yes
Bucciol & Piovesan, 2011: request	93	38%	yes
Houser, Vetter, & Winter, 2012: receivers who earned 0€	96	65%	no
Houser, Vetter, & Winter, 2012: receivers who earned 2€	75	41%	no
Houser, Vetter, & Winter, 2012: receivers who earned ≥4€	80	50%	no
Houser, Vetter, & Winter, 2012: proposers	251	45%	no
Houser, Vetter, & Winter, 2012: no intentions treatment, 0€	92	41%	no

Note. n = number of participants, M_r = standardized report, Data = primary data shared.

Table A4.1.2 (continued)

Study: Condition	n	M_r	Data
Houser, Vetter, & Winter, 2012: no intentions treatment, 2€	71	52%	no
Houser, Vetter, & Winter, 2012: no intentions treatment, ≥4€	75	36%	no
Fosgaard, Hansen, & Piovesan, 2013: 5 wins, handwritten	53	32%	yes
Fosgaard, Hansen, & Piovesan, 2013: 5 wins, pre-printed	53	28%	yes
Fosgaard, Hansen, & Piovesan, 2013: 10 wins, handwritten	51	69%	yes
Fosgaard, Hansen, & Piovesan, 2013: 10 wins, pre-printed	52	42%	yes
Hilbig & Hessler, 2013: target number 1	127	10%	yes
Hilbig & Hessler, 2013: target number 2	127	22%	yes
Hilbig & Hessler, 2013: target number 3	128	33%	yes
Hilbig & Hessler, 2013: target number 4	128	34%	yes
Hilbig & Hessler, 2013: target number 5	128	32%	yes
Hilbig & Hessler, 2013: target number 6	127	9%	yes
Ploner & Regner, 2013: hidden roll, DG, philanthropy	60	87%	yes
Ploner & Regner, 2013: hidden roll, DG, VCG punishment	32	69%	yes
Ploner & Regner, 2013: hidden roll, DG, stand-alone	96	67%	yes
Ploner & Regner, 2013: hidden roll, bonus, philanthropy	64	66%	yes
Ploner & Regner, 2013: hidden roll, bonus, VCG punishment	32	88%	yes
Ploner & Regner, 2013: hidden roll, bonus, stand-alone	32	69%	yes
Abeler, Becker, & Falk, 2014: 1-coin, telephone	658	–11%	yes
Gino & Wiltermuth, 2014: study 4	178	24%	no
Muñoz-Izquierdo et al., 2014: no penalty	90	27%	yes
Muñoz-Izquierdo et al., 2014: penalty	90	40%	yes
Muñoz-Izquierdo et al., 2014: altruistic penalty	90	16%	yes
Winterich, Mittal, & Morales, 2014: study 1a, neutral	98	4%	no
Winterich, Mittal, & Morales, 2014: study 1a, disgust	98	26%	no
Dubois, Rucker, & Galinsky, 2015: self-beneficial condition	75	23%	yes
Hilbig & Zettler, 2015: study 2	88	34%	yes
Hilbig & Zettler, 2015: study 3	185	15%	yes
Hilbig & Zettler, 2015: study 4, concealed game	50	100%	yes
Hilbig & Zettler, 2015: study 6, standard condition	107	34%	yes
Hilbig & Zettler, 2015: study 6, common goods condition	101	25%	yes
Pascual-Ezama et al., 2015: SRT, Austria	30	–20%	yes
Pascual-Ezama et al., 2015: SRT, Belgium	30	13%	yes

Table A4.1.2 (continued)

Study: Condition	n	M_r	Data
Pascual-Ezama et al., 2015: SRT, Colombia	30	13%	yes
Pascual-Ezama et al., 2015: SRT, Denmark	30	40%	yes
Pascual-Ezama et al., 2015: SRT, Finland	30	27%	yes
Pascual-Ezama et al., 2015: SRT, Germany	30	47%	yes
Pascual-Ezama et al., 2015: SRT, Greece	30	20%	yes
Pascual-Ezama et al., 2015: SRT, India	30	–7%	yes
Pascual-Ezama et al., 2015: SRT, Indonesia	30	20%	yes
Pascual-Ezama et al., 2015: SRT, Italy	30	20%	yes
Pascual-Ezama et al., 2015: SRT, Japan	30	13%	yes
Pascual-Ezama et al., 2015: SRT, Netherlands	30	33%	yes
Pascual-Ezama et al., 2015: SRT, Spain	30	53%	yes
Pascual-Ezama et al., 2015: SRT, Turkey	30	7%	yes
Pascual-Ezama et al., 2015: SRT, UK	30	27%	yes
Pascual-Ezama et al., 2015: SRT, USA	30	40%	yes
Pascual-Ezama et al., 2015: WRT, Austria	30	53%	yes
Pascual-Ezama et al., 2015: WRT, Belgium	30	–7%	yes
Pascual-Ezama et al., 2015: WRT, Colombia	30	33%	yes
Pascual-Ezama et al., 2015: WRT, Denmark	30	–13%	yes
Pascual-Ezama et al., 2015: WRT, Finland	30	33%	yes
Pascual-Ezama et al., 2015: WRT, Germany	30	0%	yes
Pascual-Ezama et al., 2015: WRT, Greece	30	20%	yes
Pascual-Ezama et al., 2015: WRT, India	30	33%	yes
Pascual-Ezama et al., 2015: WRT, Indonesia	30	13%	yes
Pascual-Ezama et al., 2015: WRT, Italy	30	–7%	yes
Pascual-Ezama et al., 2015: WRT, Japan	30	20%	yes
Pascual-Ezama et al., 2015: WRT, Netherlands	30	20%	yes
Pascual-Ezama et al., 2015: WRT, Spain	30	34%	yes
Pascual-Ezama et al., 2015: WRT, Turkey	30	33%	yes
Pascual-Ezama et al., 2015: WRT, UK	30	–33%	yes
Pascual-Ezama et al., 2015: WRT, USA	30	7%	yes
Pascual-Ezama et al., 2015: VRT, Austria	30	20%	yes
Pascual-Ezama et al., 2015: VRT, Belgium	30	7%	yes
Pascual-Ezama et al., 2015: VRT, Colombia	30	–13%	yes

Table A4.1.2 (continued)

Study: Condition	*n*	M_r	Data
Pascual-Ezama et al., 2015: VRT, Denmark	30	−20%	yes
Pascual-Ezama et al., 2015: VRT, Finland	30	7%	yes
Pascual-Ezama et al., 2015: VRT, Germany	30	7%	yes
Pascual-Ezama et al., 2015: VRT, Greece	30	−27%	yes
Pascual-Ezama et al., 2015: VRT, India	30	0%	yes
Pascual-Ezama et al., 2015: VRT, Indonesia	30	−20%	yes
Pascual-Ezama et al., 2015: VRT, Italy	30	13%	yes
Pascual-Ezama et al., 2015: VRT, Japan	30	13%	yes
Pascual-Ezama et al., 2015: VRT, Netherlands	30	0%	yes
Pascual-Ezama et al., 2015: VRT, Spain	30	26%	yes
Pascual-Ezama et al., 2015: VRT, Turkey	30	13%	yes
Pascual-Ezama et al., 2015: VRT, UK	30	40%	yes
Pascual-Ezama et al., 2015: VRT, USA	30	27%	yes
Zettler, Hilbig, Moshagen, & de Vries, 2015: study 2	134	17%	yes
Conrads et al., 2016: individual treatment	114	65%	yes
Conrads et al., 2016: team treatment	67	40%	yes
Dieckmann, Grimm, Unfried, Utikal, & Valmasoni, 2016: D	203	34%	yes
Dieckmann, Grimm, Unfried, Utikal, & Valmasoni, 2016: F	202	54%	yes
Dieckmann, Grimm, Unfried, Utikal, & Valmasoni, 2016: I	202	49%	yes
Dieckmann, Grimm, Unfried, Utikal, & Valmasoni, 2016: NL	204	65%	yes
Dieckmann, Grimm, Unfried, Utikal, & Valmasoni, 2016: ESP	204	45%	yes
Houser et al, 2016: Parent alone, prize for parent	61	10%	yes
Hugh-Jones, 2016: high treatment, Argentina	40	35%	yes
Hugh-Jones, 2016: high treatment, Brazil	52	38%	yes
Hugh-Jones, 2016: high treatment, Switzerland	40	20%	yes
Hugh-Jones, 2016: high treatment, China	50	76%	yes
Hugh-Jones, 2016: high treatment, Denmark	33	33%	yes
Hugh-Jones, 2016: high treatment, United Kingdom	42	14%	yes
Hugh-Jones, 2016: high treatment, Greece	58	38%	yes
Hugh-Jones, 2016: high treatment, India	51	53%	yes
Hugh-Jones, 2016: high treatment, Japan	50	56%	yes
Hugh-Jones, 2016: high treatment, South Korea	62	55%	yes
Hugh-Jones, 2016: high treatment, Portugal	54	11%	Yes

Table A4.1.2 (continued)

Study: Condition	n	M_r	Data
Hugh-Jones, 2016: high treatment, Russian Federation	49	43%	yes
Hugh-Jones, 2016: high treatment, Turkey	38	47%	yes
Hugh-Jones, 2016: high treatment, USA	36	11%	yes
Hugh-Jones, 2016: high treatment, South Africa	47	–11%	yes
Hugh-Jones, 2016: low treatment, Argentina	58	31%	yes
Hugh-Jones, 2016: low treatment, Brazil	49	47%	yes
Hugh-Jones, 2016: low treatment, Switzerland	33	27%	yes
Hugh-Jones, 2016: low treatment, China	51	65%	yes
Hugh-Jones, 2016: low treatment, Denmark	56	21%	yes
Hugh-Jones, 2016: low treatment, United Kingdom	47	-6%	yes
Hugh-Jones, 2016: low treatment, Greece	36	6%	yes
Hugh-Jones, 2016: low treatment, India	48	63%	yes
Hugh-Jones, 2016: low treatment, Japan	49	63%	yes
Hugh-Jones, 2016: low treatment, South Korea	41	61%	yes
Hugh-Jones, 2016: low treatment, Portugal	45	33%	yes
Hugh-Jones, 2016: low treatment, Russian Federation	57	44%	yes
Hugh-Jones, 2016: low treatment, Turkey	47	15%	yes
Hugh-Jones, 2016: low treatment, USA	45	47%	yes
Hugh-Jones, 2016: low treatment, South Africa	45	20%	yes
Thielmann, Hilbig, Zettler, & Moshagen, 2016: study 2	152	33%	yes

Table A4.1.3

Integrated Die-Roll Tasks

Study: Condition	n	M_r	Data
Lammers, Stapel, & Galinsky, 2010: study 1, high power	17	41%	yes
Lammers, Stapel, & Galinsky, 2010: study 1, low power	18	19%	yes
Gino & Ariely, 2011: study 3, control, low justification	36	4%	no
Gino & Ariely, 2011: study 3, control, high justification	36	42%	no
Gino & Ariely, 2011: study 3, creative mindset, high justification	36	53%	no
Gino & Ariely, 2011: study 3, creative mindset, high justification	36	57%	no
Gino & Ariely, 2011: study 4, control, low justification	40	–23%	no

Note. n = number of participants, M_r = standardized report, Data = primary data shared.

Table A4.1.3 (continued)

Study: Condition	n	M_r	Data
Gino & Ariely, 2011: study 4, control, high justification	40	15%	no
Gino & Ariely, 2011: study 4, creative mindset, low justification	40	41%	no
Gino & Ariely, 2011: study 4, creative mindset, high justification	40	47%	no
Shalvi, Dana, Handgraaf, & De Dreu, 2011: single roll	62	19%	yes
Shalvi, Dana, Handgraaf, & De Dreu, 2011: multiple roll	67	38%	yes
Shalvi, Handgraaf, & De Dreu, 2011: 3.50€ exit	25	20%	yes
Shalvi, Handgraaf, & De Dreu, 2011: 2.50€ exit	30	32%	yes
Piff, Stancato, Côté, Mendoza-Denton, & Keltner, 2012: study 6	189	4%	yes
Shalvi, Eldar, & Bereby-Meyer, 2012: study 1, low time pressure	38	15%	yes
Shalvi, Eldar, & Bereby-Meyer, 2012: study 2, high time pressure	34	42%	yes
Shalvi, Eldar, & Bereby-Meyer, 2012: study 1, low time pressure	33	–3%	yes
Shalvi, Eldar, & Bereby-Meyer, 2012: study 2, high time pressure	39	35%	yes
Wibral, Dohmen, Klingmüller, Weber, & Falk, 2012: placebo	45	45%	yes
Wibral, Dohmen, Klingmüller, Weber, & Falk, 2012: testost.	46	–5%	yes
Conrads et al., 2013: individual	156	32%	yes
Conrads et al., 2013: team	132	54%	yes
Conrads et al., 2013: team-mixed, individual	130	34%	yes
Conrads et al., 2013: team-mixed, team	136	45%	yes
Fischbacher & Föllmi-Heusi, 2013: baselines	389	41%	yes
Fischbacher & Föllmi-Heusi, 2013: high stakes	80	34%	yes
Fischbacher & Föllmi-Heusi, 2013: 4.9	125	38%	yes
Fischbacher & Föllmi-Heusi, 2013: externality	78	27%	yes
Fischbacher & Föllmi-Heusi, 2013: double anonymous	137	37%	yes
Gino, Krupka, & Weber, 2013: mandatory regulation	30	17%	no
Gino, Krupka, & Weber, 2013: no regulation	30	44%	No
Gino, Krupka, & Weber, 2013: voluntary regulation	30	48%	no
Gravert, 2013: random income	57	8%	yes
Shalvi & Leiser, 2013: religious track	65	9%	yes
Shalvi & Leiser, 2013: regular track	61	27%	yes
Utikal & Fischbacher, 2013: students	19	54%	yes
Utikal & Fischbacher, 2013: nuns	12	–33%	yes
Abeler, Becker, & Falk, 2014: 4-coin, telephone	94	–6%	yes
Abeler, Becker, & Falk, 2014: 4-coin-lab, telephone	170	32%	yes

Table A4.1.3 (continued)

Study: Condition	*n*	M_r	Data
Abeler, Becker, & Falk, 2014: 4-coin-lab, click	180	39%	yes
Arbel, Bar-El, Siniver, & Tobol, 2014: study 1	205	27%	yes
Arbel, Bar-El, Siniver, & Tobol, 2014: study 3, incentive to lie	194	27%	yes
Clot, Grolleau, & Ibanez, 2014: control	49	3%	yes
Clot, Grolleau, & Ibanez, 2014: good deed	49	26%	yes
Conrads et al., 2014: treatment 1	159	37%	yes
Conrads et al., 2014: treatment 3	159	48%	yes
Conrads et al., 2014: treatment 5	160	54%	yes
Gunia, Barnes, & Sah, 2014: study 2, MTurk	99	21%	yes
Gunia, Barnes, & Sah, 2014: study 2, students	43	26%	yes
Ruffle & Tobol, 2014	427	15%	yes
Chou, 2015: study 1, e-signature	30	28%	no
Chou, 2015: study 1, handwritten signature	28	11%	no
Conrads & Lotz, 2015: face-to-face	60	37%	yes
Conrads & Lotz, 2015: phone	60	38%	yes
Conrads & Lotz, 2015: computerized, lab	60	36%	yes
Conrads & Lotz, 2015: computerized, remote	66	43%	yes
Jacobsen & Piovesan, 2015: baseline	50	1%	yes
Jacobsen & Piovesan, 2015: tax framing	50	18%	yes
Jacobsen & Piovesan, 2015: explanation	49	−4%	yes
Muehlheusser, Roider, & Wallmeier, 2015: individual	108	39%	yes
Muehlheusser, Roider, & Wallmeier, 2015: team	60	39%	yes
Cadsby, Du, & Song, 2016: die-roll self	90	34%	yes
Dai, Galeotti, & Villeval, 2016: fine collection office	35	26%	yes
Dai, Galeotti, & Villeval, 2016: station	244	43%	yes
Gächter & Schulz, 2016: Austria	66	30%	yes
Gächter & Schulz, 2016: China, high stakes	138	41%	yes
Gächter & Schulz, 2016: China, low stakes	99	43%	yes
Gächter & Schulz, 2016: Colombia	104	35%	yes
Gächter & Schulz, 2016: Czech Republic	77	35%	yes
Gächter & Schulz, 2016: Georgia	97	36%	yes
Gächter & Schulz, 2016: Germany	69	20%	yes
Gächter & Schulz, 2016: Guatemala	193	31%	yes

Table A4.1.3 (continued)

Study: Condition	*n*	M_r	Data
Gächter & Schulz, 2016: Indonesia	76	35%	yes
Gächter & Schulz, 2016: Italy	82	23%	yes
Gächter & Schulz, 2016: Kenya	92	37%	yes
Gächter & Schulz, 2016: Lithuania	71	18%	yes
Gächter & Schulz, 2016: Malaysia	64	29%	yes
Gächter & Schulz, 2016: Morocco	138	56%	yes
Gächter & Schulz, 2016: Netherlands	84	30%	yes
Gächter & Schulz, 2016: Poland	110	38%	yes
Gächter & Schulz, 2016: Slovakia	87	25%	yes
Gächter & Schulz, 2016: South Africa	92	29%	yes
Gächter & Schulz, 2016: Spain	54	31%	yes
Gächter & Schulz, 2016: Sweden	82	19%	yes
Gächter & Schulz, 2016: Tanzania	140	58%	yes
Gächter & Schulz, 2016: Turkey	244	39%	yes
Gächter & Schulz, 2016: United Kingdom	197	19%	yes
Gächter & Schulz, 2016: Vietnam	112	38%	yes
Schurr & Ritov, 2016: control	23	3%	yes
Schurr & Ritov, 2016: study 1, winners	20	35%	yes
Schurr & Ritov, 2016: study 1, losers	23	−13%	yes
Schurr & Ritov, 2016: study 2, winners	19	38%	yes
Schurr & Ritov, 2016: study 2, losers	19	3%	yes
Schurr & Ritov, 2016: study 3a, winners	29	−20%	yes
Schurr & Ritov, 2016: study 3a, losers	22	6%	yes
Schurr & Ritov, 2016: study 3b, winners	23	5%	yes
Schurr & Ritov, 2016: study 3b, losers	21	22%	yes

Table A4.1.4

Integrated Matrix Tasks

Study: Condition	*n*	M_r	Data
Mazar, Amir, & Ariely 2008: study 1, Ten Commandments	116	−11%	yes
Mazar, Amir, & Ariely 2008: study 1, ten books	113	7%	yes

Note. n = number of participants, M_r = standardized report, Data = primary data shared.

