Essays in Moral Decisions

Jareef Bin Martuza

Dissertation submitted to the Department of Strategy and Management at the Norwegian School of Economics for partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD)

March 2024

ABSTRACT

People are often faced with decisions about right and wrong. Undoubtedly, individual dispositions and material incentives can influence those decisions to a large extent. However, beyond personal characteristics and expected costs versus benefits, what else may affect moral decisions? Using a combination of scenario-based and incentivized experiments, this dissertation comprises five articles that present causal evidence on how moral intentions and behavior can systematically vary across contexts.

Article 1, conducted in the registered report format, tried to replicate and extend Conway and Peetz's (2012) influential hypothesis that recalling behaviors from the recent (distant) past should lead to compensatory (consistent) moral behavior. With one of the largest single-lab studies (N = 5,091), investigating sequential moral behavior, we robustly show that recalling moral behavior led to higher prosocial intentions than recalling either immoral or neutral behavior, irrespective of recalling from the recent or distant past.

Article 2 examines how the mere size of an organization can affect dishonest behavior against it. Across eight scenario-based and incentive-aligned experiments (combined N = 5,670), we find that people are more likely to both intend to and actually cheat big businesses than small businesses for selfish gain, rendering a meta-analytic effect size corresponding to .31 of a standard deviation. Further, based on mediation analyses, we also suggest that one important explanation of this biased dishonesty is that people perceive big businesses as less moral and less vulnerable than small businesses.

Article 3 investigates intergroup bias in selfish and coalitional dishonesty. In two experiments, we tempted Democrat and Republican voters to double their earnings (or the earnings of someone else) by self-reporting a correct guess of a die-roll. In Experiment 1 (N = 1,176), we found that individuals were equally likely to cheat their political ingroup and

outgroup members to double their payoffs. In Experiment 2 (N = 1,710), participants lied at a significantly higher rate to benefit an ingroup member than to benefit an outgroup member (9 percentage points).

Article 4 aims to answer a simple question: Do people believe that others are similarly, more, or less dishonest than they truly are? In a research program spanning three years (2022-24) and a total of 31 different effects (combined N = 8,127), we tempted participants to cheat without any repercussions or detection risks, and asked them to estimate what percentage of other participants would lie in the same situation. Our meta-analysis revealed a significant overestimation of others' dishonesty by an average of 14 percentage points, suggesting the world is less dishonest than people think.

Article 5 tests for gender bias in interpersonal dishonesty by recruiting a total of 3,166 participants from nine countries and providing them an opportunity to cheat and increase their payoffs at the cost of another male, female, or sex-unmentioned participants. Overall, females were cheated significantly less (22%) than sex-unmentioned participants. Interestingly, the effect was significantly stronger among female decision-makers, who cheated other females substantially less (53.6%) than other male participants.

Theoretically, the dissertation contributes to the moral decision-making literature across topics such as sequential moral behavior, organizational perceptions, intergroup relations, beliefs, and gender biases. The findings are relevant to fellow researchers studying both basic judgment and decision-making, and those in applied settings such as organizations and marketplaces. Methodologically, most studies in the dissertation used incentivized economic experiments to study psychological phenomena, providing behavioral evidence to respective research questions. Practically, articles in this dissertation can inform managers, policymakers, and society at large on how everyday people make moral decisions.

ACKNOWLEDGEMENTS

The one(s)	*Grateful for*
Helge Thorbjørnsen	Giving my first break, trusting me with and supervising several projects, reviewing the gender paper & the introduction section, and co-authoring 3 dissertation articles
Hallgeir Sjåstad	Introducing me to registered reports, teaching me to write for leading journals, and co-authoring 3 dissertation articles
Olivia Kim	Co-authoring Article 1 and bouncing morality research ideas every other time running into each other
Siv Skard	Introducing me to (dis)honesty research and patiently mentoring and editing during my first-ever research paper
Esra Aslan	Giving instant feedback on most ideas, reviewing drafts, testing many of my Qualtrics, and being the best office mate
Hege Landsvik	Reviewing the gender paper, testing many of my Qualtrics
Experimental Research Group	The Friday meets where I learned new things every time
Sven Arne Haugland	Helping me settle in and encouraging me to explore courses
Einar Breivik	Instructing fundamental courses about empirical research
Magne Supphellen	Giving helpful feedback during the midterm evaluation

Distribusjonsøkonomiske forskningsprosjekter, CEE, SOL, and SNF	Funding several of my studies
Iffat Tarannum, Simen Bø, Denise Utochkin, Lars Moen, and Silas Braun	Testing Qualtrics, discussing ideas now and then
Katrine Berg Nødtvedt	Being the role model PhD student to look up to
Gavan J. Fitzsimons, Tanya L. Chartrand, and Duke's consumer behavior group	Being wonderful hosts during my research stay, high-level feedback on research projects, many tips & and tricks to design creative studies, and inspiring me to strive for more
All my reviewers	Taking their time to give feedback on manuscripts
Participants of La Londe, EMAC, and EACR	Being open-minded, helpful, and interested in discussing all things research, academia, and other things fun
R, jamovi, Google Scholar, and the OSF	Providing many things that I take for granted, and enabling Good Samaritans to share their research and materials
Ritsumeikan APU, BI Norwegian, and RUG	Generous scholarships, great student life, and courses from liberal arts to data science- which shaped many of my priors
Alina Silitrari	Giving the best pep talks, feedback on ideas, and a thousand other things that brought me joy
Jesmin Ara Sultana and A.K.M. Martuza Ahmed	Making and raising me, planting the idea of a PhD, and motivating me to power through the ups and downs

Table of Contents

List of Articles

- **Article 1:** Martuza, J., & Kim, O. (2024). Does Conceptual Abstraction Moderate Whether Past Moral Deeds Motivate Consistency or Compensatory Behavior? A Registered Replication and Extension of Conway and Peetz (2012). *Forthcoming at the Personality and Social Psychology Bulletin*.