Table A4.1.4 (continued)

Study: Condition	*n*	M_r	Data
Mazar, Amir, & Ariely 2008: $0.50, control vs. recycle	61	17%	yes
Mazar, Amir, & Ariely 2008: $2, control vs. recycle	77	10%	yes
Mazar, Amir, & Ariely 2008: study 3, control vs. recycle	300	16%	yes
Mazar, Amir, & Ariely 2008: study 4	44	16%	yes
Mazar, Amir, & Ariely 2008: study 5, four matrices solved	52	7%	yes
Mazar, Amir, & Ariely 2008: study 5, eight matrices solved	56	8%	yes
Rhyne, 2008: USA, no cheating vs. cheating to self	85	2%	no
Rhyne, 2008: China, no cheating vs. cheating to self	98	1%	no
Gino, Ayal, & Ariely, 2009: study 1, control vs. shredder	76	39%	no
Gino, Ayal, & Ariely, 2009: study 2, control vs. shredder	61	40%	no
Mead et al., 2009: study 1, no depletion	71	3%	no
Mead et al., 2009: study 1, depletion	62	18%	no
Gino, Norton, & Ariely, 2010: study 1a, authentic sunglasses	43	6%	no
Gino, Norton, & Ariely, 2010: study 1a, counterfeit sunglasses	42	28%	no
Gino, Norton, & Ariely, 2010: study 1b, authentic sunglasses	46	6%	no
Gino, Norton, & Ariely, 2010: study 1b, counterfeit sunglasses	45	21%	no
Gino, Norton, & Ariely, 2010: study 3, control	33	12%	no
Gino, Norton, & Ariely, 2010: study 3, authentic sunglasses	33	9%	no
Gino, Norton, & Ariely, 2010: study 3, counterfeit sunglasses	34	28%	no
Zhong, Bohns, & Gino, 2010: study 1, control room	42	6%	no
Zhong, Bohns, & Gino, 2010: study 1, dim room	42	33%	no
Gino & Ariely, 2011: study 2, control	56	9%	no
Gino & Ariely, 2011: study 2, creative mindset	56	21%	no
Gino, Schweitzer, Mead, & Ariely, 2011: study 1, no depletion	51	8%	no
Gino, Schweitzer, Mead, & Ariely, 2011: study 2, no depletion	49	6%	no
Gino, Schweitzer, Mead, & Ariely, 2011: study 2, depletion	48	18%	no
Gino & Margolis, 2011: study 3, promotion focus	41	30%	no
Gino & Margolis, 2011: study 3, prevention focus	41	8%	no
Gino & Margolis, 2011: study 3, aspiration, promotion focus	34	38%	no
Gino & Margolis, 2011: study 3, aspiration, prevention focus	34	12%	no
Gino & Margolis, 2011: study 3, compliance, promotion focus	34	22%	no
Gino & Margolis, 2011: study 3, compliance, prevention focus	34	4%	no
Shu, Gino, & Bazerman, 2011: study 3, no honor code	70	44%	no

Table A4.1.4 (continued)

Study: Condition	n	M_r	Data
Shu, Gino, & Bazerman, 2011: study 3, honor code	70	17%	no
Shu, Gino, & Bazerman, 2011: study 4, read honor code	22	22%	no
Shu, Gino, & Bazerman, 2011: study 4, signed honor code	22	4%	no
Shu, Gino, & Bazerman, 2011: study 4, control	23	44%	no
Friesen & Gangadharan, 2012	114	11%	yes
Gino & Galinsky, 2012: study 3, shared attributes	41	30%	no
Gino & Galinsky, 2012: study 3, control	41	11%	no
Shu & Gino, 2012: study 1	56	13%	no
Shu & Gino, 2012: study 2	78	12%	no
Shu, Mazar, Gino, Ariely, & Bazerman, 2012: study 1, top	35	8%	no
Shu, Mazar, Gino, Ariely, & Bazerman, 2012: study 1, bottom	33	36%	no
Shu, Mazar, Gino, Ariely, & Bazerman, 2012: study 1, control	33	24%	no
Shu, Mazar, Gino, Ariely, & Bazerman, 2012: study 2, top	30	15%	no
Shu, Mazar, Gino, Ariely, & Bazerman, 2012: study 2, bottom	30	31%	no
Gamliel & Peer, 2013: control vs. shredder	68	12%	yes
Gino, Ayal, & Ariely, 2013: study 1, individual	64	30%	no
Gino, Ayal, & Ariely, 2013: study 1, dyad	64	52%	no
Gino, Ayal, & Ariely, 2013: study 1, group	64	68%	no
Gino, Ayal, & Ariely, 2013: study 2, self only, high payoff	36	11%	no
Gino, Ayal, & Ariely, 2013: study 2, self only, low payoff	35	10%	no
Gino, Ayal, & Ariely, 2013: study 2, self and other payoff	36	27%	no
Gino, Krupka, et al., 2013: study 1, no regulation vs. mandatory	60	25%	no
Gravert, 2013: performance income	48	41%	yes
Gunia, Barnes, & Sah, 2014: study 1	48	5%	yes
Kouchaki, Gino, & Jami, 2014: study 3, light backpack	35	15%	yes
Kouchaki, Gino, & Jami, 2014: study 3, heavy backpack	36	5%	yes
Kouchaki & Wareham, 2015: study 4, morning	27	15%	yes
Kouchaki & Wareham, 2015: study 4, afternoon	21	28%	yes
Rigdon & D'Esterre, 2014: self-grading, non-competitive	48	11%	yes
Rigdon & D'Esterre, 2014: self-grading, competitive	52	9%	yes
Cai, Huang, Wu, & Kou, 2015: study 1, eyes	66	9%	yes
Cai, Huang, Wu, & Kou, 2015: study 1, control	65	9%	yes
Faravelli, Friesen, & Gangadharan, 2015: study 2, piece rate	119	25%	yes

Table A4.1.4 (continued)

Study: Condition	*n*	M_r	Data
Faravelli, Friesen, & Gangadharan, 2015: study 3, tournament r.	119	25%	yes
Faravelli, Friesen, & Gangadharan, 2015: study 4, piece rate	77	27%	yes
Faravelli, Friesen, & Gangadharan, 2015: study 4, tournament r.	41	36%	yes
Kouchaki & Wareham, 2015: study 1, exclusion	20	38%	yes
Kouchaki & Wareham, 2015: study 1, inclusion	18	21%	yes
Lee, Gino, Jin, Rice, & Josephs, 2015: pilot study	82	16%	yes
Lee, Gino, Jin, Rice, & Josephs, 2015: main study	117	17%	yes
Grolleau, Kocher, & Sutan, 2016: gain frame	300	9%	yes
Grolleau, Kocher, & Sutan, 2016: loss frame	300	41%	yes
Hildreth, Gino, & Bazerman, 2016: study 1a, loyalty	28	2%	yes
Hildreth, Gino, & Bazerman, 2016: study 1a, control	27	4%	yes
Hildreth, Gino, & Bazerman, 2016: study 1b, loyalty	33	2%	yes
Hildreth, Gino, & Bazerman, 2016: study 1b, control	30	13%	yes
Hildreth, Gino, & Bazerman, 2016: study 2b	88	2%	yes
Hildreth, Gino, & Bazerman, 2016: study 3a, loyalty	30	2%	yes
Hildreth, Gino, & Bazerman, 2016: study 3a, control	29	6%	yes
Hildreth, Gino, & Bazerman, 2016: study 3b, pledge	31	5%	yes
Hildreth, Gino, & Bazerman, 2016: study 3b, no pledge	39	2%	yes
Hildreth, Gino, & Bazerman, 2016: study 3b, control	36	7%	yes
Hildreth, Gino, & Bazerman, 2016: study 5a, loyalty, low comp.	48	4%	yes
Hildreth, Gino, & Bazerman, 2016: study 5a, loyalty, high comp.	51	14%	yes
Hildreth, Gino, & Bazerman, 2016: study 5a, control, low comp.	52	13%	yes
Hildreth, Gino, & Bazerman, 2016: study 5a, control, high comp.	55	17%	yes
Hildreth, Gino, & Bazerman, 2016: study 5b, loyalty, low comp.	50	4%	Yes
Hildreth, Gino, & Bazerman, 2016: study 5b, loyalty, high comp.	53	12%	yes
Hildreth, Gino, & Bazerman, 2016: study 5b, control, low comp.	55	10%	yes
Hildreth, Gino, & Bazerman, 2016: study 5b, control, high com.	50	11%	yes

Appendix 4.2

Classification of Experiments and Calculation of the Dishonesty Measures

Sender–receiver games. In sender–receiver games, a completely honest participant sample would always send the truthful message, where $t = t_{min} = t_{max} = 0\%$. The actual mean claim of the sample M_r is thus the percentage of false messages, where $M_r = m = M_{liars}$. A study with $m = 0\%$ honest messages thus converts into $M_r = 0\%$ and a study with $m = 100\%$ converts to $M_r = +100\%$. The standardized report for sender–receiver games can never take negative values. We address this methodological concern further below.

Coin-flip tasks. Coin-flip tasks covered all experiments (1) in which reporting was measured dichotomously—that is, involved a decision between claiming to have won a bonus and going empty handed—, and in which (2) the "true" observed outcome had to be randomly generated. In most coin-flip tasks, participants tossed a single coin once and reported the outcome. In such situations, a completely honest sample would report a win about half of the time, $t = 50\%$. Reporting a win results in the highest possible claim, $t_{max} = 100\%$, and reporting a loss, the lowest possible claim, $t_{min} = 0\%$. A study with a mean claim of $m = 0\%$ would thus convert to $M_r = -100\%$; a study with $m = 50\%$ would convert to $M_r = 0\%$; and a study with $m = 100\%$ would convert to $M_r = +100\%$.

Some coin-flip tasks used multiple coin tosses to generate the "true" observed outcome. Here, participants earned a bonus for reporting a specific target number of wins (e.g., two heads in three tosses). If they did not report exactly the target number, they left empty handed (Hilbig & Zettler, 2015; Houser et al., 2012; Thielmann et al., 2016; Zettler et al., 2015). Other designs of coin-flip tasks randomly generated the "true" outcome via some form of dice game. Unlike die-roll tasks, such tasks enforced a dichotomous decision between reporting a win and a loss. For example, participants could win a flat bonus if they reported having rolled an even number (Ploner & Regner, 2013); if the sum of all pips reached a threshold (Conrads et al., 2016; Dubois et al., 2015); or if the die showed a specific target number, such as a 4 (Hilbig & Hessler, 2013; Hilbig & Zettler, 2015). The standardized report M_r and the rate of liars M_{liars} were always adapted to the specific experimental design. For example, a completely honest sample in a target number dice game would report a win about one in six times; therefore, $t = 1/6$.

To estimate the rate of liars in coin-flip tasks, we assumed that participants would always claim a win if they tossed one. Accordingly, the rate of liars M_{liars} was estimated as:

$$M_{liars} = 0\% \quad \text{if } m < t \tag{A4.2.1}$$

$$M_{liars} = \frac{m-t}{1-t} \quad \text{if } m \geq t \tag{A4.2.2}$$

where m is the actual report per study and t is the expected reported percentage of winning sides if participants were honest.

Estimating M_{liars} decreases the number of utilized observations per primary study because only a subset of participants tossed the losing side. We thus adjusted the sample size n per coin-flip task to be the estimated number of participants who tossed the losing side:

$$n = N \times (100\% - t) \tag{A4.2.3}$$

where N is the total sample size per coin-flip task and t is the expected reported percentage of winning sides if participants were honest.

Die-roll tasks. Die-roll tasks included all experiments in which the reported outcome was (1) continuous and (2) the "true" observed outcome was either known (as in Piff et al., 2012) or randomly generated. Most die-roll tasks used some form of dice game to generate the "true" outcome. For example, in a task with a six-sided die where each pip translates into \$1, the minimum report is $t_{min} = 1$, the maximum report is $t_{max} = 6$, and the expected report if everyone was honest is $t = 3.5$—that is, (1 + 2 + 3 + 4 + 5 + 6) / 6; assuming the die is fair. In such a study, an average report of $m = 6$ would convert to $M_r = 100\%$, indicating that all participants who could cheat, cheated to the maximal degree; an average report of $m = 3.5$ would convert to $M_r = 0\%$, indicating honest reporting; and an average report of $m = 1$ would convert to $M_r = -100\%$.

Instead of dice, two experiments used repeated coin tosses, with participants being asked to report how many times they obtained the winning side (Abeler et al., 2014; Conrads & Lotz, 2015). We classified these tasks as die-roll tasks due to their continuous measure of dishonesty. Notably, they were also one-shot tasks, because participants reported their observations only at a single instance. Most die-roll tasks converted pips into money by using either a complete linear payoff function (e.g., \$1 for each reported pip) or a linear payoff function that excluded the highest pip (e.g., \$1 for each reported pip, but reporting a 6 paid \$0). Only a single die-roll task used an alternative, nonlinear payoff function

(Jacobsen & Piovesan, 2015: reporting a 1 paid DKK 10, a 2 paid DKK 20, a 3 paid DKK 30, a 4 paid DKK 35, a 5 paid DKK 40, and a 6 paid DKK 45; DKK = Danish Krone). To calculate the standardized report for this study, we let m, t, and M_r refer to the reported outcomes (e.g., reporting a 5), not the associated payoffs (e.g., claiming DKK 40).

To calculate the rates liars in die-roll tasks, we assumed that (1) all studies used fair dice and (2) participants would always claim at least the number they actually observed. Accordingly, we calculated the percentage of honest responses M_{honest} per primary experiment as:

$$M_{honest} = \frac{C(t_{\min})}{E(t_{\min})} \qquad \text{if } E(t_{\min}) \leq C(t_{\min}) \qquad \text{(A4.2.4)}$$

$$M_{honest} = 0 \qquad \text{if } E(t_{\min}) > C(t_{\min}) \qquad \text{(A4.2.5)}$$

where $C(t_{min})$ is the actual percentage of claims of the lowest paying option, and $E(t_{min})$ is the expected percentage of claims of the lowest paying option if everyone was honest. For example, in a die-roll task with a six-sided die and a linear payoff function, the lowest option has a probability of $E(t_{min}) = 1/6$. If, for example, only $C(t_{min}) = 1/12$ of the participants report the lowest pip, then the estimated percentage of honest participants is $M_{honest} = 50\%$.

We assumed that the percentages of honest responses M_{honest} and dishonest responses M_{liars} per experiment would add up to 100%. Accordingly:

$$100\% - M_{honest} = M_{liars} \qquad \text{(A4.2.6)}$$

where M_{honest} is the percentage of honest responses and M_{liars} is the percentage of dishonest responses.

Two things are worth mentioning about *estimating* the rates of liars in die-roll tasks. First, M_{honest} is the lower bound for the percentage of honest reporting (assuming that participants would always claim at least the number they actually observed). This is because participants who actually obtained t_{min} were maximally tempted to cheat (e.g., there might be participants who cheat when they observe the lowest outcome but would not cheat if they observe a medium outcome). Accordingly, the current estimation of M_{liars} is an upper bound to the real rate of liars in the die-roll task. Second, estimating the rate of liars decreases the number of utilized observations per primary study because only a subset of participants obtained the lowest paying option, which is the basis for calculating M_{honest}. For all analyses that involve the rates of liars M_{liars} we thus adjusted the sample size n per die-roll task to be the estimated number of participants who obtained the lowest paying option per primary study:

$$n = N \times E\left(t_{\min}\right) = N \times E\left(t_{\max}\right) \tag{A4.2.7}$$

where N is the total sample size per die-roll task and $E(t_{min})$ is the expected percentage of claims of the lowest paying option if everyone was honest, and $E(t_{max})$ is the expected percentage of claims of the highest paying option if everyone was honest. In all experiments $E(t_{min})$ was equal to $E(t_{max})$.

Matrix tasks. In all *matrix tasks*, participants were shown a series of 12 three-digit numbers (e.g., 6.41) and had to identify the two numbers that would add up to exactly 10. Honest participants would report the exact number of actually solved matrices. For matrix tasks that measured at the individual level, t reflects the "true" performance per experiment. For matrix tasks that measured at the aggregate level, we assumed that the control group (which could not possibly cheat) and the experimental group (which could cheat) solved an equal number of matrices. Hence, the "true" performance of the control group was assumed to be equal to the "true" performance of the experimental group, t. If, for example, on average $t = 5$ out of 20 matrices were solved, an honest participant sample would also claim an average of $m = 5$ matrices. The minimum is to claim no matrix was solved, $t_{min} = 0$. The maximum is to claim all matrices were solved; for example, $t_{max} = 20$ for a total of 20 matrices. In a study with 20 matrices and $t = 5$, an average claim of $m = 20$ would thus convert to $M_r = +100\%$, an average claim of $m = 5$ would convert to $M_r = 0\%$, and an average claim of $m = 0$ would convert to $M_r = -100\%$.

To calculate the rates of liars we limited the analysis to matrix tasks that measured dishonesty at the individual level and that either reported the rate of liars or for which we had the primary data. This held for only 77 of the 96 matrix tasks. For each experiment, we excluded all participants who had solved (1) all matrices and those who (2) reported having solved *fewer* matrices than was actually the case. These steps were necessary to exclude (1) participants who could not cheat to their own advantage because they had already solved all matrices (i.e., all subjects s_i whose performance p_i was equal to the maximum claim t_{max}); and (2) participants who misreported in such a way that they earned less (i.e., all subjects s_i whose claim c_i was smaller than their performance p_i). We thus defined the eligible sample for calculating the rates of liars per primary study S^* as:

$$S \supseteq S^{*} = \left\{ s_i \,\middle|\, p_i < t_{\max} \cap p_i \le c_i \right\} \tag{A4.2.8}$$

where S denotes all observations per matrix task and S^* is the set of all subjects s_i whose performance p_i was smaller than the maximum possible claim t_{max} and whose claims c_i were at least equal to the number of matrices they had solved.

Accordingly, the sample size n per primary study is the number of eligible observations, which was defined as:

$$n = \left| S^{*} \right|. \tag{A4.2.9}$$

For each primary study, the percentage of honest responses M_{honest} was calculated as the fraction of the sample size n whose claims were equal to their true performance:

$$S^{*} \supseteq honest = \left\{ s_i \,\middle|\, c_i = p_i \right\} \tag{A4.2.10}$$

$$M_{honest} = \frac{\left| honest \right|}{n} \tag{A4.2.11}$$

where the honest reports per primary study, *honest*, are the reports of all participants s_i whose claims c_i equaled the number of matrices they had solved p_i.

It follows that the percentage of honest responses M_{honest} and dishonest responses M_{liars} per experiment add up to 100%. Accordingly:

$$100\% - M_{honest} = M_{liars} \tag{A4.2.12}$$

where M_{honest} is the percentage of honest responses and M_{liars} is the percentage of dishonest responses.

Further remarks. The calculation of the standardized report may cause biased estimates, that is, M_r and $\mathrm{M}_{\mathrm{liars}}$ may systematically over- or underestimate the degree of dishonest responses per experiment. Biases can occur at two levels: *within* and *between* paradigms. *Within*-paradigm bias refers to variation in aspects of the experimental design that were specific to each paradigm. In die-roll tasks, for example, the "true" observed outcomes could be continuously uniform (e.g., the outcome of a one-shot die roll) or normally distributed (e.g., the outcome of two die rolls). We examine the potential reasons for within-paradigm bias in Appendix 4.3. *Between*-paradigm bias can occur due to systematic differences in one or more of the four experimental paradigms that have methodological or theoretical reasons. Methodological reasons pertain to the way in which the standardized report was computed. For example, M_r can take negative values in die-roll, coin-flip, and matrix tasks, whereas in sender–receiver games $M_r = 0\%$ is the minimum. To account for between-paradigm biases, we either analyzed the paradigms separately or used dummy variables to control for experimental paradigm. Theoretical reasons for between-paradigms biases cannot be addressed via statistical methods. For example, the four paradigms may in fact measure different facets of dishonest behavior: standardized reports may represent the sending of false messages (sender–receiver games), the deviation from the expected value of a randomly generated figure (die-roll and coin-flip tasks), or one's actual performance (matrix task).

Monetary units. All monetary values were converted from local currency units to US dollars using 2015 purchasing power parity (PPP). The data for the PPP conversion were taken from the World Economic Outlook Database (International Monetary Fund, 2015). The 2015 PPP conversion rate for Argentine peso was missing from the database; we therefore used its 2011 conversion rate. Two series of experiments used Lindt Lindor chocolate truffles as the incentive for reporting a win in coin-flip tasks (Muñoz-Izquierdo et al., 2014; Pascual-Ezama et al., 2015). In March 2017, a pack of 60 Lindt Lindor chocolate truffles cost US$15.99 on Amazon.com. We therefore assumed that the incentive was worth $0.27 (= $15.99/60). In cases where the exact incentive size was unknown (Gunia et al., 2014) or other nonmonetary incentives were employed (Hilbig & Zettler, 2015; Hildreth et al., 2016, study 2b; Ruffle & Tobol, 2014), we let the incentive size equal the mean incentive of the respective experimental paradigm.

Appendix 4.3

Supplementary Analyses with Paradigm-Specific Features

In this section, we analyze distinct experimental characteristics that are limited to specific experimental paradigms (e.g., variations in the setup practices of sender–receiver games). We also inspect the response distributions in experimental paradigms with continuous outcome measures (i.e., die-roll tasks and matrix tasks).

Sender–receiver games. It is worth noting that the mean rate of 49% false messages may not represent the rate of intentionally misleading messages. In a seminal investigation, Sutter (2009) found that only about 70% of participants in the role of sender believed that their advice would be followed. Almost all of the 30% skeptical participants sent truthful messages, believing that their partners in the role of receivers would *not* follow their advice (sophisticated truth telling); only a minority sent false messages believing that their advice would *not* be followed. In our data set, the rate of advice following among receivers was indeed close to 70% (k = 45, n = 2,619, M_{follow} = 68%). To reduce or even eliminate the possibility of sophisticated truth telling, researchers have commonly used two experimental modifications. Some give more than two options for receivers to choose between in sender–receiver games. This setup makes it more difficult for receivers to find the option with the higher payoff by random guessing. The rationale is that the greater the number of options, the more receivers should trust senders. As a consequence, sophisticated truth tellers will reveal themselves by switching from sending true messages in sender–receiver games with two options to sending false messages in sender–receiver games with three (Erat & Gneezy, 2012; Wang & Murnighan, 2016). A linear mixed effects model with random intercepts between experiments tested whether the number of options in sender–receiver games was associated with more lying. As predicted, every additional option raised the percentage of false messages by 2% (Table A4.3.1). An alternative method to reduce sophisticated truth telling is to inform participants that receivers have already decided to follow their advice (e.g., Cohen et al., 2009). However, we could not confirm that this approach was associated with fewer false messages (Table A4.3.1).

Table A4.3.1

Increasing the Number of Options Increased Misreporting in Sender–Receiver Games

Dependent variable	M_r
Intercept	42.21%*** (2.54)
Number of options	2.23%*** (0.59)
Guaranteed implementation	1.20% (5.24)
Observations	$k = 150$ $n = 7{,}463$
Residual heterogeneity	$I^2 = 87\%$ $\tau^2 = 0.03$
Heterogeneity accounted for	$R^2 = 20\%$

Note. Linear mixed effects model with random intercepts between experiments. Unless denoted otherwise, values refer to regression estimates with standard errors in parentheses.

Adding the number of options to the regression model presented in Table 4.3 as a covariate did not qualitatively change the results, with two notable exceptions: maximal externality (p = .052) and maximal gain (p = .081) fell below conventional levels (Table A4.3.2). The model fit improved only slightly (from R^2 = 14% to R^2 = 19%).

Table A4.3.2

Predictors of the Standardized Report in Sender–Receiver Games: Regression Analyses With and Without the Number of Options as Covariate

Predictor (reference category)	M_r	M_r
Intercept	10.85**** (16.64)	22.87 (16.63)
Investigative setting (laboratory)		
Online/telephone	−2.50% (7.36)	0.09% (7.25)
Field experiment	−7.32% (5.60)	−4.64% (5.52)
Participant characteristics (non-economics students)		
Nonstudents	5.31% (6.72)	5.58% (6.55)
Economics students	1.88% (4.74)	2.39% (4.63)
Primary data shared	0.24% (4.93)	9.24% (5.49)
Year of publication	−0.52% (0.83)	−1.12% (0.83)
Maximal externality	−1.39%* (0.57)	1.09% (0.56)
Maximal gain	1.19%* (0.60)	1.03% (0.59)
Number of options	—	2.47%*** (0.73)
Observations	$k = 150$ $n = 7{,}463$	$k = 150$ $n = 7{,}463$
Residual heterogeneity	$I^2 = 88\%$ $\tau^2 = 0.03$	$I^2 = 87\%$ $\tau^2 = 0.03$
Heterogeneity accounted for	$R^2 = 14\%$	$R^2 = 19\%$

Note. Linear regression models with random effects at the experiment level. Unless denoted otherwise, values refer to beta weights with standard errors in parentheses, with * $p < .05$, ** $p < .01$, *** $p < .001$, and **** $p < .0001$.

Die-roll tasks. In the following, we focus on die-roll tasks in which participants roll the die once. If participants were honest, rolling once would yield continuous uniform distributions of the reported score (see Figure A4.3.1). Participants' actual reporting was, however, far from this. Over-reporting of high scores and underreporting of low scores was common. The highest pip score was reported almost four times as often as the lowest score ($M_{max.\ pip}$ = 31% vs. $M_{min.\ pip}$ = 8%). By contrast, medium scores were reported about as frequently as would be expected from honest reporting. This does not mean, however, that participants who obtained medium scores were necessarily more honest than the rest. The total mean across all observations was largely consistent with "justified dishonesty" theory (Gächter & Schulz, 2016; Shalvi, Dana, et al., 2011), according to which participants respond as if they had rolled twice and then chose to report the larger outcome—although they were unambiguously instructed to report only the outcome of the first roll. According to justified dishonesty theory, even people who obtain medium scores will be willing to over-report if the second roll yields a greater score than the first.

Matrix tasks. It has been suggested that participants in the matrix task inflate their results only to the extent that they can maintain a positive self-concept of being an honest person in general (Mazar et al., 2008). That is, people shy away from reporting that they have solved all matrices (maximal lying) and instead inflate the number of solved matrices to some degree (truth stretching; see also Schweitzer & Hsee, 2002; Tenbrunsel & Messick, 2004). Our results are largely in line with this idea (Figure A4.3.2), although other explanations are possible (see Discussion). The 45% of participants who reported dishonestly in matrix tasks for which we had the primary data on average claimed solving "only" an additional 4 matrices out of an average of 14 unsolved matrices.

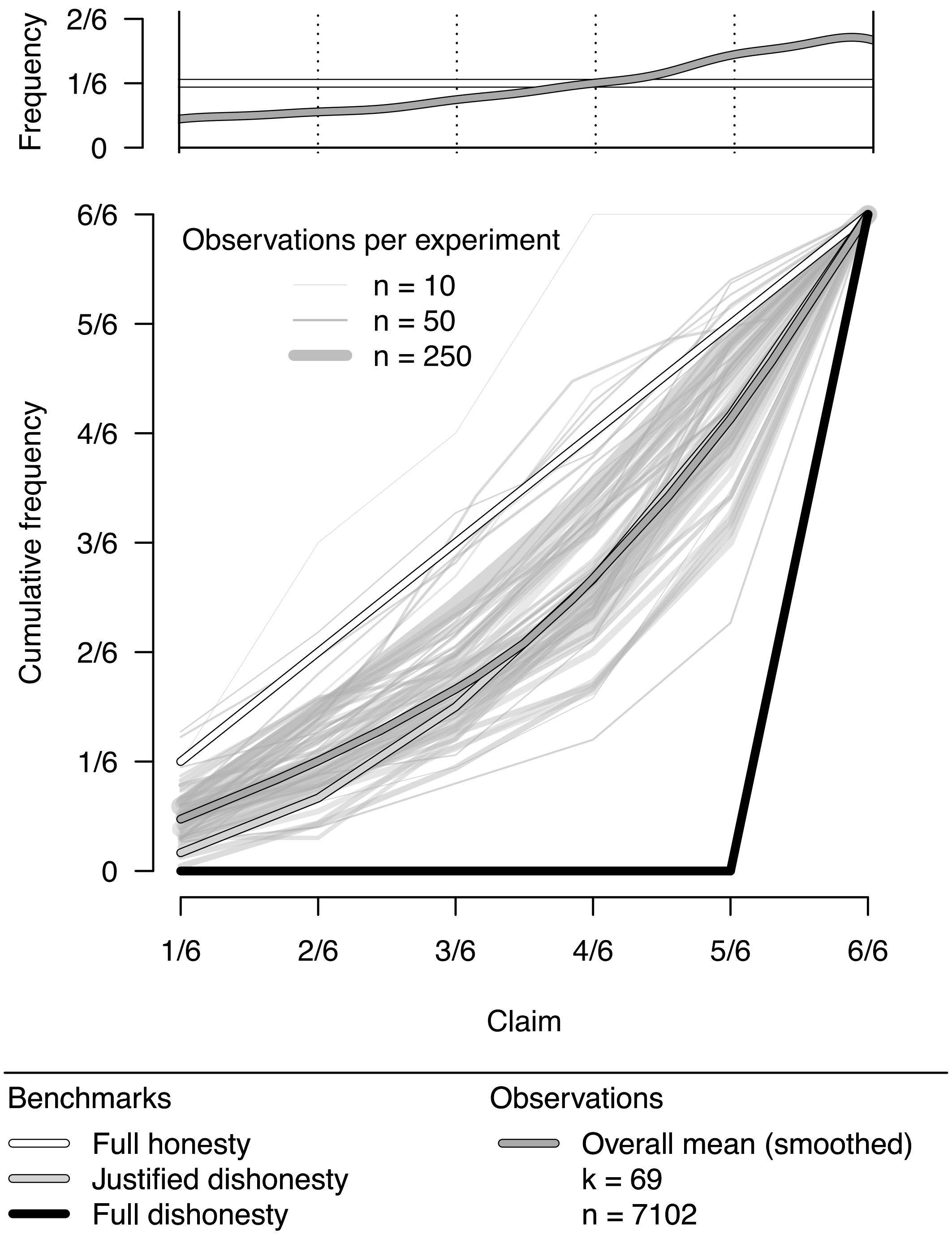


Figure A4.3.1. *Distribution of reported scores in die-roll tasks with a single roll.* The top graph shows the density function of the reported scores. The main graph shows the cumulative distribution functions (CDF) of the reported score. The gray lines depict the CDFs per primary experiment. The more observations per experiment, the wider the gray line. The black line represents the mean CDF for all experiments. The mean CDF was smoothed by local polynomial regression fitting and weighted by the number of observations per reported score. The full dishonesty benchmark (white) depicts the CDF for honest reporting. The justified dishonesty benchmark (gray) represents the CDF for rolling twice and then reporting the greater score of the two rolls, instead of reporting the first score, as instructed. The full dishonesty benchmark (black) depicts the CDF for a sample of maximal liars.

In the main text, the regression model predicting the standardized report in matrix tasks had a relatively poor model fit (R^2 = 20%; Table 4.3), potentially due to the outcome measure itself. A better outcome measure than the standardized report may have been the absolute number of unsolved matrices claimed as solved. In order to validate the regression model presented in Table 4.3, we therefore fitted three additional regression models to matrix tasks by means of a 2×2 design (Table A4.3.3): the dependent variable was either the standardized report (M_r) or the number of unsolved matrices claimed as solved; we either added or removed covariates (including the total number of matrices per task: total matrices; the percentage of matrices without a solution: percent unsolvable; and the time in minutes that participants were allotted for each matrix: time per matrix). Essentially, the results of all four models were qualitatively similar, with one exception: In contrast to the uncontrolled model predicting the standardized report in Table 4.3, field setting reached conventional levels of significance ($p < .001$) once additional covariates were integrated in Table A4.3.3. Overall, the model fits improved only slightly with the additional controls. If anything, predicting the standardized report had a better fit than predicting the number of unsolved matrices claimed as solved. Note that the standardized report and the number of unsolved matrices claimed as solved were highly correlated (Spearman's $\rho = 0.99$, $p < .001$), suggesting that the two measures provided essentially the same assessment of dishonesty.

To further investigate the effects of demographics on dishonest behavior (as reported in the main text), we fitted a number of linear mixed models with random intercepts per experiment to matrix tasks in which the participants' age varied by at least 5 standard deviations ($SD_{age} > 5$). The results presented in Table A4.3.4 indicate that low performers cheated more than high performers did. Every additional matrix solved decreased the absolute number of unsolved matrices claimed as solved by 0.24 and lying by 2.81%. Overall, men solved an average of 1.17 matrices more than women. However, gender did not predict any of the measures of dishonest behavior. By contrast, age negatively predicted three of the four measures: Every year of life lowered the absolute number of unsolved matrices claimed as solved by 0.03 and lying by 0.57%. At the same time, age did not predict performance, suggesting that older participants performed as well as their younger peers did.