- **Article 2:** Martuza, J., Sjåstad, H., & Thorbjørnsen, H. (2024). Business-Size Bias in Moral Concern: People are More Dishonest Against Big than Small Organizations. Working paper.

- **Article 3:** Martuza, J., Sjåstad, H., & Thorbjørnsen, H. (2024). Intergroup Bias in Dishonesty: Selfish Versus Coalitional Lying. Working paper.

- **Article 4:** Martuza, J., Sjåstad, H., & Thorbjørnsen, H. (2024). Beliefs Versus Reality: People Overestimate the Actual Dishonesty of Others. Working paper.

- **Article 5:** Martuza, J. (2024). A Registered Report on Gender Bias in Interpersonal Dishonesty: Are Females and Males Cheated Differently?. Working paper. (Stage 1 proposal has been peer reviewed and accepted in principle at the Social Psychological and Personality Science. Experiment conducted. Stage 2 is yet to be peer-reviewed.)

Chapters 2-6 are based on separate scientific papers that have been published or are in preparation for submission. So, between Chapters 2-6, some theoretical and methodological overlap exists, and each may be read independently.

Chapter 1. Background

1.1. Moral decisions and behavior

We often make decisions about right and wrong. *What* is considered as right and wrong may vary across social norms, cultures, ideologies, and individuals (Haidt, 2001). Nonetheless, moral decisions often involve a decision-maker trading off personal benefits against the welfare of others (Crockett, 2016). For example, one may forgo buying luxury items and donate to charity, volunteer at a soup kitchen instead of watching television, or underreport income to pay less taxes. Of course, moral decisions can comprise both omission and commission of prosocial and/or dishonest behaviors. That is, within each set of possible alternatives, one can make both moral *and* immoral decisions (Tenbrunsel & Smith-Crowe, 2008).

Beyond moral philosophy, scholars in several domains including psychology (Malle, 2021), economics (Sen et al., 2020), biology (Kurzban et al., 2015), sociology (Alan Fine, 2019), anthropology (Mattingly & Throop, 2018), and marketing (Campbell & Winterich, 2018) have devoted considerable time, energy, and resources to understanding what morality is. Whereas moral judgments are assessments about the rightness or wrongness of (intended) actions and acting entities (Malle, 2021), moral *decisions* involve choices and behavior that can affect the self and other entities (Crockett, 2016). Consider a scenario where one can misreport private information for personal benefit. It can be underreporting income in tax returns, inserting erroneous information in insurance claims forms, etc. In these situations, as an observer, one may judge the act (misreporting) and/or the person (cheater) on moral dimensions, such as the extent the action or person is bad/wrong. However, if one is presented with the opportunity, they will be "making a decision" whether to cheat or not.

Scholars studying moral decision-making have studied various dimensions of moral behavior such as prosociality (Keltner et al., 2014), cooperation (Tomasello & Vaish, 2013), fairness (Cappelen et al., 2007), altruism (Fehr & Fischbacher, 2003), and honesty (Cohn et al., 2019). Although the definitions and/or conceptualizations of these dimensions often overlap, they can have some operational distinctions. For instance, prosocial behaviors that are intended to benefit others have been argued as being one of the key artifacts of human evolution over thousands of years (Hare, 2017). This evolutionary perspective proposes morality as a form of cooperation, where individuals suppress or align their self-interests with others (Tomasello & Vaish, 2013). Relatedly, the modern morality-as-cooperation perspective (Curry, 2016; Curry et al., 2019; Greene, 2015; Rai & Fiske, 2011) also proposes morality as a facilitator of cooperation in societies (Curry et al., 2019). Fairness- referring to just, equitable, and impartial treatment- has attracted considerable attention from moral decision-making scholars across economics and psychology (Cappelen et al., 2007; Feess et al., 2021; Houser et al., 2012; Jihwan Chae et al., 2022; Leib et al., 2021). Altruism is seen as one of the highest forms of moral behavior, where the decision-maker performs behavior for the benefit of others even at a *cost* to the self (Kurzban et al., 2015). Finally, honesty is widely seen as behaving truthfully even when tempted by material benefits (or the lack of material costs), and is a key component in how modern economies and society function smoothly (Abeler et al., 2019).

Despite these nuances, categorizing and differentiating between different dimensions of moral decisions is not the purpose of my dissertation. I primarily study the prosociality and honesty aspects of moral decisions. Further, making or not making distinctions between them would not aid or impede comprehension of my dissertation. Even more so, I study systematic *differences* in people's prosocial and (dis)honest behavior in different situations. So, it is not what the moral decision or behavior is that is the object of investigation. Rather, in the moral domain, my dissertation asks how those behaviors systematically *vary* as a function of the context. For instance, when are individuals more likely to behave prosocially? Which individuals are more likely to be cheated?