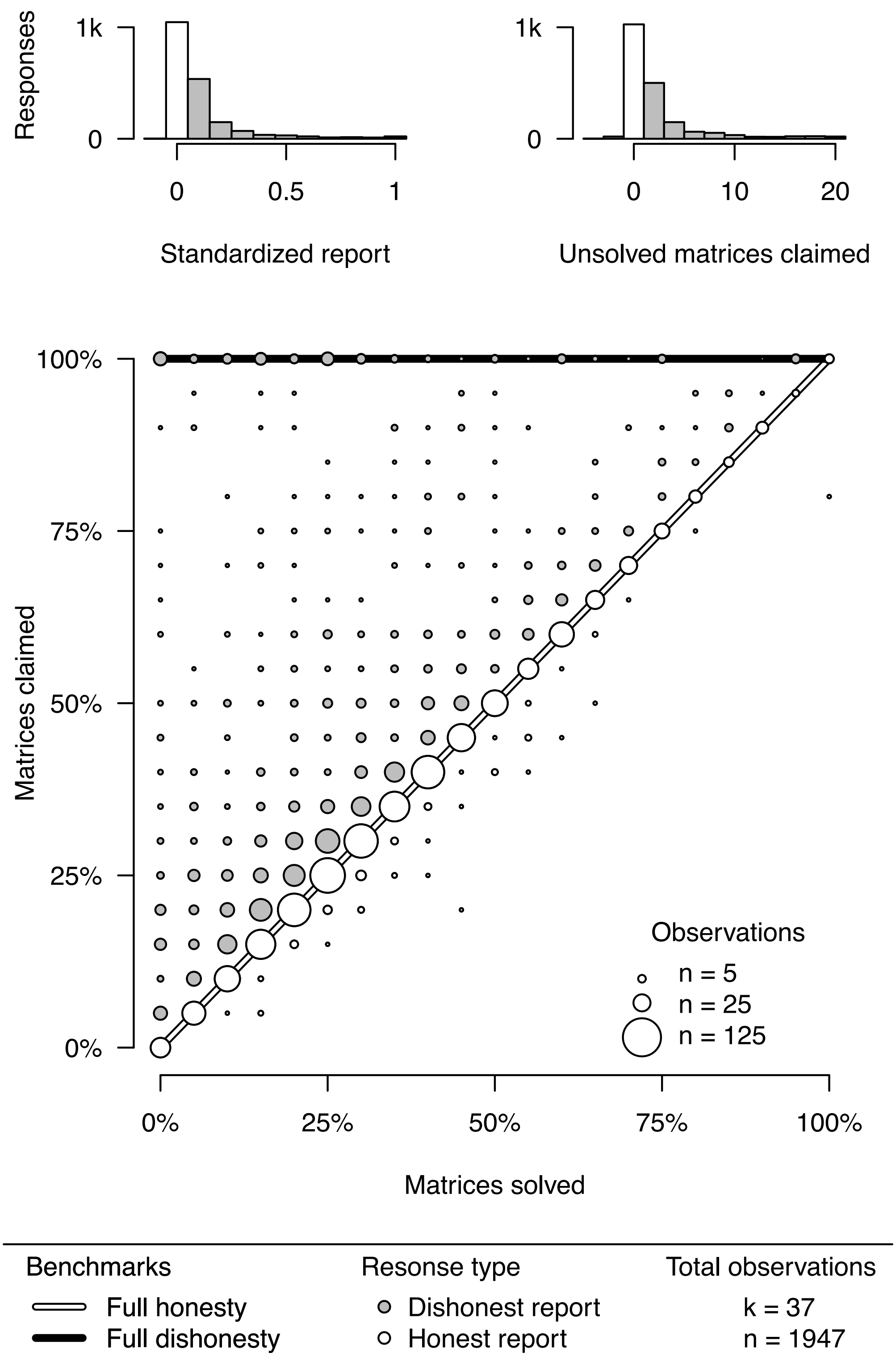


Figure A4.3.2. *Number of claimed and solved matrices in the matrix task.* The main plot depicts the proportion of matrices claimed as solved as a function of the proportion actually solved for experiments that measured performance and claims per participant and for which we had the primary data. Bigger dots stand for more observations. The top-left histogram depicts the distribution of standardized reports. The top-right histogram depicts the number of unsolved matrices claimed as solved—that is, the number of actually solved matrices subtracted from the number of matrices reported as solved.

Table A4.3.3

Predictors of Different Outcome Measures in Matrix Tasks: Regression Analyses for Standardized Report and Absolute Number of Unsolved Matrices Claimed as Solved

Dependent variable (reference category)	M_r	M_r	Unsolved matrices claimed	Unsolved matrices claimed
Intercept	−15.67 (14.32)	−14.78 (13.96)	−114.29 (204.89)	−108.02 (203.27)
Investigative setting (lab)				
Online/telephone	−4.84% (13.71)	−8.21% (13.54)	−0.72 (2.01)	−1.22 (1.98)
Field experiment	−9.47% (7.67)	−37.18%*** (11.05)	−1.85 (2.01)	−5.44*** (1.60)
Participants (non-econ. st.)				
Nonstudents	−3.21% (13.10)	0.25% (12.99)	−0.80 (1.91)	−0.47 (1.89)
Economics students	9.58% (7.47)	−8.38% (9.02)	1.03 (1.03)	−1.16 (1.25)
Primary data shared	−10.16%** (3.78)	−11.63%** (3.75)	−1.11* (0.54)	−1.31* (0.54)
Year of publication	0.79% (0.71)	0.82% (0.69)	0.06 (0.10)	0.06 (0.10)
Maximal externality	0.08% (0.25)	0.07% (0.25)	0.03 (0.04)	0.02 (0.04)
Maximal gain	−0.02% (0.06)	−0.03% (0.06)	−0.00 (0.01)	−0.00 (0.01)
Total matrices	—	−6.87%*** (2.08)	—	−0.62* (0.26)
Percent unsolvable	—	35.86%** (12.73)	—	5.85** (1.84)
Time per matrix	—	−55.08% (50.11)	—	−6.65 (7.08)
Observations	$k = 96$ $n = 5{,}709$	$k = 96$ $n = 5{,}709$	$k = 96$ $n = 5{,}709$	$k = 96$ $n = 5{,}709$
Residual heterogeneity	$I^2 = 76\%$ $\tau^2 = 0.01$	$I^2 = 75\%$ $\tau^2 = 0.01$	$I^2 = 78\%$ $\tau^2 = 2.53$	$I^2 = 76\%$ $\tau^2 = 2.36$
Heterogeneity accounted for	$R^2 = 20\%$	$R^2 = 25\%$	$R^2 = 11\%$	$R^2 = 17\%$

Note. Linear regression models with random effects at the experiment level. Unless denoted otherwise, values refer to beta weights with standard errors in parentheses, with * $p < .05$, ** $p < .01$, *** $p < .001$, and **** $p < .0001$

Table A4.3.4

Effects of Demographics on Performance and Dishonest Behavior in Matrix Tasks

Dependent variable	Matrices solved	Unsolved matrices claimed as solved	M_{liars}
Intercept	5.43**** (0.47)	4.02**** (0.66)	88.47%**** (9.40)
Matrices solved	—	−0.24**** (0.05)	−2.81%** (0.72)
Female	−1.17**** (0.29)	−0.14 (0.34)	2.16% (4.56)
Age	−0.02 (0.01)	−0.03* (0.01)	−0.57%** (0.20)
Observations	$k = 10$ $n = 461$	$k = 10$ $n = 461$	$k = 10$ $n = 461$

Note. Linear mixed models with random effects between the experiments. Unless denoted otherwise, values refer to beta weights with standard errors in parentheses with * $p < .05$; ** $p < .01$; *** $p < .001$; **** $p < .0001$.

Appendix 4.4

Rates of Truth Stretchers and Maximal Liars

In die-roll and matrix tasks, two types of liars can potentially be distinguished: those who inflated their observation/their performance to the maximum possible degree (maximal liars; e.g., people who falsely claimed to have rolled a 6; people who falsely claimed to have solved all matrices) and those whose claims remained below the possible maximum (truth stretchers; e.g., people who reported a 3 instead of the actually observed 2; people who claimed they had solved 6 matrices instead of the actually 4). In this section, we compare liar types and we reveal striking differences between die-roll and matrix tasks. We begin by defining maximal liars and truth stretchers.

In die-roll tasks, the rate of maximal liars M_{max} can be estimated as:

$$M_{\max} = \frac{C\left(t_{\max}\right) - E\left(t_{\max}\right)}{1 - E\left(t_{\max}\right)} \tag{A4.4.1}$$

where $C(t_{max})$ is the percentage of claims of the highest paying option and $E(t_{max})$ is the expected percentage of claims of the highest paying option if everyone was honest. Notably, $C(t_{max}) - E(t_{max})$ has to be divided by $1 - E(t_{max})$ to take into account that some participants did not have the chance to lie maximally because they had already obtained the highest score. For example, in die-roll tasks with a six-sided die and a complete linear payoff function, about $E(t_{max}) = 1/6$ of participants rolled the highest paying option (i.e., a 6) and were not tempted to claim anything but a 6.

A truth stretcher is a liar whose claim remained below the possible maximum. Hence, the rate of truth stretchers M_{ts} can be estimated as:

$$M_{ts} = M_{liars} - M_{\max} \tag{A4.4.2}$$

where M_{liars} is the rate of liars (see Appendix 4.2).

It should be noted that M_{max} in die-roll tasks is an upper bound because M_{max} comprises participants who observed the second highest score (= $t_{max} - 1$), the third highest score (= $t_{max} - 2$), etc. It is impossible to determine, for example, whether a maximal liar who observed the second highest score would also lie maximally if he/she had observed the lowest score.

To calculate M_{max} and M_{ts} for matrix tasks, we limited the dataset to the 36 matrix tasks that measured dishonest behavior at the individual level and for which we had the primary data. Within these experiments, maximal liars are defined as those participants who claimed to have solved all matrices but who actually did not solve all matrices:

$$S^* \supseteq \max = \{ s_i | c_i = t_{\max} \}, \tag{A4.4.3}$$

$$M_{\max} = \frac{|max|}{n}, \tag{A4.4.4}$$

where a maximal liar *max* is a participant s_i whose claim c_i was equal to the maximum possible claim t_{max}. The rate of maximal liars M_{max} is thus the percentage of *max* divided by all eligible participants n (see Appendix 4.2).

The rate of truth stretchers can be calculated as the fraction of all members of S^* who claimed to have solved more matrices than they actually did but less than the maximum number:

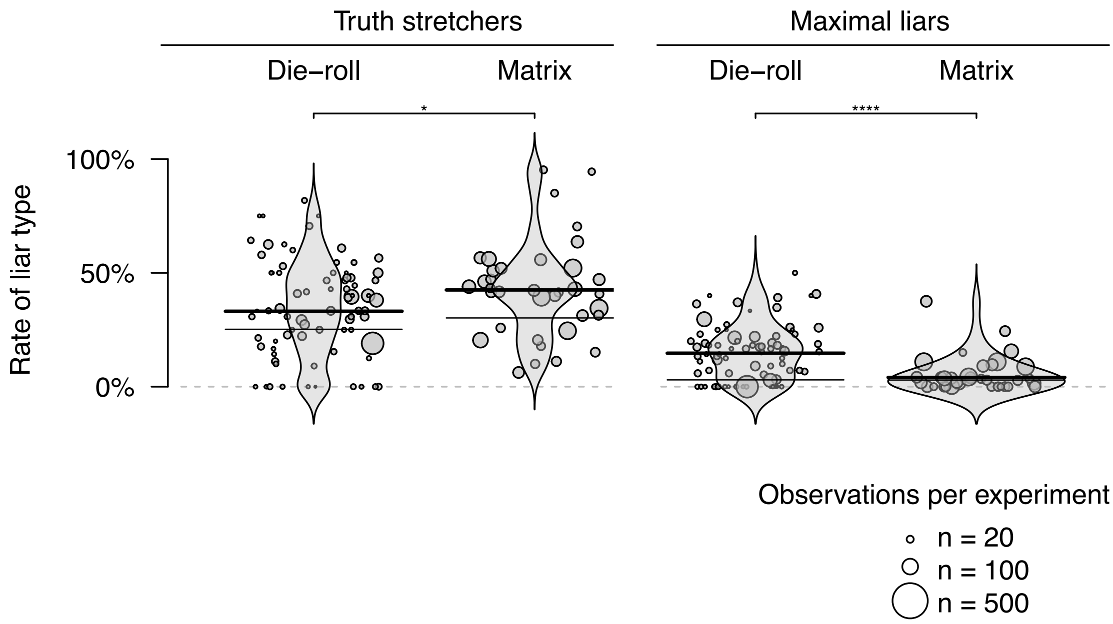
$$S^* \supseteq ts = \{ s_i | p_i < c_i < t_{\max} \}, \tag{A4.4.5}$$

$$M_{ts} = \frac{|ts|}{n}, \tag{A4.4.6}$$

where a truth stretcher *ts* is a participant s_i whose claim c_i was greater than her performance p_i but smaller than the maximum possible claim t_{max}. The rate of truth stretchers M_{ts} is thus the percentage of *ts* divided by all eligible participants n.

Similar to M_{max} in die-roll tasks, M_{max} in matrix tasks represents an upper bound. It is impossible to say if participants who solved all but one matrix (= $p_i + 1 = t_{max}$) or all but two matrices (= $p_i + 2 = t_{max}$), etc., would also lie maximally if they solved fewer matrices.

Differentiating according to liar type (M_{max} and M_{ts}) reveals striking differences between die-roll and matrix tasks. As shown in Figure A4.4.1, there were marked differences in the rates of truth stretchers, $Q(1) = 4.56$, $p = .033$, and maximal liars, $Q(1) = 11.97$, $p < .001$, across the two paradigms. In die-roll tasks, $M_{ts} = 33\%$ of participants stretched the truth and $M_{max} = 15\%$ lied maximally. In matrix tasks, $M_{ts} = 42\%$ stretched the truth but only $M_{max} = 4\%$ lied maximally. Within die-roll tasks and within matrix tasks, the correlation between the rate of truth stretchers and the rate of maximal liars was close to zero, suggesting that maximal lying and truth stretching were relatively independent and neither co-occurred nor replaced each other (die-roll tasks: $\rho = -0.02$, $p = .865$; matrix tasks: $\rho = -0.12$, $p = .469$; Spearman rank correlations per experiment). Trim and fill analysis found some indication of publication bias (Figure A4.4.1).



	Truth stretchers: Die−roll	Truth stretchers: Matrix	Maximal liars: Die−roll	Maximal liars: Matrix
Random effect [95% CI]	M_{ts} = 33% [28%, 38%] z = 13.58 p < .001	M_{ts} = 42% [35%, 50%] z = 11.71 p < .001	M_{max} = 15% [9%, 20%] z = 5.07 p < .001	M_{max} = 4% [2%, 6%] z = 4.07 p < .001
Observations	k = 77 n = 1428	k = 36 n = 1890	k = 77 n = 1428	k = 36 n = 1890
Heterogenity	I^2 = 72% τ^2 = 0.17	I^2 = 92% τ^2 = 0.21	I^2 = 92% τ^2 = 0.22	I^2 = 75% τ^2 = 0.05
Trim and fill [95% CI]	M_{ts} = 25% [20%, 30%] z = 10.21 p < .001	M_{ts} = 30% [22%, 38%] z = 7.51 p < .001	M_{max} = 3% [0%, 6%] z = 1.88 p = .059	M_{max} = 3% [1%, 5%] z = 3.11 p = .002
Added studies	k_{add} = +19	k_{add} = +12	k_{add} = +35	k_{add} = +3

Figure A4.4.1. *Violin plots showing rates of maximal liars and truth stretchers by experimental paradigm.* The bars above the plots summarize the results of tests for subgroup differences, with ** $p < .01$, **** $p < .0001$. See caption to Figure 2 for further explanation.

One reason why more maximal liars were observed in die-roll tasks than in matrix task is because estimating maximal liars in die-roll tasks yields to a greater extend an upper bound than calculating maximal liars in matrix tasks. Another reason could be that there is a greater chance of being eligible for the highest reward in die-roll tasks. As mentioned in the main text, 15 times more people were eligible for the highest reward in die-roll tasks than in matrix tasks ($W = 114$, $p < .001$; one-sided Wilcoxon test). Linear regression analyses indeed showed that eligibility for the highest reward predicted the rate of maximal liars—in die-roll tasks, in matrix tasks, and also when the two tasks were combined (Table A4.4.1). At the same time, eligibility for the highest reward did not predict the rate of liars as such. Taken together, these results suggest that liars shy away from maximal lying when doing so appears implausible. This might explain why die-roll tasks yielded higher rates of maximal liars than matrix tasks did.