In that vein, an important stream of research in moral psychology studies *how* people form moral judgments (Greene & Haidt, 2002; Haidt, 2007; Malle, 2021) and what may drive systematic differences in people's judgments (e.g., Forbes & Stellar, 2022; Pizarro & Tannenbaum, 2012; Uhlmann et al., 2015; Weidman et al., 2020). For instance, research shows that people judge close others more leniently than distant others for the *same* transgressions (Forbes & Stellar, 2022; Weidman et al., 2020). This hints at how contextual factors, such as the actor's relationship to the observer, may play an important role in shaping the moral acceptability of immoral actions, even when the action is the same.

If contextual factors, beyond the extent of benefit or harm, can affect moral judgments, then, might we expect that external factors may create systematic differences in moral decisions as well? The bounded rationality perspective (Simon, 1990) on moral behavior proposes that moral behavior is shaped not by character traits or rational deliberation only, but is influenced by the interplay between people's minds and the environment (Gigerenzer, 2010). Further, the social intuitionist approach to understanding morality proposes that people make moral judgments quickly, based on intuition of what feels right and wrong, rather than detailed moral reasoning (Haidt, 2001). If the same applies to moral decisions, then people's intuitions of what feels right and wrong may also affect to what extent people would *behave* morally.

Please consider a situation where the decision context may likely affect intuitions of what feels right and wrong to do. Imagine you volunteered to clean up the local beach on a Sunday. The next day (Monday), you may "feel more moral" for having performed a prosocial behavior the previous day (Conway & Peetz, 2012). That Monday, you are requested to donate to a charity: How much would you donate? Whatever the amount may be, would it be different in a counterfactual world where you had not volunteered your time the day before? Might it also differ from another counterfactual world where instead of volunteering, you misreported on your tax returns to keep more of your income?

Of course, when people make moral decisions, it is difficult to access their minds and know exactly why they behave the way they behave. As experimental researchers, we mostly observe people's decisions in the context of our studies. Based on those observations, reasonable theories are proposed to explain what people did and why they did it. Drawing on the earlier example, if most people donate more after having done something good, one could conclude that perhaps doing good makes people do more good afterward. Conversely, if people donate less after having done something good, one could conclude that doing good makes people do less good afterward. We will discuss this in more detail in section *1.2 sequential moral behavior*.

Beyond how moral decisions may be shaped by the context one makes the decision, let us also consider how it may also be influenced by *whom* the decision affects. For instance, are there systematic differences in who evokes more moral concern? A burgeoning literature on prosocial behavior has established that people are more generous to some than others (De Dreu et al., 2022). For example, people are more likely to support small businesses over big businesses (Paharia et al., 2011), share more resources with their ingroup than their outgroup (Halevy et al., 2008), and are more concerned about moral harm to women than men (Reynolds et al., 2020).

If people are more morally concerned about some than others, might they find it more acceptable to cheat some individuals and/or organizations more than others? The answer may not be so straightforward. First, consider how prosocial behavior toward another entity versus cheating another entity may be different in essence. When you behave prosocially toward someone, you do something that benefits them, which often may come at some personal cost to yourself, or perhaps to another entity. There are multiple explanations for why people behave prosocially in the first place, including evolutionary accounts (Hare, 2017), reciprocal expectations (Greene, 2013), and even warm glow (Crumpler & Grossman, 2008).

Conversely, when you cheat someone, you do something that benefits *you*, and your benefit would come at the cost of *whom* you cheat. In studying cheating or dishonest behavior, my dissertation adopts a dyadic view in thinking about both perpetrators and victims of dishonest behavior (Gray & Wegner, 2009). Further, the act of dishonesty has to be such that the material benefits derived by the perpetrator equals the material costs inflicted on the victim (Jiang, 2013). In this dissertation, I will use cheating and dishonest behavior interchangeably, all pointing to the same notion of lying for personal gain at the cost of someone else. One exception is Article 3, where we study both self-interested and other-benefitting dishonesty, that is, systematic differences in individuals lying for personal gain and lying to benefit certain others.

Broadly, prosocial behavior is fundamentally driven by the intent to contribute positively to the welfare of others, often manifesting through acts of altruism or community support. On the contrary, dishonest behavior primarily serves the individual's interests, pursuing personal gains at the cost of someone else. To that end, cheating Entity X more than Entity Y may not mirror being more prosocial to Entity Y than Entity X because the *motivations* may differ. For example, a customer may be more likely to fraudulently return items to a large firm than a small firm, while also being more likely to support a small firm during economic downturns than a large firm. Although the behaviors may seem to be directionally the same in terms of greater moral concern toward small businesses, the *mechanisms* for systematic differences may not be the same. Whereas differences in support may be driven by a greater desire to contribute positively to small businesses, differences in cheating may be driven by greater rationalization of dishonest behavior affecting large businesses.

We can also examine decisions to be dishonest through the influential lens of cognitive dissonance (Festinger, 1957). In his seminal work, Leon Festinger proposed that people experience psychological discomfort when they hold contradictory beliefs, values, or attitudes related to their decision-making. In the moral domain, most people want to see themselves as good, decent, and honest people, and presumably only cheat to the extent that it allows maintenance of a positive self-concept (Mazar et al., 2008). In the context of dishonesty, two strong forces may be at odds when considering the decision to cheat. On one end, the temptation of material gains influences attitudes in favor of cheating. On the other end, the motive to see oneself as a moral and honest person may influence attitudes against cheating. Turning back to Festinger, people are motivated to reduce the *dissonance* of holding contradictory attitudes, and they may try to do so by either changing their beliefs or justifying their behavior (Festinger, 1957).

Although I do not examine differences between belief changes and self-justifications, let us nonetheless consider a *third* force that may help the person reduce their dissonance regardless of being honest (and forgoing material gains) or dishonest (potentially harming moral self-image). This third force may be any rationale or intuition that may help reduce the cognitive dissonance of having competing motives: acquiring material gains vs. maintaining an honest self-concept. In that vein, I propose that this third force may be a function of *whom* the perpetrator considers to cheat against (or for). That is, the third force may make it more likely to reduce dissonance of behaving honestly and forgoing for some entities (e.g., small businesses, ingroup, women), versus behaving dishonestly and compromising moral standards for others (e.g., big businesses, outgroup, men). We discuss this in more detail in section *1.3 Dishonesty across victims.*

Finally, how dishonest do people think others are? Scholars across domains, from economics (Schotter & Trevino, 2014) to political science (Levi, 2022) to sociology (Kluegel & Smith, 1981) to psychology (Coltheart et al., 2011), have made considerable efforts in studying beliefs. Given the extensive study of moral decisions in this dissertation, serendipitously, the large set of studies enabled us to examine one simple question: Do people believe that others are similarly, more, or less dishonest than they truly are? Section *1.4 Moral beliefs about others* will examine this in greater detail.

1.2. Sequential moral behavior

People do not make moral decisions in isolation (Schwabe et al., 2018). Moral behaviors are often influenced by the social context, where the agent's recent behavioral history in the moral domain may play an important role (West & Zhong, 2015). That is, the moral decision at hand may be influenced by *prior* behaviors in the moral domain (Mullen & Monin, 2016). For example, helping a friend move furniture during the weekend may influence if and how much one donates to charity when solicited the next week. Theorizing on sequential moral choices, Huber and colleagues (2008) draw from Max Weber's (1958) contrast between Catholicism and Calvinism: Whereas the Catholic could use good deeds to atone for particular sins, the Calvinist could not compensate for some bad deeds with other good deeds. However, beyond what ought to do, how do everyday people make moral decisions they face sequentially? Does doing something good or bad lead to more of the same or the opposite?

Table 1. Illustrating sequential moral behavior at two points in time. For a review, please see (Mullen & Monin, 2016).		
	Moral behavior at T2	**Immoral behavior at T2**
Moral behavior at T1	Positive moral consistency	Moral self-licensing (Compensatory)
Immoral behavior at T1	Moral cleansing (Compensatory)	Negative moral consistency

Similar to the teachings of Jon Calvin, behavioral consistency theories propose that people are inclined to maintain consistency in their actions (Festinger, 1954; Heider, 1946). Then, if a person makes a moral (immoral) decision at T1, they are more likely to make another moral (immoral) decision at T2 (please see Table 1 for a simplified illustration). This perspective has received considerable support in subsequent work; individuals who invested heavily in failing projects invested even more (Arkes & Blumer, 1985), people who had previously agreed to a smaller request were more likely to agree to a later larger request (Freedman & Fraser, 1966), and people who bought organic products at one point in time were more likely to buy *more* organic products afterward (Juhl et al., 2017).

Conversely, the past two decades have also accumulated a growing body of research on compensatory moral behavior. In this dissertation, compensatory moral behavior refers to when an initial moral or immoral act leads to a subsequent behavior of the opposite moral valence (Mullen & Monin, 2016). Within this framework, moral licensing refers to a decrease in the likelihood of subsequent moral behavior after performing an initial moral behavior (Merritt et al., 2010), while moral cleansing refers to an increase in the likelihood of subsequent moral behavior after performing an initial immoral behavior (Zhong & Liljenquist, 2006).

Several studies show staunch support for moral licensing in several contexts. For example, performing ethical behavior(s) can give individuals a perceived license to behave less ethically afterward (Effron et al., 2009; Effron & Conway, 2015; Monin & Miller, 2001; Mullen & Monin, 2016; Sachdeva et al., 2009). Some notable examples include how expressing disagreement with racist statements (Monin & Miller, 2001) or endorsing Barack Obama (Effron et al., 2009) increased preference for hiring a white person, recalling past moral actions increased future prosocial intentions (Jordan et al., 2011), and buying environmentally friendly products increased likelihood to cheat or steal (Mazar & Zhong, 2010).

Then, what about the mirror opposite of licensing, moral cleansing? It has long been known that performing immoral actions may have negative psychological consequences (Klass, 1978) and can spur the agent to act in ways to restitute their moral self-image (Tetlock et al., 2000). Research shows that recalling past instances of having unsafe sex increased the likelihood of donating to a homeless shelter (Stone et al., 1997), participating in a mock Milgram experiment increased social cooperativeness (Carlsmith & Gross, 1969), and lying to increase payoffs in a deception gain increased donation amounts to charity (Gneezy et al., 2014).

Considering moral licensing and cleansing together highlights how people balance their behaviors to maintain a positive moral self-image (West & Zhong, 2015). Further, the effect may be symmetrical such that past moral deeds can increase the likelihood of being less moral in the future, and vice versa (Jordan et al., 2011; Zhong et al., 2009, 2010). Summarizing this compensatory perspective, Sachdeva and colleagues (2009) proposed that moral behavior can be thought of as "being embedded within a larger system that contains competing forces. Moral or immoral action may emerge from an attempt to find balance among these forces." (p-528).