Table A4.4.1

Eligibility for the Highest Reward Predicted the Rate of Maximal Liars but not the Rate of Liars

Dependent variable	M_{max}	M_{max}	M_{max}	M_{liars}	M_{liars}	M_{liars}
Paradigm	Both	Die-roll	Matrix	Both	Die-roll	Matrix
Intercept	1.38% (3.35)	3.05% (3.60)	−7.41% (5.19)	54.57%*** (8.05)	55.40%*** (8.56)	30.41%** (10.44)
Percentage eligible for highest reward	76.71%*** (20.90)	66.00%** (21.44)	273.85%** (93.06)	−27.61% (48.42)	−33.12% (51.64)	51.62% (163.56)
Matrix task (dummy)	0.38% (3.82)	—	—	−19.61% (10.65)	—	—
Observations	$k = 83$ $n = 1{,}700$	$k = 77$ $n = 1{,}428$	$k = 6$ $n = 272$	$k = 83$ $n = 1{,}700$	$k = 77$ $n = 1{,}428$	$k = 6$ $n = 272$
Residual heterogeneity	$I^2 = 36\%$, $\tau^2 = 0.00$	$I^2 = 28\%$ $\tau^2 = 0.00$	$I^2 = 74\%$ $\tau^2 = 0.00$	$I^2 = 79\%$ $\tau^2 = 0.04$	$I^2 = 79\%$ $\tau^2 = 0.05$	$I^2 = 72\%$ $\tau^2 = 0.01$
Heterogeneity accounted for	$R^2 = 45\%$	$R^2 = 66\%$	$R^2 = 33\%$	$R^2 = 18\%$	$R^2 = 13\%$	$R^2 = 0\%$

Note. Linear regression models with random effects between the experiments. To prevent floor effects, we limited the analysis of matrix tasks to experiments in which at least one participant solved all matrices. Unless denoted otherwise, values refer to regression weights with standard errors in parentheses and with * $p < .05$, ** $p < .01$, *** $p < .001$.

Bibliography

Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*, 96–104. https://doi.org/10.1016/j.jpubeco.2014.01.005

Abeler, J., Nosenzo, D., & Raymond, C. (2016). *Preferences for truth-telling* (CESIfo Working Paper No. 6087). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2866381

Ahmed, A. M. (2010). What is in a surname? The role of ethnicity in economic decision making. *Applied Economics*, *42*(21), 2715–2723. https://doi.org/10.1080/00036840801964609

Ahmed, A. M., & Hammarstedt, M. (2011). The effect of subtle religious representations on cooperation. *International Journal of Social Economics*, *38*(11), 900–910. https://doi.org/10.1108/03068291111171405

Ahmed, A. M., & Salas, O. (2011). Implicit influences of Christian religious representations on dictator and prisoner's dilemma game decisions. *The Journal of Socio-Economics*, *40*(3), 242–246. https://doi.org/10.1016/j.socec.2010.12.013

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *CXV*(3), 715–753. https://doi.org/10.1162/003355300554881

Alexander, C. N., & Weil, H. G. (1969). Players, persons, and purposes: Situational meaning and the prisoner's dilemma game. *Sociometry*, *32*(2), 121–144. https://doi.org/10.2307/2786258

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*, 464–477. https://doi.org/10.2307/2234133

Andreoni, J. (1995). Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, *CX*(1), 1–21. https://doi.org/10.2307/2118508

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–753. https://doi.org/10.1111/1468-0262.00302

Angelova, V., & Regner, T. (2013). Do voluntary payments to advisors improve the quality of financial advice? An experimental deception game. *Journal of Economic Behavior and Organization*, *93*, 205–218. https://doi.org/10.1016/j.jebo.2013.03.022

Aoki, K., Akai, K., & Onoshiro, K. (2013). *An apology for lying* (ISER Discussion Paper). Retrieved from http://econpapers.repec.org/paper/dprwpaper/0786.htm

Arbel, Y., Bar-El, R., Siniver, E., & Tobol, Y. (2014). Roll a die and tell a lie: What affects honesty? *Journal of Economic Behavior and Organization*, *107*, 153–172. https://doi.org/10.1016/j.jebo.2014.08.009

Bacharach, M. (1999). Interactive team reasoning: A contribution to the theory of co-operation. *Research in Economics*, *53*, 117–147. https://doi.org/10.1006/reec.1999.0188

Bacharach, M. (2006). *Beyond individual choice: Teams and frames in game theory.* (N. Gold & R. Sudgen, Eds.). Princeton, NJ: Princeton University Press.

Bacharach, M., & Bernasconi, M. (1997). The variable frame theory of focal points: An experimental study. *Games and Economic Behavior*, *19*(1), 1–45. https://doi.org/10.1006/game.1997.0546

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin*, *112*(3), 461–484.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes and Intergroup Relations*, *12*(4), 533–547. https://doi.org/10.1177/1368430209105040

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556–1581. https://doi.org/10.1037/a0037737

Bandura, A. (1965). Influence of model's reinforcement contigencies on the acquisition of immitative responses. *Journal of Personality and Social Psychology*, *1*(6), 589–595. https://doi.org/10.1037/h0022070

Banerjee, P., & Chakravarty, S. (2014). Psychological ownership, group affiliation and other-regarding behaviour: Some evidence from dictator games. *Global Economics and Management Review*, *19*(1–2), 3–15. https://doi.org/10.1016/j.gemrev.2014.12.001

Bardsley, N. (2008). Dictator game giving: Altruism or artefact? *Experimental Economics*, *11*(2), 122–133. https://doi.org/10.1007/s10683-007-9172-2

Bardsley, N. (2010). Sociality and external validity in experimental economics. *Mind and Society*, *9*, 119–138. https://doi.org/10.1007/s11299-010-0075-0

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*(6), 1014–1027. https://doi.org/10.1037//0022-3514.81.6.1014

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, *72*(6), 1335–1348. https://doi.org/10.1037/0022-3514.72.6.1335

Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology*, *29*, 909–924. https://doi.org/10.1002/(SICI)1099-0992(199911)29:73.0.CO;2-L

Battigalli, P., & Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, *144*(1), 1–35. https://doi.org/10.1016/j.jet.2008.01.004

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, *76*(2), 169–217. https://doi.org/10.1007/978-1-349-62853-7_2

Benjamin, D. J., Choi, J. J., & Fisher, G. (2013). *Religious identity and economic behavior* (Working paper). https://doi.org/10.3386/w15925

Berg, J., Dickhaut, J., & McCabe, K. A. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*, 122–142. https://doi.org/10.1006/game.1995.1027

Berger, P. L., & Luckmann, T. (1966). *The social construction of reality. A treatise in the* *sociology of knowledge*. London, United Kingdom: Penguin.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912–915. https://doi.org/10.1038/nature04981

Bernold, E., Gsottbauer, E., Ackermann, K. A., & Murphy, R. O. (2015). *Social framing and cooperation: The roles and interaction of preferences and beliefs* (Available at SSRN). https://doi.org/10.2139/ssrn.2557927

Bettenhausen, K. L., & Murnighan, J. K. (1991). The development of an intragroup norm and the effects of interpersonal and structural challenges. *Administrative Science Quarterly*, *36*(1), 20–35. https://doi.org/10.2307/2393428

Bicchieri, C. (1990). Norms of cooperation. *Ethics*, *100*(4), 838–861. https://doi.org/10.1086/293237

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge, United Kingdom: Cambridge University Press.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making*, *22*, 191–208. https://doi.org/10.1002/bdm.621

Bicchieri, C., & Zhang, J. (2012). An embarrassment of riches: Modeling social preferences in ultimatum games. In U. Mäki (Ed.), *Philosophy of Economics* (pp. 577–595). San Diego, CA: North Holland. https://doi.org/10.1016/B978-0-444-51676-3.50020-8

Biel, A., & Thøgersen, J. (2007). Activation of social norms in social dilemmas: A review of the evidence and reflections on the implications for environmental behaviour. *Journal of Economic Psychology*, *28*(1), 93–112. https://doi.org/10.1016/j.joep.2006.03.003

Binmore, K. G. (1999). Why experiment in economics? *The Economic Journal*, *109*(453), F16–F24. https://doi.org/10.1111/1468-0297.00399

Bischoff, I., & Frank, B. (2011). Good news for experimenters: Subjects are hard to influence by instructors' cues. *Economics Bulletin*, *31*(4), 3221–3225. https://doi.org/10.1016/s0165-1765(98)00162-1

Biziou-van-Pol, L., Haenen, J., Novaro, A., Liberman, A. O., & Capraro, V. (2015). Does telling white lies signal pro-social preferences? *Judgment and Decision Making*, *10*(6), 538–548.

Blais, A., & Young, R. (1999). Why do people vote? An experiment in rationality. *Public Choice*, *99*(1), 39–55. https://doi.org/10.1023/A:1018341418956

Bohnet, I., & Frey, B. S. (1999). The sound of silence in prisoner's dilemma and dictator games. *Journal of Economic Behavior and Organization*, *38*(1), 43–57. https://doi.org/10.1016/S0167-2681(98)00121-8

Bolton, G. E. (1991). A comparative model of bargaining: Theory and evidence. *The American Economic Review*, *81*(5), 1096–1136.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166–193. https://doi.org/10.1257/aer.90.1.166

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, United Kingdom: John Wiley & Sons.

Bosch-Domènech, A., & Silvestre, J. (2015). *The role of frames, numbers and risk in the* *frequency of cooperation* (Working paper). Retrieved from http://www.econ.upf.edu/docs/papers/downloads/1501.pdf

Bouma, J. A., & Ansink, E. (2013). The role of legitimacy perceptions in self-restricted resource use: A framed field experiment. *Forest Policy and Economics*, *37*, 84–93. https://doi.org/10.1016/j.forpol.2013.01.006

Bouma, J. A., Joy, K. J., Paranjape, S., & Ansink, E. (2014). The influence of legitimacy perceptions on cooperation: A framed field experiment. *World Development*, *57*, 127–137. https://doi.org/10.1016/j.worlddev.2013.12.007

Brañas-Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology*, *28*(4), 477–486. https://doi.org/10.1016/j.joep.2006.10.001

Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovářík, J., & Ponti, G. (2010). Altruism and social integration. *Games and Economic Behavior*, *69*(2), 249–257. https://doi.org/10.1016/j.geb.2009.10.014

Brandts, J., & Schwieren, C. (2009). *Frames and economic behavior: An experimental study* (Working paper). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=26685F8CF1BD9EC8DBB28DCF91593553?doi=10.1.1.533.4454&rep=rep1&type=pdf

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology*, *50*(3), 543–549. https://doi.org/10.1037/0022-3514.50.3.543

Browning, C. R. (2001). *Ordinary men. Reserve police battalion 101 and the final solution in Poland*. London, United Kingdom: Penguin.

Bruner, J. S. (1957). Going beyond the information given. In J.S. Bruner, E. Brunswik, L. Festinger, F. Heider, K. F. Muenzinger, C. E. Osgood, & D. Rapaport (Eds.), *Contemporary approaches to cognition* (pp. 41–69). Cambridge, MA: Harvard University Press.

Brunk, G. G. (1980). The impact of rational participation models on voting attitudes. *Public Choice*, *35*(5), 549–564. https://doi.org/10.1007/BF00140085

Brunswik, E. (1943). Organismic achievement and environmental probability. *The Psychological Review*, *50*(3), 255–272. https://doi.org/10.1037/h0060889

Brunswik, E. (1944). Distal focussing of perception: Size-constancy in a representative sample of situations. *Psychological Monographs*, *56*(1), 1–49. https://doi.org/10.1037/h0093505

Bucciol, A., & Piovesan, M. (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology*, *32*(1), 73–78. https://doi.org/10.1016/j.joep.2010.12.001

Buchan, N. R., Croson, R. T. A., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the investment game. *Journal of Economic Behavior & Organization*, *68*, 466–476. https://doi.org/10.1016/j.jebo.2007.10.006

Burks, S. V, & Krupka, E. L. (2012). A multimethod approach to identifying norms and normative expectations within a corporate hierarchy. Evidence from the financial services industry. *Management Science*, *58*(1), 203–217. https://doi.org/10.1287/mnsc.1110.1478

Butler, D. J., Burbank, V. K., & Chisholm, J. S. (2011). The frames behind the games: Player's perceptions of prisoners dilemma, chicken, dictator, and ultimatum games. *The Journal of Socio-Economics*, *40*(2), 103–114. https://doi.org/10.1016/j.socec.2010.12.009

Cadsby, C. B., Du, N., & Song, F. (2016). In-group favoritism and moral decision-making. *Journal of Economic Behavior and Organization*, *128*, 59–71. https://doi.org/10.1016/j.jebo.2016.05.008

Cadsby, C. B., & Maynes, E. (1998). Choosing between a socially efficient and free-riding equilibrium: Nurses versus economics and business students. *Journal of Economic Behavior and Organization*, *37*(2), 183–192. https://doi.org/10.1016/S0167-2681(98)00083-3

Cai, W., Huang, X., Wu, S., & Kou, Y. (2015). Dishonest behavior is not affected by an image of watching eyes. *Evolution and Human Behavior*, *36*(2), 110–116. https://doi.org/10.1016/j.evolhumbehav.2014.09.007

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Camerer, C. F., & Fehr, E. (2002). Measuring social norms and preferences using experimental games: A guide for social scientists. In J. Henrich, R. T. Boyd, S. Bowles, C. F. Camerer, E. Fehr, H. Gintis, & R. McElreath (Eds.), *Foundations of human sociality: Experimental and ethnographic evidence from 15 small-scale societies* (pp. 55–95). New York, NY: Oxford University Press.

Camerer, C. F., & Thaler, R. H. (1995). Anomalies: Ultimatums, dictators and manners. *Journal of Economic Perspectives*, *9*(2), 209–219. https://doi.org/10.1257/jep.9.2.209

Cappelen, A. W., Nygaard, K., Sørensen, E. Ø., & Tungodden, B. (2015). Social Preferences in the Lab: A Comparison of Students and a Representative Population. *The Scandinavian Journal of Economics*, *117*(4), 1306–1326. https://doi.org/10.1111/sjoe.12114

Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2013). When do we lie? *Journal of Economic Behavior and Organization*, *93*, 258–265. https://doi.org/10.1016/j.jebo.2013.03.037

Capraro, V., Smyth, C., Mylona, K., & Niblo, G. A. (2014). Benevolent characteristics promote cooperative behaviour among humans. *PLOS ONE*, *9*(8), e102881. https://doi.org/10.1371/journal.pone.0102881

Carpenter, J. P., Burks, S., & Verhoogen, E. (2005). Comparing students to workers: The effect of social framing on behavior in distribution games. *Research in Experimental Economics*, *10*, 261–290. https://doi.org/10.1016/S0193-2306(04)10007-0

Carpenter, J., Rücker, G., & Schwarzer, G. (2009). copas. An R package for fitting the copas selection model. *The R Journal*, *1*(2), 31–36.

Carter, J. R., & Irons, M. D. (1991). Are economists different, and if so, why? *Journal of Economic Perspectives*, *5*(2), 171–177. https://doi.org/10.1257/jep.5.2.171

Chen, Y., Li, S., Liu, T., & Shih, M. (2014). Which hat to wear? Impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, *84*, 58–86. https://doi.org/10.1016/j.geb.2013.12.002

Cherry, T. L. (2001). Mental accounting and other-regarding behavior: Evidence from the lab. *Journal of Economic Psychology, 22*(5), 605–615. https://doi.org/10.1016/S0167-4870(01)00058-7

Childs, J. (2012a). Demonstrating the need for effective business ethics. An alternative approach. *Business and Society Review, 117*(2), 221–232. https://doi.org/10.1111/j.1467-8594.2012.00406.x

Childs, J. (2012b). Gender differences in lying. *Economics Letters, 114*(2), 147–149. https://doi.org/10.1016/j.econlet.2011.10.006

Childs, J. (2013). Personal characteristics and lying. An experimental investigation. *Economics Letters, 121*(3), 425–427. https://doi.org/10.1016/j.econlet.2013.09.005

Chou, E. Y. (2015). What's in a name? The toll e-signatures take on individual honesty. *Journal of Experimental Social Psychology, 61*, 84–95. https://doi.org/10.1016/j.jesp.2015.07.010

Clot, S., Grolleau, G., & Ibanez, L. (2014). Smug alert! Exploring self-licensing behavior in a cheating game. *Economics Letters, 123*(2), 191–194. https://doi.org/10.1016/j.econlet.2014.01.039

Cohen, T. R., Gunia, B. C., Kim-Jun, S. Y., & Murnighan, J. K. (2009). Do groups lie more than individuals? Honesty and deception as a function of strategic self-interest. *Journal of Experimental Social Psychology, 45*(6), 1321–1324. https://doi.org/10.1016/j.jesp.2009.08.007

Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP scale. A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology, 100*(5), 947–966. https://doi.org/10.1037/a0022641

Cohn, A., Maréchal, M. A., & Noll, T. (2015). Bad boys: How criminal identity salience affects rule violation. *Review of Economic Studies, 82*(4), 1289–1308. https://doi.org/10.1093/restud/rdv025

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge, MA: Harvard University Press. https://doi.org/10.2307/2579680

Cone, J., & Rand, D. G. (2014). Time pressure increases cooperation in competitively framed social dilemmas. *PLOS ONE, 9*(12), e115756. https://doi.org/10.1371/journal.pone.0115756

Conrads, J., Ellenberger, M., Irlenbusch, B., Ohms, E. N., Rilke, R. M., & Walkowitz, G. (2016). Team goal incentives and individual lying behavior. *Die Betriebswirtschaft, 76*(1), 103–123.

Conrads, J., Irlenbusch, B., Rilke, R. M., Schielke, A., & Walkowitz, G. (2014). Honesty in tournaments. *Economics Letters, 123*(1), 90–93. https://doi.org/10.1016/j.econlet.2014.01.026

Conrads, J., Irlenbusch, B., Rilke, R. M., & Walkowitz, G. (2013). Lying and team incentives. *Journal of Economic Psychology, 34*, 1–7. https://doi.org/10.1016/j.joep.2012.10.011

Conrads, J., & Lotz, S. (2015). The effect of communication channels on dishonest behavior. *Journal of Behavioral and Experimental Economics, 58*, 88–93. https://doi.org/10.1016/j.socec.2015.06.006

Cooper, D. J., & Kagel, J. H. (2003). The impact of meaningful context on strategic play in signaling games. *Journal of Economic Behavior and Organization, 50*(3), 311–337.

https://doi.org/10.1016/S0167-2681(02)00025-2

Copas, J. B., & Shi, J. Q. (2000). Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics*, *1*(3), 247–262. https://doi.org/10.1093/biostatistics/1.3.247

Copas, J. B., & Shi, J. Q. (2001). A sensitivity analysis for publication bias in systematic reviews. *Statistical Methods in Medical Research*, *10*(1), 251–265. https://doi.org/10.1177/096228020101000402

Cronk, L. (2007). The influence of cultural framing on play in the trust game: A Maasai example. *Evolution and Human Behavior*, *28*(5), 352–358. https://doi.org/10.1016/j.evolhumbehav.2007.05.006

Cubitt, R. P., Drouvelis, M., & Gächter, S. (2011). Framing and free riding: Emotional responses and punishment in social dilemma games. *Experimental Economics*, *14*(2), 254–272. https://doi.org/10.1007/s10683-010-9266-0

Dai, Z., Galeotti, F., & Villeval, M. C. (2016). Cheating in the lab predicts fraud in the field. An experiment in public transportations. *Management Science*, Advance online publication. https://doi.org/10.1287/mnsc.2016.2616

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. https://doi.org/10.1007/s00199-006-0153-z

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies. Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4), 377–383. https://doi.org/10.1037/h0025589

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193. https://doi.org/10.1146/annurev.ps.31.020180.001125

de Haan, T., & van Veldhuizen, R. (2015). Willpower depletion and framing effects. *Journal of Economic Behavior and Organization*, *117*, 47–61. https://doi.org/10.1016/j.jebo.2015.06.002

de Waal, F. (2005). *Our inner ape. A leading primatologist explains why we are who we are*. New York, NY: Riverhead Trade.

Dehue, F. M. J., McClintock, C. G., & Liebrand, W. B. G. (1993). Social value related response latencies: Unobtrusive evidence for individual differences in information processing. *European Journal of Social Psychology*, *23*, 273–293. https://doi.org/10.1002/ejsp.2420230305

DeScioli, P., & Krishna, S. (2013). Giving to whom? Altruism in different types of relationships. *Journal of Economic Psychology*, *34*(2013), 218–228. https://doi.org/10.1016/j.joep.2012.10.003

Deutsch, M. (1957). *Conditions affecting cooperation*. Washington, DC: Office of Naval Research.

Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, *2*(4), 265–279. https://doi.org/10.1177/002200275800200401

Deutsch, M. (1960). The effect of motivational orientation upon trust and suspicion. *Human Relations*, *13*(2), 123–139. https://doi.org/10.1177/001872676001300202

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, *130*(6), 959–988. https://doi.org/10.1037/0033-2909.130.6.959

Dieckmann, A., Grimm, V., Unfried, M., Utikal, V., & Valmasoni, L. (2016). On trust in honesty and volunteering among Europeans. Cross-country evidence on perceptions and behavior. *European Economic Review*, *19*, 225–253. https://doi.org/10.1016/j.euroecorev.2016.01.011

Douglas, M. (1970). *Natural symbols: Explorations in cosmology*. London, United Kingdom: Barrie & Rockliff/The Cresset Press.

Dreber, A., Ellingsen, T., Johannesson, M., & Rand, D. G. (2013). Do people care about social context? Framing effects in dictator games. *Experimental Economics*, *16*(3), 349–371. https://doi.org/10.1007/s10683-012-9341-9

Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters*, *99*(1), 197–199. https://doi.org/10.1016/j.econlet.2007.06.027

Drouvelis, M., Metcalfe, R., & Powdthavee, N. (2015). Can priming cooperation increase public good contributions? *Theory and Decision*, *79*(3), 479–492. https://doi.org/10.1007/s11238-015-9481-4

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2015). Social class, power, and selfishness: When and why upper and lower class individuals behave unethically. *Journal of Personality and Social Psychology*, *108*(3), 436–449. https://doi.org/10.1037/pspi0000008

Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, *73*(2), 459–478. https://doi.org/10.1016/j.geb.2011.02.003

Duval, S., & Tweedie, R. (2000). A non-parametric "trim and fill" method of assessing publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

Eiser, J. R., & Bhavnani, K.-K. (1974). The effect of situational meaning on the behaviour of subjects in the prisoner's dilemma game. *European Journal of Social Psychology*, *4*(1), 93–97. https://doi.org/10.1002/ejsp.2420040108

Ellingsen, T., Johannesson, M., Mollerstrom, J., & Munkhammar, S. (2012). Social framing effects: Preferences or beliefs? *Games and Economic Behavior*, *76*(1), 117–130. https://doi.org/10.1016/j.geb.2012.05.007

Elliott, C. S., Hayward, D. M., & Canon, S. (1998). Institutional framing: Some experimental evidence. *Journal of Economic Behavior and Organization*, *35*(4), 455–464. https://doi.org/10.1016/S0167-2681(98)00047-X

Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, *3*(4), 99–117. https://doi.org/10.1257/jep.3.4.99

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7

Engel, C., & Rand, D. G. (2014). What does "clean" really mean? The implicit framing of decontextualized experiments. *Economics Letters*, *122*(3), 386–389. https://doi.org/10.1016/j.econlet.2013.12.020

Erat, S. (2013). Avoiding lying. The case of delegated deception. *Journal of Economic Behavior and Organization*, *93*, 273–278. https://doi.org/10.1016/j.jebo.2013.03.035

Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, *58*(4), 723–733. https://doi.org/10.1287/mnsc.1110.1449

Eriksson, K., & Strimling, P. (2014). Spontaneous associations and label framing have similar effects in the public goods game. *Judgement and Decision Making*, *9*(5), 360–372. https://doi.org/10.1007/bf01669207

Eriksson, K., Strimling, P., & Coultas, J. C. (2014). Bidirectional associations between descriptive and injunctive norms. *Organizational Behavior and Human Decision Processes*, *129*, 59–69. https://doi.org/10.1016/j.obhdp.2014.09.011

"Ex-PM Olmert released from prison after 16 months." (2017, July 2). The Times of Israel. Retrieved from http://www.timesofisrael.com/ex-pm-olmert-to-go-free-after-16-months-in-prison/

Exadaktylos, F., Espín, A., & Brañas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports*, *3*, 1213. https://doi.org/10.1038/srep01213

Faravelli, M., Friesen, L., & Gangadharan, L. (2015). Selection, tournaments, and dishonesty. *Journal of Economic Behavior and Organization*, *110*, 160–175. https://doi.org/10.1016/j.jebo.2014.10.019

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2004), 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, *13*(1), 1–25. https://doi.org/10.1007/s12110-002-1012-7

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140. https://doi.org/10.1038/415137a

Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism: Experimental evidence and new theories. In S. C. Kolm & J. M. Ythier (Eds.), *Handbook of the economics of giving, altruism and reciprocity* (Vol. 1, pp. 615–691). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S1574-0714(06)01008-6

Ferraro, F., Pfeffer, J., & Sutton, R. (2005). Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review*, *30*(1), 8–24. https://doi.org/10.5465/AMR.2005.15281412

Fiddick, L., & Cummins, D. (2007). Are perceptions of fairness relationship-specific? The case of noblesse oblige. *Quarterly Journal of Experimental Psychology*, *60*(1), 16–31. https://doi.org/10.1080/17470210600577266

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise. An experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547. https://doi.org/10.1111/jeea.12014

Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods games. *American Economic Review*, *100*(1), 541–556. https://doi.org/10.1257/aer.100.1.541

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404. https://doi.org/10.1016/S0165-1765(01)00394-9

Fischbacher, U., & Heusi, F. (2008). *Lies in disguise. An experimental study on cheating* (TWI Research Paper). Retrieved from http://econpapers.repec.org/paper/twirespas/0040.htm

Fiske, A. P. (1991a). *Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing*. New York, NY: Free Press.

Fiske, A. P. (1991b). The cultural relativity of selfish individualism: Anthropological evidence that humans are inherently sociable. In M. S. Clark (Ed.), *Prosocial Behavior. Review of Personality and Social Psychology* (pp. 176–214). Thousand Oaks, CA: Sage.

Fiske, A. P., & Tetlock, P. E. (1997). Taboo trade-offs: Reactions to transactions that transgress the spheres of justice. *Political Psychology*, *18*(2), 255–297. https://doi.org/10.1111/0162-895x.00058

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill.

Forsythe, R., Horowitz, J., Savin, N., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*(3), 347–369. https://doi.org/10.1006/game.1994.1021

Fosgaard, T. R. (2016). *Students cheat more. Comparing dishonesty of a student and a representative sample in the laboratory. Manuscript in preparation*.

Fosgaard, T. R., Hansen, L. G., & Piovesan, M. (2013). Separating will from grace. An experiment on conformity and awareness in cheating. *Journal of Economic Behavior and Organization*, *93*, 279–284. https://doi.org/10.1016/j.jebo.2013.03.027

Frank, R. H., Gilovich, T., & Regan, D. T. (1993). Does studying economics inhibit cooperation? *Journal of Economic Perspectives*, *7*(2), 159–171. https://doi.org/10.1257/jep.7.2.159

Frank, R. H., Gilovich, T., & Regan, D. T. (1996). Do economists make bad citizens? *Journal of Economic Perspectives*, *10*(1), 187–192. https://doi.org/10.1257/jep.10.1.187

Frey, B. S., Pommerehne, W. W., & Gygi, B. (1993). Economic indoctrination or selection? Some empirical results. *The Journal of Economic Education*, *24*(3), 271–281. https://doi.org/10.2307/1183127

Friedman, D., & Sunder, S. (1994). *Experimental methods: A primer for economists*. Cambridge, United Kingdom: Cambridge University Press.

Friesen, L., & Gangadharan, L. (2012). Individual level evidence of dishonesty and the gender effect. *Economics Letters*, *117*(3), 624–626. https://doi.org/10.1016/j.econlet.2012.08.005

Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, *531*(7595), 1–11. https://doi.org/10.1038/nature17160

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453–466. https://doi.org/10.1037/0022-3514.85.3.453

Gamliel, E., & Peer, E. (2013). Explicit risk of getting caught does not affect unethical behavior. *Journal of Applied Social Psychology*, *43*(6), 1281–1288. https://doi.org/10.1111/jasp.12091

Gerkey, D. (2013). Cooperation in context: Public goods games and post-Soviet collectives in Kamchatka, Russia. *Current Anthropology*, *54*(2), 144–176.

https://doi.org/10.1086/669856

Gerlach, P. (2017). The games economists play: Why economics students behave more selfishly than other students. *PLOS ONE*, *12*(9), e0183814. https://doi.org/https://doi.org/10.1371/journal. pone.0183814

Gerlach, P., & Jaeger, B. (2016). Another frame, another game? Explaining framing effects in economic games. In *Norms, Actions, Games* (pp. 1–10). Toulouse: Institute for Advanced Studies.

Gerlach, P., Jaeger, B., & Hertwig, R. (2017). Cooperation needs interpretation. A meta-analysis of context frames in social dilemma games. *Manuscript in Preparation*.

Gerlach, P., Teodorescu, K., & Hertwig, R. (2017). The truth about lies. A meta-analysis on dishonest behavior. *Manuscript in Preparation*.

Gigerenzer, G., & McElreath, R. (2003). Social intelligence in games: Comment. *Journal of Institutional and Theoretical Economics*, *159*, 188–194. https://doi.org/10.1628/0932456032975050

Gino, F. (2015). Understanding ordinary unethical behavior. Why people who value morality act immorally. *Current Opinion in Behavioral Sciences*, *3*, 107–111. https://doi.org/10.1016/j.cobeha.2015.03.001

Gino, F., & Ariely, D. (2011). The dark side of creativity. Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, *102*(3), 445–459. https://doi.org/10.1037/a0026406

Gino, F., & Ariely, D. (2016). Dishonesty explained. What leads moral people act immorally. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 322–344). New York, NY: Guilford Press.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior. The effect of one bad apple on the barrel. *Psychological Science*, *20*(3), 393–8. https://doi.org/10.1111/j.1467-9280.2009.02306.x

Gino, F., Ayal, S., & Ariely, D. (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of Economic Behavior and Organization*, *93*, 285–292. https://doi.org/10.1016/j.jebo.2013.04.005

Gino, F., & Galinsky, A. D. (2012). Vicarious dishonesty. When psychological closeness creates distance from one's moral compass. *Organizational Behavior and Human Decision Processes*, *119*(1), 15–26. https://doi.org/10.1016/j.obhdp.2012.03.011

Gino, F., Krupka, E. L., & Weber, R. A. (2013). License to cheat. Voluntary regulation and ethical behavior. *Management Science*, *59*(10), 2187–2203. https://doi.org/10.1287/mnsc.1120.1699

Gino, F., & Margolis, J. D. (2011). Bringing ethics into focus. How regulatory focus and risk preferences influence (un)ethical behavior. *Organizational Behavior and Human Decision Processes*, *115*(2), 145–156. https://doi.org/10.1016/j.obhdp.2011.01.006

Gino, F., & Mogilner, C. (2014). Time, money, and morality. *Psychological Science*, *25*(2), 414–421. https://doi.org/10.1177/0956797613506438

Gino, F., Norton, M. I., & Ariely, D. (2010). The counterfeit self. The deceptive costs of faking it. *Psychological Science*, *20*(10), 1–9. https://doi.org/10.1177/0956797610366545

Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation. How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, *115*(2), 191–203. https://doi.org/10.1016/j.obhdp.2011.03.001

Gino, F., & Wiltermuth, S. S. (2014). Evil genius? How dishonesty can lead to greater creativity. *Psychological Science*, *25*(4), 973–981. https://doi.org/10.1177/0956797614520714

Gintis, H. (2007). A framework for the unification of the behavioral sciences. *The Behavioral and Brain Sciences*, *30*(1), 1-16-61. https://doi.org/10.1017/S0140525X07000581

Glätzle-Rützler, D., & Lergetporer, P. (2015). Lying and age. An experimental study. *Journal of Economic Psychology*, *46*, 12–25. https://doi.org/10.1016/j.joep.2014.11.002

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384–394. https://doi.org/10.1257/0002828053828662

Gold, N., & Sugden, R. (2007a). Collective intentions and team agency. *The Journal of Philosophy*, *104*(3), 109–137.

Gold, N., & Sugden, R. (2007b). Theories of team agency. In F. Peter & H. B. Schmid (Eds.), *Rationality and commitment* (pp. 280–312). Oxford, United Kingdom: Oxford University Press.

Gomes, C. M., & McCullough, M. E. (2015). The effects of implicit religious primes on dictator game allocations: A preregistered replication experiment. *Journal of Experimental Psychology: General*, *144*(6), e94–e104. https://doi.org/10.1037/xge0000027

Gravert, C. (2013). How luck and performance affect stealing. *Journal of Economic Behavior and Organization*, *93*, 301–304. https://doi.org/10.1016/j.jebo.2013.03.026

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12506–12511. https://doi.org/10.1073/pnas.1000505107

Grimm, V., Utikal, V., & Valmasoni, L. (2015). *In-group favoritism and discrimination among multiple out-groups* (Discussion Paper No. 05/2015). Retrieved from https://ideas.repec.org/p/zbw/iwqwdp/052015.html

Grinberg, M., Hristova, E., & Borisova, M. (2012). Cooperation in prisoner's dilemma game: Influence of social relations. *Proceedings of CogSci 2012*, 408–413.