However, a stream of null findings cast doubts on the robustness of compensatory moral behavior. This includes studies where writing about positive traits (Sachdeva et al., 2009) did not lead to lower donations (Blanken et al., 2014), green consumption (Mazar & Zhong, 2010) did not increase subsequent cheating (Urban et al., 2019), and exposure to organic food (Eskine, 2013) did not reduce altruistic intentions (Moery & Calin-Jageman, 2016). Nonetheless, these null findings may also be interpreted as moral consistency being more likely under some conditions and compensatory moral behavior in others.

Indeed, research examining moderators shows that consistent (compensatory) moral behavior is more likely when people adopt an abstract (concrete) mental construal (Brown et al., 2011; Conway & Peetz, 2012), have a rule-based (outcome-based) ethical orientation (Cornelissen et al., 2013), have a goal commitment (progress) mindset (Susewind & Hoelzl, 2014), gain social recognition of the initial good deed (Susewind & Walkowitz, 2020), focus on the future (past) organizational citizenship behavior (Griep et al., 2021), or have promotion (prevention) regulatory focus (Lalot et al., 2022). This may explain why although two meta-analyses find the classical moral licensing effect to be d = .31 (Blanken et al., 2015) and d = .32 (Simbrunner & Schlegelmilch, 2017), the effect may be overestimated (Kuper & Bott, 2019). A "many-labs" project suggests a more modest effect size of d = .14 (Ebersole et al., 2016). So, the basic question of whether doing something good or bad leads to more of the same or the opposite remains unclear and warrants rigorous empirical examinations.

In that vein, Article 1 in my dissertation is based on the largest single-lab study (N = 5091) examining sequential moral behavior. In the registered report format, we tried to replicate and extend the influential hypothesis by Conway and Peetz (2012) that *conceptual abstraction* moderates if past moral deeds lead to consistent or compensatory behavior. The original hypothesized interaction was *not* replicated: conceptual abstraction did not moderate the effect of recalling moral vs. immoral behavior on prosocial intentions. By adding two neutral conditions, we were able to attribute how the differences after recalling moral or immoral behavior compared to baseline prosocial intentions, overcoming a common shortcoming of most previous studies on sequential moral behavior (Mullen & Monin, 2016). Overall, our results robustly show that recalling moral behavior led to higher prosocial intentions than recalling either immoral or neutral behavior, irrespective of recalling from the recent or distant past. We found no evidence for compensatory moral behavior, only for positive moral consistency.

1.3. Dishonest behavior across victims

While honesty is positively associated with trust, cooperation, and economic prosperity (Gächter & Schulz, 2016), acts of dishonesty can inflict great costs on society. The Federal Bureau of Investigation (2020) estimates the cost of non-health insurance fraud, such as by falsely reporting numbers in applications or claims forms, in the United States at over 40 billion USD. Merchandise returns fraud, that is returning purchased items- often after use- under false premises for refunds, cost US retailers $23.2 billion in 2021 (NRF, Appriss Retail, 2022).

Despite the severe economic costs, factors that drive dishonest behavior remain underexamined. Across the board, interventions to fight dishonesty have shown limited success (Skowronek, 2022); and even those that have shown promise, most were tested *without* unpacking the psychological underpinnings of their effectiveness (Hertwig & Mazar, 2022). As a result, there is a severe gap in our understanding of what *drives* dishonesty, which impedes informed decisions and policies.

Why do people behave dishonestly? Drawing on the classical economic model, Becker's (1968) theory of rational crime predicts that people will behave dishonestly if the expected material benefits exceed the costs. Material benefits may be monetary or other forms of reward, whereas costs may take the form of punitive punishment and/or reputational consequences. According to this rational actor model, the decision-maker would consider the costs and benefits, and their subjective probabilities. In its essence, this perspective simply proposes that people would behave honestly or dishonestly depending on the *incentives* at play.

Although incentives play an important role in driving dishonest behavior, it is now well-established in psychology and behavioral economics that individuals are also motivated by factors beyond their self-interests (Fehr & Fischbacher, 2003; Gibson et al., 2013). A meta-analysis of incentive-compatible “honesty tasks” shows that people cheat far less than the traditional economic model would predict (Abeler et al., 2019). Violating the rational actor model, even when people are presented with opportunities to cheat completely anonymously and without any repercussions, people do *not* cheat to the maximum level- basically leaving money on the table as per the rational actor model (Gerlach et al., 2019; Rosenbaum et al., 2014).

Although experimental research on dishonest behavior (for a comprehensive review, see Gerlach et al., 2019) has examined the effects of incentives (Balasubramanian et al., 2017; Charness et al., 2019; Conrads et al., 2013; Wang & Murnighan, 2017; Wiltermuth, 2011), variations in the environment the decision takes place (Ackert et al., 2011; Ayal et al., 2019; Capraro, 2017; Chou, 2015; Cohn et al., 2015, 2022; Kocher et al., 2018; Leib et al., 2021; Shalvi et al., 2011, 2012), and the characteristics of the perpetrator (Capraro, 2018; Hilbig & Zettler, 2015; Pascual-Ezama et al., 2020; Utikal & Fischbacher, 2013; Vincent et al., 2013), the fundamental characteristics of the *victim* have been mostly overlooked. In the domain of dishonesty, the decision to cheat can be seen as a zero-sum situation, where the perpetrator can lie and increase their payoffs at the *cost* of the victim. Given this dyadic nature of dishonesty, where a perpetrator stands to gain certain benefits at the expense of a victim who loses an equal objective value, overlooking *who* is more (vs.) likely to get cheated would be ignoring half the story.