Grolleau, G., Kocher, M. G., & Sutan, A. (2016). Cheating and loss aversion. Do people cheat more to avoid a loss? *Management Science*, *62*(12), 3428–3438. https://doi.org/10.1287/mnsc.2015.2313

Gu, J., Zhong, C.-B., & Page-Gould, E. (2013). Listen to your heart. When false somatic feedback shapes moral behavior. *Journal of Experimental Psychology*, *142*(2), 307–312. https://doi.org/10.1037/a0029549

Gunia, B. C., Barnes, C. M., & Sah, S. (2014). The morality of larks and owls. Unethical behavior depends on chronotype as well as time of day. *Psychological Science*, *25*(12), 2272–2274. https://doi.org/10.1177/0956797614541989

Gunia, B. C., Wang, L., Huang, L., Wang, J., & Murnighan, J. K. (2012). Contemplation and conversation. Subtle influences on moral decision making. *Academy of Management Journal*, *55*(1), 13–33. https://doi.org/10.5465/amj.2009.0873

Gylfason, H. F., Arnardottir, A. A., & Kristinsson, K. (2013). More on gender differences in lying. *Economics Letters*, *119*(1), 94–96. https://doi.org/10.1016/j.econlet.2013.01.027

Gylfason, H. F., Halldorsson, F., & Kristinsson, K. (2016). Personality in Gneezy's cheap talk game. The interaction between honesty-humility and extraversion in predicting deceptive behavior. *Personality and Individual Differences*, *96*, 222–226. https://doi.org/10.1016/j.paid.2016.02.075

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology*, *69*(3), 339–348. https://doi.org/10.1016/j.tpb.2005.09.005

Handgraaf, M. J. J., van Dijk, E., Vermunt, R. C., Wilke, H. A. M., & de Dreu, C. K. W. (2008). Less power or powerless? Egocentric empathy gaps and the irony of having little versus no power in social decision making. *Journal of Personality and Social Psychology*, *95*(5), 1136–1149. https://doi.org/10.1037/0022-3514.95.5.1136

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*(3859), 1243–1248. https://doi.org/10.1126/science.162.3859.1243

Harrell, A. (2012). Do religious cognitions promote prosociality? *Rationality and Society*, *24*(4), 463–482. https://doi.org/10.1177/1043463112463930

Haucap, J., & Müller, A. (2014). *Why are economists so different? Nature, nurture, and gender effects in a simple trust game* (Discussion paper). Düsseldorf Institute for Competitive Economics. Retrieved from http://www.dice.hhu.de/fileadmin/redaktion/Fakultaeten/Wirtschaftswissenschaftliche_Fakultaet/DICE/Discussion_Paper/DP-136_Haucap_Mueller.pdf

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (2004). *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*. New York, NY: Oxford University Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *The American Economic Review*, *91*(2), 73–78. https://doi.org/10.1257/aer.91.2.73

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., ... Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*(6), 795-815-855. https://doi.org/10.1017/S0140525X05000142

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Henrich, J., Mcelreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., ... Ziker, J. (2006). Supporting online material for "Costly Punishment across human societies." *Science*, *312*(1767). https://doi.org/10.1126/science.1127333

Hershfield, H. E., Cohen, T. R., & Thompson, L. (2012). Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. *Organizational Behavior and Human Decision Processes*, *117*(2), 298–310. https://doi.org/10.1016/j.obhdp.2011.11.002

Hertel, G., & Fiedler, K. (1994). Affective and cognitive influences in a social dilemma game. *European Journal of Social Psychology*, *145*, 131–145. https://doi.org/10.1002/ejsp.2420240110

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, *24*(3), 383-403-451.

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org

Hilbig, B. E., & Hessler, C. M. (2013). What lies beneath. How the distance between truth and lie drives dishonesty. *Journal of Experimental Social Psychology*, *49*(2), 263–266. https://doi.org/10.1016/j.jesp.2012.11.010

Hilbig, B. E., & Zettler, I. (2015). When the cat's away, some mice will play. A basic trait account of dishonest behavior. *Journal of Research in Personality*, *57*, 72–88. https://doi.org/10.1016/j.jrp.2015.04.003

Hildreth, J. A. D., Gino, F., & Bazerman, M. (2016). Blind loyalty? When group loyalty makes us see evil or engage in it. *Organizational Behavior and Human Decision Processes*, *132*, 16–36. https://doi.org/10.1016/j.obhdp.2015.10.001

Hoffman, E., McCabe, K. A., & Smith, V. L. (1996a). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory*, *25*, 289–301. https://doi.org/10.1017/cbo9780511528347.010

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. L. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, *7*(3), 346–380. https://doi.org/10.1006/game.1994.1056

Hoffman, E., McCabe, K., & Smith, V. L. (1996b). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, *86*(3), 653–660. https://doi.org/10.1017/cbo9780511528347.009

Hoffman, E., & Spitzer, M. L. (1985). Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *The Journal of Legal Studies*, *14*(2), 259–297. https://doi.org/10.1086/467773

Hoffmann, A., Diedenhofen, B., Verschuere, B., & Musch, J. (2015). A strong validation of the crosswise model using experimentally-induced cheating behavior. *Experimental Psychology*, *62*(6), 403–414. https://doi.org/10.1027/1618-3169/a000304

Holm, H. J., & Kawagoe, T. (2010). Face-to-face lying. An experimental study in Sweden and Japan. *Journal of Economic Psychology*, *31*(3), 310–321. https://doi.org/10.1016/j.joep.2010.01.001

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 1–36. https://doi.org/10.1007/s10683-011-9273-9

Houser, D., List, J. A., Piovesan, M., Samek, A., & Winter, J. (2016). Dishonesty: From parents to children. *European Economic Review*, *82*, 242–254. https://doi.org/10.1016/j.euroecorev.2015.11.003

Houser, D., Vetter, S., & Winter, J. (2012). Fairness and cheating. *European Economic Review, 56*(8), 1645–1655. https://doi.org/10.1016/j.euroecorev.2012.08.001

Hristova, E., Grinberg, M., Georgieva, I., & Borisova, M. (2013). Cooperation in prisoner's dilemma game: Influence of players' social roles. In *Proceedings of CogSci 2013*. Retrieved from https://mindmodeling.org/cogsci2012/papers/0082/paper0082.pdf

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods, 11*(2), 193–206. https://doi.org/10.1037/1082-989X.11.2.193

Hugh-Jones, D. (2016). Honesty, beliefs about honesty, and economic growth in 15 countries. *Journal of Economic Behavior and Organization, 127*, 99–114. https://doi.org/10.1016/j.jebo.2016.04.012

Hurkens, S., & Kartik, N. (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics, 12*(2), 180–192. https://doi.org/10.1007/s10683-008-9208-2

Innes, R., & Mitra, I. (2013). Is dishonesty contagious? *Economic Inquiry, 51*(1), 722–734. https://doi.org/10.1111/j.1465-7295.2012.00470.x

International Monetary Fund. (2015). World economic and financial surveys: World Economic Outlook Database. Retrieved from https://www.imf.org/external/pubs/ft/weo/2015/01/weodata/index.aspx

Isaac, M. R., McClue, K. F., & Plott, C. R. (1985). Public goods provision in an experimental environment. *Journal of Public Economics, 26*, 51–74. https://doi.org/10.1016/0047-2727(85)90038-6

Jacobsen, C., Fosgaard, T. R., & Pascual-Ezama, D. (2017). Why do we lie? A practical guide to the dishonesty literature. *Journal of Economic Surveys*, Advance online publication. https://doi.org/10.1111/joes.12204

Jacobsen, C., & Piovesan, M. (2015). Tax me if you can. An artifactual field experiment on dishonesty. *Journal of Economic Behavior and Organization, 124*, 7–14. https://doi.org/10.1016/j.jebo.2015.09.009

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology, 32*(5), 865–889. https://doi.org/10.1016/j.joep.2011.05.007

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *The Journal of Business, 59*(4), S285–S300. https://doi.org/10.1086/296367

Kallgren, C. A., Reno, R. R., & Cialdini, R. B. (2000). A focus theory of normative conduct: When norms do and do not affect behavior. *Personality and Social Psychology Bulletin, 26*(8), 1002–1012. https://doi.org/10.1177/01461672002610009

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations. A theory of interdependence*. New York, NY: Wiley.

Kerr, N. L. (1995). Norms in social dilemmas. In D. A. Schroeder (Ed.), *Social Dilemmas. Perspectives on Individuals and Groups* (pp. 31–47). Westport, CT: Greenwood.

Kilduff, G., Galinksy, A., Gallo, E., & Reade, J. (2015). Whatever it takes to win: Rivalry increases unethical behavior. *Academy of Management Journal, 59*(5), 1508–1534.

https://doi.org/10.5465/amj.2014.0545

Kirchgässner, G. (2005). (Why) are economists different? *European Journal of Political Economy*, *21*(3), 543–562. https://doi.org/10.1016/j.ejpoleco.2005.05.003

Kölle, F., Gächter, S., & Quercia, S. (2014). The ABC of cooperation in voluntary contribution and common pool extraction games. In *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik*. Retrieved from http://econpapers.repec.org/RePEc:zbw:vfsc14:100417

Kouchaki, M., Gino, F., & Jami, A. (2014). The burden of guilt. Heavy backpacks, light snacks, and enhanced morality. *Journal of Experimental Psychology: General*, *143*(1), 414–424. https://doi.org/10.1037/a0031769

Kouchaki, M., Smith-Crowe, K., Brief, A. P., & Sousa, C. (2013). Seeing green. Mere exposure to money triggers a business decision frame and unethical outcomes. *Organizational Behavior and Human Decision Processes*, *121*(1), 53–61. https://doi.org/10.1016/j.obhdp.2012.12.002

Kouchaki, M., & Smith, I. H. (2014). The morning morality effect. The influence of time of day on unethical behavior. *Psychological Science*, *25*(1), 95–102. https://doi.org/10.1177/0956797613498099

Kouchaki, M., & Wareham, J. (2015). Excluded and behaving unethically. Social exclusion, physiological responses, and unethical behavior. *Journal of Applied Psychology*, *100*(2), 547–556. https://doi.org/10.1037/a0038034

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, *46*(5), 1044–1057. https://doi.org/10.1037/0022-3514.46.5.1044

La Barbera, F., Ferrara, P. C., & Boza, M. (2014). Where are we coming from versus who we will become: The effect of priming different contents of European identity on cooperation. *International Journal of Psychology*, *49*(6), 480–487. https://doi.org/10.1002/ijop.12073

Lambsdorff, J. G., & Frank, B. (2010). Bribing versus gift-giving: An experiment. *Journal of Economic Psychology*, *31*(3), 347–357. https://doi.org/10.1016/j.joep.2010.01.004

Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy. Moralizing in reasoning, immorality in behavior. *Psychological Science*, *21*(5), 737–44. https://doi.org/10.1177/0956797610368810

Lanteri, A. (2008a). (Why) do selfish people self-select in economics? *Erasmus Journal for Philosophy & Economics*, *1*(1), 1–23.

Lanteri, A. (2008b). *The moral trial. On the ethics of economics. Dissertation*. Erasmus University Rotterdam. Retrieved from http://repub.eur.nl/pub/12050/

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111–194). Princeton, NJ: Princeton University Press.

Lee, J. J., Gino, F., Jin, E. S., Rice, L. K., & Josephs, R. A. (2015). Hormones and ethics. Understanding the biological basis of unethical conduct. *Journal of Experimental Psychology: General*, *144*(5), 891–897. https://doi.org/10.1037/xge0000099

Lee, J. J., Im, D. K., Parmar, B. L., & Gino, F. (2015). *Thick as thieves? Dishonest behavior and egocentric social networks* (Harvard Business School NOM Unit

Working Paper). https://doi.org/10.1007/BF00138862

Leibbrandt, A., & López-pérez, R. (2012). An exploration of third and second party punishment in ten simple games. *Journal of Economic Behavior and Organization, 84*(3), 753–766. https://doi.org/10.1016/j.jebo.2012.09.018

Lesorogol, C. K. (2007). Bringing norms in: The role of context in experimental dictator games. *Current Anthropology, 48*(6), 920–926. https://doi.org/10.1086/523017

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior & Human Decision Processes, 76*(2), 149–188. https://doi.org/10.1006/obhd.1998.2804

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics, 1*(3), 593–622. https://doi.org/10.1006/redy.1998.0023

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives, 21*(2), 153–174. https://doi.org/10.1257/jep.21.2.153

Lewis, A., Bardis, A., Flint, C., Mason, C., Smith, N., Tickle, C., & Zinser, J. (2012). Drawing the line somewhere. An experimental study of moral compromise. *Journal of Economic Psychology, 33*(4), 718–725. https://doi.org/10.1016/j.joep.2012.01.005

Liberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Personality and Social Psychology Bulletin, 30*(9), 1175–1185. https://doi.org/10.1177/0146167204264004

Liebrand, W. B. G. (1983). A classification of social dilemma games. *Simulation & Games, 14*(2), 123–138. https://doi.org/10.1177/104687818301400201

Liebrand, W. B. G., & McClintock, C. G. (1988). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality, 2*(3), 217–230. https://doi.org/10.1002/per.2410020304

Liu, C.-J., & Li, S. (2009). Contextualized self: When the self runs into social dilemmas. *International Journal of Psychology, 44*(6), 451–458. https://doi.org/10.1080/00207590902757377

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology, 57*(3), 426–441. https://doi.org/10.1037//0022-3514.57.3.426

Loomes, G. (1999). Some lessons from past experiments and some challanges for the future. *The Economic Journal, 109*(453), F35–F45. https://doi.org/10.1111/1468-0297.00401

López-Pérez, R., & Spiegelman, E. (2009). *Do economists lie more?* (Working paper No. 1885–6888). Madrid. Retrieved from http://www.uam.es/departamentos/economicas/analecon/especifica/mimeo/wp20124.pdf

López-Pérez, R., & Spiegelman, E. (2013). Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics, 16*, 233–247. https://doi.org/10.1007/s10683-012-9324-x

Lopez, M. C., Murphy, J. J., Spraggon, J. M., & Stranlund, J. K. (2012). Comparing the effectiveness of regulation and pro-social emotions to enhance cooperation: Experimental evidence from fishing communities in Colombia. *Economic Inquiry, 50*(1), 131–142. https://doi.org/10.1111/j.1465-7295.2010.00344.x

"Luiz Inácio Lula da Silva. The Rise and fall of a Brazilian leader." (2017, July 12). New York Times. Retrieved from https://www.nytimes.com/2017/07/12/world/americas/luiz-inacio-lula-da-silva-the-rise-and-fall-of-a-brazilian-leader.html

Lundquist, T., Ellingsen, T., Gribbe, E., & Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior and Organization, 70*(1–2), 81–92. https://doi.org/10.1016/j.jebo.2009.02.010

Mai, K. M., Ellis, A. P. J., & Welsh, D. T. (2015). The gray side of creativity. Exploring the role of activation in the link between creative personality and unethical behavior. *Journal of Experimental Social Psychology, 60*, 76–85. https://doi.org/10.1016/j.jesp.2015.05.004

Marglin, S. A. (2008). *The dismal science. How thinking like an economist undermines community*. Cambridge, MA: Harvard University Press.

Marwell, G., & Ames, R. E. (1981). Economists free ride, does anyone else? *Journal of Public Economics, 15*, 295–310. https://doi.org/10.1016/0047-2727(81)90013-X

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people. A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633–644. https://doi.org/10.1509/jmkr.45.6.633

McClure, M. J., Bartz, J. A., & Lydon, J. E. (2013). Uncovering and overcoming ambivalence: The role of chronic and contextually activated attachment in two-person social dilemmas. *Journal of Personality, 81*(1), 103–117. https://doi.org/10.1111/j.1467-6494.2012.00788.x

Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., & Ariely, D. (2009). Too tired to tell the truth. Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology, 45*(3), 594–597. https://doi.org/10.1016/j.jesp.2009.02.004

Mellers, B. A., Haselhuhn, M. P., Tetlock, P. E., Silva, J. C., & Isen, A. M. (2010). Predicting behavior in economic games by looking through the eyes of the players. *Journal of Experimental Psychology, 139*(4), 743–755. https://doi.org/10.1037/a0020280

Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review, 8*(2), 193–210. https://doi.org/10.2307/4609267

Messick, D. M. (1999). Alternative logics for decision making in social settings. *Journal of Economic Behavior and Organization, 39*(1), 11–28. https://doi.org/10.1016/S0167-2681(99)00023-2

Milgram, S. (1974). *Obedience to authority. An experimental view*. New York, NY: Harper Collins.

Miller, D. T. (1999). The norm of self-interest. *The American Psychologist, 54*(12), 1053–1060. https://doi.org/10.1037/0003-066X.54.12.1053

Muehlheusser, G., Roider, A., & Wallmeier, N. (2015). Gender differences in honesty. Groups versus individuals. *Economics Letters, 128*, 25–29. https://doi.org/10.1016/j.econlet.2014.12.019

Muñoz-Izquierdo, N., Liaño, B. G.-G. de, Rin-Sánchez, F. D., & Pascual-Ezama, D. (2014). *Economists: Cheaters with altruistic instincts* (Munich Personal RePEc Archive). Retrieved from https://mpra.ub.uni-muenchen.de/60678/

Nietzsche, F. W. (1988). Frühjahr 1888 [Spring 1888]. In *Nachgelassene Fragmente [Posthumous Fragments]*. Munich, Germany: DTV.

Olson, M. (1965). *The logic of collective action. Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Ortmann, A. (2005). Field experiments in economics: Some methodological caveats. In G. W. Harrison, J. Carpenter, & J. A. List (Eds.), *Field experiments in economics* (pp. 51–70). Amsterdam, The Netherlands: Elsevier.