The influential moral typecasting theory (Gray & Wegner, 2009) provides a useful general framework for understanding how different entities, in case of interactions in the moral domain, are categorized as moral agents or patients. According to the moral typecasting theory, those categorized as agents are perceived more as agentic performers of actions and less as recipients of actions. Conversely, those categorized as patients are perceived more as recipients of actions and to have less agency. This categorization of a moral dyad (Schein & Gray, 2018) is mutually exclusive. That is, in a given interaction, agentic characteristics of an entity can make it easier to see them as a moral agent and less as a moral patient, and vice versa (Gray & Wegner, 2009).

In the case of victims of moral harm- such as being cheated- decision-makers may *typecast* the potential victim as an agentic moral perpetrator or a vulnerable moral victim depending on their salient characteristics. Because of mutual exclusivity in a moral dyad, agentic perpetrators (vulnerable victims) are perceived to be less sensitive to pain and suffering (Reynolds et al., 2020; Shepherd et al., 2019), and so evoke less moral concerns (Dijker, 2010; Gray & Wegner, 2009). To that end, the characteristics of victims that signal their agency or patiency in the moral context can serve as a useful lens to understand *who* gets cheated more and why.

The literature on how victim characteristics affect dishonest behavior is starting to grow. From the early days which relied on field observations and interviews to investigate how people are more willing to cheat some than others (Mars, 1985; Smigel, 1956), recent research leans toward experimental investigations and aims to provide *causal* evidence using behavioral measures. The limited existing research specifically examining *victim* characteristics in dishonest behavior finds that when tempted with the same risk-free gains, individuals are more likely to cheat another person than an organization (Soraperra et al., 2019), an identifiable rather than non-identifiable victim (Yam & Reynolds, 2016), a harmless corporation than a harmful corporation (Rotman et al., 2018), and a brand with a non-cute than a cute logo (Septianto & Kwon, 2021). Although this stream of research suggests that salient characteristics of the potential victim can significantly affect dishonest behavior against them, several important comparisons of *who gets cheated more and why* are yet to be rigorously examined.

Of course, different victim characteristics can also be confounded with factors such as expected costs and/or reputational concerns. For example, cheating an organization with the resources to enforce costly punishment may increase the *expected* costs of cheating. A different way of considering this could be that cheating a resource-abundant organization may be perceived to have low expected costs because the victim organization may not find it worthwhile to enforce punishments. Beyond how cheating different victims- for the same potential benefit- may have different expected costs, can *mere* victim characteristics affect dishonesty against them? That is, if the expected benefits and costs are held constant, can there be systematic differences in dishonest behavior such that some entities get cheated more and others cheat less?

Consider the following model, inspired by Mota (2023), that formalizes how victim characteristics can affect dishonest behavior:

$$D_{pv} = \beta_{op} + \beta_1 I + \beta_2 C_v + \varepsilon_{pv}$$

In this equation:

- D_{pv} represents the decision of perpetrator *p* to cheat victim *v*.
- β_{op} accounts for individual moral character of perpetrator *p*.
- $\beta_1 I$ represents the incentives, i.e., the expected gains minus costs from cheating.
- $\beta_2 C_v$ reflects the internal costs of cheating victim *v* with characteristics *C*.
- ε_{pv} represents the residual effects, covering unobserved variables, measurement errors, randomness, and model specification errors in the decision-making process of perpetrator *p* towards victim *v*.

Assume that the $\beta_1 I_p$ is held constant, with equal potential gains from cheating victims, for example, *v1* and *v2,* under conditions of non-detection, no reputational harm, and no punishment. Then, for the average decision-maker, the primary influencing factor of dishonest behavior would be $\beta_2 C_v$, which is a function of the characteristics of the victim *v*. If the levels of dishonesty systematically vary between two victims with characteristics C_1 and C_2, it would suggest that these characteristics impact the internal cost of dishonesty. For example, higher dishonesty against victim *v1* with characteristics C_1 compared to victim *v2* with characteristics C_2, it would imply systematically lower internal costs of cheating victim *v1* than victim *v2*.

Although the aforementioned conceptualization is simply an abstraction from reality, increasingly, similar situations are faced by everyday people. Whether it is filing for taxes, insurance forms, returning times, and so on, people are often faced with decisions where they can choose to take advantage of private information- that is, information that only they can be aware of- and exploit information asymmetries to benefit dishonestly (Tennyson, 1997; Van Zant & Kray, 2014). For example, what is to stop people from underreporting income without paper trails, exaggerating the value of lost items when filing an insurance claim, or returning items under false premises? Arguably, many of these cases involve virtually zero detection, punishment, and reputational costs, and so beyond the moral character of the decision-maker, factors such as the characteristic*s* of *whom* they are cheating (e.g., a big vs. small businesses, men vs. women, ingroups vs. outgroups) may play an important role. Our minds are tuned to make sense of the slightest bit of available information. So, what characteristics of potential victims are *salient* can shape how we behave towards them.

Although promising in terms of recent advancements, there is much left to be examined in the domain of victim characteristics and dishonesty, the biggest focus of investigation of this dissertation: To what extent does it matter who gets cheated, in influencing dishonest behaviors against them?

In that spirit, Article 2 examines how the mere size of an organization can influence people's dishonest behavior against them. Informed by moral typecasting theory, we suggest people are less likely to perceive big businesses as vulnerable victims than small businesses, making it seem more acceptable to cheat them for selfish gain. We studied this "business-size bias" across eight experiments (*N* = 5,670). First, using a series of scenario-based experiments, we found that people intended to be more dishonest against big than small businesses, mediated by lower vulnerability perceptions. Then, three experiments using established incentive-aligned paradigms from behavioral economics found that people were also more likely to *behave* dishonestly toward big businesses for personal gain. Unlike the general principle of equality under the law, our findings suggest that people operate with at least two different moral standards depending on the size of the organization they are interacting with.

Article 3 explores systematically biased dishonest behavior in an intergroup context, considering both selfish and coalitional lying. As individual decisions naturally occur in a social context, we hypothesized that the acceptability of dishonest behavior may depend on the group identity of whom it affects. We used different adaptations of the "mind game" paradigm to provide anonymous Democrat and Republican voting U.S. American participants an economic incentive to lie without any detection risk, and randomly varied the group identity of the victim (Experiment 1: N = 1,177) and beneficiary (Experiment 2: N = 1,710). We found that although people lied at the same rate for personal gain, irrespective of gaining at the cost of an outgroup or ingroup, people lied at a significantly higher rate to benefit their ingroups when their lying did not affect their outcomes.

Article 5 tested for a *gender bias* in interpersonal dishonesty. Informed by moral typecasting theory, I hypothesized that decision-makers may exhibit a gender bias in dishonesty, adopting different moral standards for dishonest behavior against males and females as victims. A large-scale and incentivized experiment (N = 3,168), with participants recruited from nine different countries tempted to cheat another same-sex, opposite-sex, or sex-unmentioned participant- with complete anonymity, zero reputational risks, and no punishment. The results showed that female targets were cheated 22% less than unmentioned-sex targets. Interestingly, female decision-makers cheated female (vs. male) targets 53.6% less, with no such difference among male decision-makers, suggesting an asymmetrical gender bias in interpersonal dishonesty.

1.4. Moral beliefs about others

People's beliefs about the world significantly shape their expectations and behaviors (Jervis, 2006). People hold beliefs regarding several aspects such as how similar they are to others (Robbins & Krueger, 2005), how favorable others are (Tarrant et al., 2012), and what to attribute the behaviors of others (Hewstone, 1990). From economists (Schotter & Trevino, 2014) to political scientists (Levi, 2022) to sociologists (Kluegel & Smith, 1981) to psychologists (Coltheart et al., 2011), scholars have made considerable efforts in studying beliefs.

How close are people's beliefs to reality? The wisdom of the crowd principle (Surowiecki, 2005) suggests that the collective predictions of a large group of diverse individuals have the potential to align with reality. In 1906, Francis Galton held a weight-judging contest where 787 people entered their guesses regarding the weight of an ox (Galton, 1907). The average guess of all entrants was astonishingly only 1lb less than the actual weight (1,198 lbs.), which far outperformed the guess of the individual winner of the contest, and those of cattle experts (Galton, 1907). In a modern reenactment, NPR (National Public Radio) posted a photo of a cow and asked people to guess its weight. More than 17,000 people responded, and the average guess was off by only 68lbs for the 1,355lb cow-that is only by 5 percent (Bui, 2015).

Indeed, the accuracy of collective judgments has been established across domains. For instance, the collective judgments of a group of radiologists far outperform those of any single radiologist (Wolf et al., 2015). In politics, the wisdom-of-crowds approach also outperforms recognition-based predictions in most instances of election forecasts (Gaissmaier & Marewski, 2011). Even for guessing the class average on an exam, classes that were bigger made more accurate predictions (Blackwell & Pickford, 2011).

However, collective judgments being more accurate than individual judgments does not imply that collective judgments will always mirror reality. This may be especially so concerning beliefs about others. We cannot access other people's minds and so it becomes difficult to be informed about what others perceive and intend to do (Oeberst & Imhoff, 2023). So, our beliefs about others may not coincide with reality. Now, individuals having wrong beliefs may not be so problematic as there is hope that when considered collectively, random errors of those who overestimate and those who underestimate may cancel each other out. Of course, this would not be the case if collective judgments were *systematically skewed* in one direction.

In the moral domain, people's beliefs are often miscalibrated. For instance, both nonexperts and professional economists failed to predict what percentage of supposedly lost wallets would be returned from field experiments across 355 cities in 40 countries (Cohn et al., 2022). Using multiple studies including scenarios experiments, recalled experiences, and live interactions, Zhao and Epley (2022) found that people systematically underestimate the prosociality of others, which impedes individuals from asking strangers for help. In a large-scale study across 60 countries, Mastroianni and Gilbert (2023) found that people *wrongly* believe that morality is declining. In a globally representative study with participants across

125 countries, Andre and colleagues (2024) found that people systematically underestimate how willing their fellow citizens are to enact actions to mitigate climate change.

The aforementioned research suggests that people seem to have a “negativity bias” when it comes to the moral behavior of others (Baumeister et al., 2001; Rozin & Royzman, 2001). In many individual settings, one can point out that the costs of wrongly trusting someone are higher than the costs of wrongly distrusting someone. This is clearly supported by the notion that people are loss averse such that for the same magnitude of a potential loss vs. gain, losses receive double the weight when making decisions under uncertainty (Kahneman & Tversky, 1979). So, it makes intuitive sense that when a potential target considers whether the other party would cheat them or not, erring on the side of having pessimistic beliefs can help in self-preservation.