Orwant, C. J., & Orwant, J. E. (1970). A comparison of interpreted and abstract versions of mixed-motive games. *Journal of Conflict Resolution*, *14*(1), 91–97. https://doi.org/10.1177/002200277001400110

Pascual-Ezama, D., Fosgaard, T. R., Cardenas, J. C., Kujal, P., Veszteg, R., Gil-Gómez de Liaño, B., ... Brañas-Garza, P. (2015). Context-dependent cheating. Experimental evidence from 16 countries. *Journal of Economic Behavior and Organization*, *116*, 379–386. https://doi.org/10.1016/j.jebo.2015.04.020

Peeters, R., Vorsatz, M., & Walzl, M. (2015). Beliefs and truth-telling: A laboratory experiment. *Journal of Economic Behavior and Organization*, *113*, 1–12. https://doi.org/10.1016/j.jebo.2015.02.009

Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a “cooperative phenotype” that is domain general and temporally stable. *Nature Communications*, *5*(4939), 1–8. https://doi.org/10.1038/ncomms5939

Piff, P. K., Stancato, D. M., Côté, S., Mendoza-Denton, R., & Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(11), 4086–4091. https://doi.org/10.1073/pnas.1118373109

Pillutla, M. M., & Chen, X.-P. (1999). Social norms and cooperation in social dilemmas: The effects of context and feedback. *Organizational Behavior & Human Decision Processes*, *78*(2), 81–103. https://doi.org/10.1006/obhd.1999.2825

Ploner, M., & Regner, T. (2013). Self-image and moral balancing. An experimental analysis. *Journal of Economic Behavior and Organization*, *93*, 374–383. https://doi.org/10.1016/j.jebo.2013.03.030

Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, *28*, 363–392. https://doi.org/10.1146/annurev.ps.28.020177.002051

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org.

Ramalingam, A. (2012). The relevance of irrelevant information in the dictator game. *Economics Bulletin*, *32*(1), 746–754. https://doi.org/10.2139/ssrn.1796227

Rand, A. (1997). Atlas shrugged. New York, NY: Signet.

Rand, D. G., Dreber, A., Haque, O. S., Kane, R. J., Nowak, M. A., & Coakley, S. (2014). Religious motivations for cooperation: An experimental investigation using explicit primes. *Religion, Brain & Behavior*, *4*(1), 31–48. https://doi.org/10.2139/ssrn.2123243

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489*(7416), 427–30. https://doi.org/10.1038/nature11467

Rand, D. G., Newman, G. E., & Wurzbacher, O. M. (2014). Social context and the dynamics of cooperative choice. *Journal of Behavioral Decision Making*, *28*(2), 159–166. https://doi.org/10.1002/bdm.1837

Rasmußen, A. (2015). Reporting behavior. A literature review of experimental studies. *Central European Journal of Operations Research*, *23*(2), 283–311. https://doi.org/10.1007/s10100-014-0379-y

Ratner, R. K., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology*, *81*(1), 5–16. https://doi.org/10.1037/0022-3514.81.1.5

Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, *88*(7–8), 1625–1644. https://doi.org/10.1016/S0047-2727(03)00021-5

Rhyne, W. J. (2008). *Culture, self-orientation, and reward structure effects. Measuring cheating behavior in China and the USA* (SSRN Working paper series). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1931751

Rigdon, M. L., & D'Esterre, A. (2014). The effects of competition on the nature of cheating behavior. *Southern Economic Journal*, *81*(4), 1012–1024. https://doi.org/10.4284/0038-4038-2012.301

Rode, J. (2010). Truth and trust in communication: Experiments on the effect of a competitive context. *Games and Economic Behavior*, *68*(1), 325–338. https://doi.org/10.1016/j.geb.2009.05.008

Roeser, K., McGregor, V. E., Stegmaier, S., Mathew, J., Kübler, A., & Meule, A. (2016). The dark triad of personality and unethical behavior at different times of day. *Personality and Individual Differences*, *88*, 73–77. https://doi.org/10.1016/j.paid.2015.09.002

Rosenbaum, S. M., Billinger, S., & Stieglitz, N. (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, *45*, 181–196. https://doi.org/10.1016/j.joep.2014.10.002

Roth, A. E. (1995). Introduction to experimental economics. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 3–109). Princeton, NJ: Princeton University Press.

Rubinstein, A. (2006). A sceptic's comment on the study of economics. *Economic Journal*, *116*(510), C1–C9. https://doi.org/10.1111/j.1468-0297.2006.01071.x

Rücker, G., Carpenter, J. R., & Schwarzer, G. (2011). Detecting and adjusting for small-study effects in meta-analysis. *Biometrical Journal*, *53*(2), 351–368. https://doi.org/10.1002/bimj.201000151

Ruffle, B. J., & Tobol, Y. (2014). Honest on Mondays: Honesty and the temporal separation between decisions and payoffs. *European Economic Review*, *65*, 126–135. https://doi.org/10.1016/j.euroecorev.2013.11.004

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7*(1), 58–92. https://doi.org/10.1177/104346319500700l004

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research, 37*(3), 409–425. https://doi.org/10.1086/651235

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature, 470*(7335), 437. https://doi.org/10.1038/470437a

Schurr, A., & Ritov, I. (2016). Winning a competition predicts dishonest behavior. *Proceedings of the National Academy of Sciences of the United States of America, 113*(7), 1754–1759. https://doi.org/10.1073/pnas.1515102113

Schwarzer, G. (2007). meta: An R package for meta-analysis. *R News, 7*(3), 40–45. https://doi.org/10.1159/000323281

Schwarzer, G., Carpenter, J., & Rücker, G. (2010). Empirical evaluation suggests Copas selection model preferable to trim-and-fill method for selection bias in meta-analysis. *Journal of Clinical Epidemiology, 63*(3), 282–288. https://doi.org/10.1016/j.jclinepi.2009.05.008

Schweitzer, M. E., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty, 25*(2), 185–201. https://doi.org/10.1023/A:1020647814263

Scott, J. H., & Rothman, M. P. (1975). The effect of an introductory economics course on student political attitudes. *The Journal of Economic Education, 6*(2), 107–112. https://doi.org/10.1080/00220485.1975.10845410

Selten, R., & Ockenfels, A. (1998). An experimental solidarity game. *Journal of Economic Behavior and Organization, 34*, 517–539. https://doi.org/10.1016/S0167-2681(97)00107-8

Shalvi, S., Dana, J., Handgraaf, M. J. J., & De Dreu, C. K. W. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes, 115*(2), 181–190. https://doi.org/10.1016/j.obhdp.2011.02.001

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23*(10), 1264–1270. https://doi.org/10.1177/0956797612443835

Shalvi, S., Handgraaf, M. J. J., & De Dreu, C. K. W. (2011). Ethical manoeuvring: Why people avoid both major and minor lies. *British Journal of Management, 22*, S16–S27. https://doi.org/10.1111/j.1467-8551.2010.00709.x

Shalvi, S., & Leiser, D. (2013). Moral firmness. *Journal of Economic Behavior and Organization, 93*, 400–407. https://doi.org/10.1016/j.jebo.2013.03.014

Shariff, A. F., & Norenzayan, A. (2007). God is watching you: Priming god concepts increases prosocial behavior in an anonymous economic game. *Psychological Science, 18*(9), 803–809. https://doi.org/10.1111/j.1467-9280.2007.01983.x

Sher, S., & McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition, 101*(3), 467–494. https://doi.org/10.1016/j.cognition.2005.11.001

Sher, S., & McKenzie, C. R. M. (2008). Framing effects and rationality. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 79–96). Oxford, United Kingdom: Oxford University Press.

Shu, L. L., & Gino, F. (2012). Sweeping dishonesty under the rug: How unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology, 102*(6), 1164–1177. https://doi.org/10.1037/a0028381

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Pychology Bulletin*, *37*(3), 330–349. https://doi.org/10.1177/0146167211398138

Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(38), 15197–15200. https://doi.org/10.1073/pnas.1209746109

Simpson, B., & Eriksson, K. (2009). The dynamics of contracts and generalized trustworthiness. *Rationality and Society*, *21*(1), 59–80. https://doi.org/10.1177/1043463108099348

Smeesters, D., Warlop, L., Van Avermaet, E., Corneille, O., & Yzerbyt, V. (2003). Do not prime hawks with doves: The interplay of construct activation and consistency of social value orientation on cooperative behavior. *Journal of Personality and Social Psychology*, *84*(5), 972–987. https://doi.org/10.1037/0022-3514.84.5.972

Smith, E. R., & Queller, S. (2001). Mental representation. In A. Tesser & N. Schwarz (Eds.), *Blackwell Handbook of Social Psychology. Intraindividual Processes* (pp. 111–133). Oxford: Blackwell.

Smithson, M., & Foddy, M. (1999). Theories and strategies for the study of social dilemmas. In M. Foddy, M. Smithson, S. Schneider, & M. Hogg (Eds.), *Resolving Social Dilemmas. Dynamic, Structural, and Integroup Aspects* (pp. 1–14). Philadelphia, PA: Psychology Press.

Stigler, G. J. (1981). Economics or ethics? *The Tanner Lectures on Human Values. Delivered at Harvard University April 24, 25, and 28,1980*, 145–191. Retrieved from http://tannerlectures.utah.edu/_documents/a-to-z/s/stigler81.pdf

Sugden, R. (1991). Rational choice: A survey of contributions from economics and philosophy. *The Economic Journal*, *101*(407), 751–785. https://doi.org/10.2307/2233854

Sugden, R. (1993). Thinking as a team: Towards an explanation of nonselfish behavior. *Social Philosophy & Policy*, *10*(1), 69–89. https://doi.org/10.1017/S0265052500004027

Sugden, R. (2000). Team preferences. *Economics and Philosophy*, *16*(2), 175–204. https://doi.org/10.1017/S0266267100000213

Sugden, R. (2015). A theory of focal points. *The Economic Journal*, *105*(430), 533–550. https://doi.org/10.2307/2235016

Sutter, M. (2009). Deception through telling the truth?! Experimental evidence from individuals and teams. *The Economic Journal*, *119*(534), 47–60. https://doi.org/10.1111/j.1468-0297.2008.02205.x

Tao, L., & Au, W. (2014). Values, self and other-regarding behavior in the dictator game.

Rationality and Society, *26*(1), 46–72. https://doi.org/10.1177/1043463113512995

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading. The role of self-deception in unethical behavior. *Social Justice Research*, *17*(2), 223–236. https://doi.org/10.1023/B:SORE.0000027411.35832.53

Ter Meer, J. (2014). *The indirect effect of monetary incentives on deception* (CGS Working Paper). Retrieved from http://www.cgs.uni-koeln.de/fileadmin/wiso_fak/cgs/pdf/working_paper/cgswp_05-04.pdf

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, *22*(13), 2113–2126. https://doi.org/10.1002/sim.1461

Thielmann, I., Hilbig, B. E., Zettler, I., & Moshagen, M. (2016). On measuring the sixth basic personality dimension: A comparison between HEXACO honesty-humility and Big Six honesty-propriety. *Assessment*, 1–13. https://doi.org/10.1177/1073191116638411

Thoemmes, F. (2015). Reversing arrows in mediation models does not distinguish plausible models. *Basic and Applied Social Psychology*, *37*(4), 226–234. https://doi.org/10.1080/01973533.2015.1049351

Thompson, M., Ellis, R., & Wildavsky, A. (1990). *Cultural theory*. Boulder, CO: Westview.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38. https://doi.org/10.18637/jss.v059.i05

Torsvik, G., Molander, A., Tjøtta, S. T., & Kobbeltvedt, T. (2011). Anticipated discussion and cooperation in a social dilemma. *Rationality and Society*, *23*(2), 199–216. https://doi.org/10.1177/1043463111404664

Transparency International. (2017). Corruption Perception Index 2016. Retrieved July 6, 2017, from https://www.transparency.org/news/feature/corruption_perceptions_index_2016

Trikalinos, T. A., & Ioannidis, J. P. (2005). Assessing the evolution of effect sizes over time. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 241–259). Chichester, United Kingdom: John Wiley & Sons.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/https://doi.org/10.1017/cbo9780511610226.003

Utikal, V., & Fischbacher, U. (2013). Disadvantageous lies in individual decisions. *Journal of Economic Behavior and Organization*, *85*(1), 108–111. https://doi.org/10.1016/j.jebo.2012.11.011

Uziel, L., & Hefetz, U. (2014). The selfish side of self-control. *European Journal of Personality*, *28*(5), 449–458. https://doi.org/10.1002/per.1972

van der Heijden, A. J., Groenen, P. J. F., Zeelenberg, R., & te Lindert, R. (2014). *Report of the Smeesters follow-up investigation committee*. Erasmus University Rotterdam. Retrieved from http://www.rsm.nl/fileadmin/Images_NEW/News_Images/2014/Report_Smeesters_follow-up_investigation_committee.final.pdf

Van Lange, P. A. M., Schippers, M., & Balliet, D. (2011). Who volunteers in psychology experiments? An empirical review of prosocial motivation in volunteering. *Personality and Individual Differences, 51*(3), 279–284. https://doi.org/10.1016/j.paid.2010.05.038

Van Zant, A. B., & Kray, L. J. (2014). "I can't lie to your face". Minimal face-to-face interaction promotes honesty. *Journal of Experimental Social Psychology, 55*, 234–238. https://doi.org/10.1016/j.jesp.2014.07.014

Vetter, S. (2012). *Empirical studies of individual behavior. Cheating, corruption, and insurance choice*. (Doctoral dissertation). Ludwig Maximilian University of Munich, Germany. Retrieved from https://edoc.ub.uni-muenchen.de/15005/

von Borgstede, C., Dahlstrand, U., & Biel, A. (1999). From ought to is: Moral norms in large-scale social dilemmas. *Göteborg Psychological Reports, 29*(5), 1–17.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wagenaar, W. A., Keren, G., & Lichtenstein, S. (1988). Islanders and hostages: Deep and surface structures of decision problems. *Acta Psychologica, 67*, 175–189. https://doi.org/10.1016/0001-6918(88)90012-1

Wang, L., Malhotra, D., & Murnighan, J. K. (2011). Economics education and greed. *Academy of Management Learning and Education, 10*(4), 643–660. https://doi.org/10.5465/amle.2009.0185

Wang, L., & Murnighan, J. K. (2016). How much does honesty cost? Small bonuses can motivate ethical behavior. *Management Science*, Advanced online publication. https://doi.org/10.1287/mnsc.2016.2480

Weber, J. M., Kopelman, S. J., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review, 8*(3), 281–307. https://doi.org/10.1207/s15327957pspr0803

Welsh, D. T., Ellis, A. P. J., Christian, M. S., & Mai, K. M. (2014). Building a self-regulatory model of sleep deprivation and deception: The role of caffeine and social influence. *Journal of Applied Psychology, 99*(6), 1268–1277. https://doi.org/10.1037/a0036202

Welzer, H. (2007). *Täter: Wie aus ganz normalen Menschen Massenmörder werden [Offenders: How normal people become mass murderers]*. Berlin, Germany: Fischer.

Wibral, M., Dohmen, T., Klingmüller, D., Weber, B., & Falk, A. (2012). Testosterone administration reduces lying in men. *PLOS ONE, 7*(10), e46774. https://doi.org/10.1371/journal.pone.0046774

Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes, 115*(2), 157–168. https://doi.org/10.1016/j.obhdp.2010.10.001

Winnicott, D. (1945). *Home is where we start from*. London, United Kingdom: Penguin.

Winterich, K. P., Mittal, V., & Morales, A. C. (2014). Protect thyself: How affective self-protection increases self-interested, unethical behavior. *Organizational Behavior and Human Decision Processes, 125*(2), 151–161. https://doi.org/10.1016/j.obhdp.2014.07.004

Wit, A. P., & Wilke, H. A. M. (1992). The effect of social categorization on cooperation in three types of social dilemmas. *Journal of Economic Psychology*, *13*(1), 135–151. https://doi.org/10.1016/0167-4870(92)90056-D

Wong, R. Y.-M., & Hong, Y.-Y. (2005). Dynamic influences of culture on cooperation in the prisoner's dilemma. *Psychological Science*, *16*(6), 429–434. https://doi.org/10.1111/j.0956-7976.2005.01552.x

Wu, Y., Loch, C., & Ahmad, G. (2011). Status and relationships in social dilemmas of teams. *Journal of Operations Management*, *29*(7–8), 650–662. https://doi.org/10.1016/j.jom.2011.03.004

Yaniv, G., & Siniver, E. (2016). The (honest) truth about rational dishonesty. *Journal of Economic Psychology*, *53*, 131–140. https://doi.org/10.1016/j.joep.2016.01.002

Young, H. P. (2003). The power of norms. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation* (pp. 389–399). Cambridge, MA: MIT Press.

Zaleskiewicz, T., Gasiorowska, A., & Kesebir, P. (2015). The Scrooge effect revisited: Mortality salience increases the satisfaction derived from prosocial behavior. *Journal of Experimental Social Psychology*, *59*, 67–76. https://doi.org/10.1016/j.jesp.2015.03.005

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*(3), 299–310. https://doi.org/10.1023/A:1026277420119

Zettler, I., Hilbig, B. E., Moshagen, M., & de Vries, R. E. (2015). Dishonest responding or true virtue? A behavioral test of impression. *Personality and Individual Differences*, *81*, 107–111. https://doi.org/10.1016/j.paid.2014.10.007

Zhong, C.-B. (2011). The ethical dangers of deliberative decision making. *Administrative Science Quarterly*, *56*(1), 1–25.

Zhong, C.-B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police. Darkness increases dishonesty and self-interested behavior. *Psychological Science*, *21*(3), 311–314. https://doi.org/10.1177/0956797609360754

Zhong, C.-B., Loewenstein, J., & Murnighan, J. K. (2007). Speaking the same language: The cooperative effects of labeling in the prisoner's dilemma. *Journal of Conflict Resolution*, *51*(3), 431–456. https://doi.org/10.1177/0022002707300834

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75–98. https://doi.org/10.1007/s10683-009-9230-z