Then, what about *third-party* beliefs about others’ dishonest behaviors? That is, how do beliefs about dishonest behaviors of others compare to reality when those do not affect those having the beliefs? Here, being overly pessimistic should not lead to any self-preservation advantages. Nevertheless, however close (or far) people’s beliefs may be from reality can influence decisions about the common good and policy support. Believing others to be less dishonest than they actually are can lead to advocating naïve behaviors and policies that may be vulnerable to exploitation by bad actors. Conversely, believing others are more dishonest than they actually are may lead to unnecessary surveillance in economic processes and social interactions.

While inaccurate beliefs about others' dishonesty at the individual level may not be so problematic, if beliefs are systematically skewed at the *aggregate* level, that can shape important marketplace and organizational functions. For example, more and more supermarkets are putting even low-value products such as deodorant, toothpaste, and soap behind lock and key (Meyersohn, 2022). In the workplace, a surge in monitoring employee activities is fueling worker distrust (Christian, 2022). For these phenomena, one can point out that the potential costs to the self of underestimating dishonesty are much higher than that of overestimating dishonesty (Blaine & Boyer, 2018). As a result, people may err on the overestimation side, which explains the prevalence of a negativity bias (Baumeister et al., 2001; Rozin & Royzman, 2001). However, what about beliefs about others' dishonesty as a *third-party*, in contexts where the self is *not* affected by others' actions? This is important to examine because many decisions regarding promoting a more honest market, workplace, and society require enacting policies that do not directly affect the self but are nonetheless shaped by beliefs about others.

Let us consider the argument that a group of individuals collectively may have correct beliefs about the dishonesty of others *on average*. Given each individual is privy to different information, deviations in either direction may cancel each other out. In that vein, the wisdom of the crowd principle (Surowiecki, 2005) posits that collective predictions of a diverse and large group of individuals are usually more accurate than those of individual experiments, and have been shown superior performance in domains such as medical diagnosis (Wolf et al., 2015), stock trading (Blackwell & Pickford, 2011), and election forecasts (Gaissmaier & Marewski, 2011). Based on this reasoning, a large sample of individuals may be able to accurately predict the extent of dishonest behaviors in a population.

However, there are competing streams of theoretical and empirical work that would predict beliefs to be *skewed* in a particular direction. On the one hand, individuals naturally assume honesty in others as a social norm (Yamagishi, 2001), introduced as the concept of “trust default”. Moreover, the truth-default theory (TDT) suggests that people typically operate under a 'truth bias', often overlooking dishonesty in most interactions (Levine, 2014, 2022). While classical research in optimism bias finds that people rate their own chances of positive (negative) events higher (lower) than others (Sharot et al., 2011; Weinstein, 1980), research also shows that people also overestimate the likelihood of positive events than negative events not only for the self but also for the general population (Dricu et al., 2022). This overestimation (underestimation) of positive (negative) events has been operationalized as *social optimism bias* (Aue et al., 2021)- a positive view of the world that is associated with experiencing positive affect in general (Fox, 2012). This suggests people might often think others are more honest than they really are.

On the other hand, people exhibit a tendency to overestimate the sheer frequency of a range of behaviors (e.g., smoking marijuana, getting drunk, attending religious services, etc.), with the authors speculating it being driven by the fact that behavior is more salient than non-behavior (Nisbett & Kunda, 1985). In fact, when people were asked about how they would act in a moral dilemma, participants in the moral behavior condition cheated significantly less than participants in a forecasting condition predicted they themselves would cheat (Teper et al., 2011). Therefore, it may be so that people similarly overestimate the dishonesty of others, even when the dishonesty does not affect them.

Another possibility may be that there is systematic heterogeneity among people regarding which direction their beliefs are skewed in. It has long been established that people use how they would behave in a particular situation to infer how others would act in the same situation (Ross et al., 1977). Indeed, this form of social projection (Ames, 2004) has been found to predict levels of expected cooperative and prosocial behavior (Fischer, 2009; Krueger, 2007, 2013). Individuals scoring high in Honesty-Humility- defined as the "tendency to be fair and genuine in dealing with others" (Ashton & Lee, 2007)- tend to project their own prosociality and trustworthiness onto other people (Pfattheicher & Böhm, 2018; Thielmann & Hilbig, 2014). With regards to negative expectations, individuals who lie more often tend to perceive others as less honest, a phenomenon known as 'deceiver's distrust' (Sagarin et al., 1998). A study on academic integrity among students found that business school students who self-reported cheating more also believed that their peers were likely to cheat a lot as well (Chapman et al., 2004). Even among adolescents, those who cheated on tests were more likely to believe that their peers would have done the same (Evans & Lee, 2014). In the domain of romantic relationships, people's rate of lying in mobile dating was positively correlated with the perceived lying rate of their partners (Markowitz & Hancock, 2018). Taken together, if an individual is honest (dishonest), they might project their inclination onto others and erroneously believe that this honesty (dishonesty) is shared more widely among others, thus underpredicting (overpredicting) the prevalence of dishonesty in the general population.

Amidst competing theoretical and empirical perspectives, Article 4 asks a simple question: Do people believe that others are similarly, more, or less dishonest than they truly are? In a pilot study, we asked 376 USA-based participants, "Out of 100 people in our Zip code, how many may try to deceive X", where X was replaced by a range of firms (e.g., Meta, Google, Walmart, etc.), with 10 randomly selected entities per participant. A random-effects naïve model showed that participants believed a significant[1] number of others (M = 22.8, SD = 23.2) in their Zip code would try to deceive firms. This suggests that people believe that nearly 1 in 4 people would cheat on average. Nonetheless, this type of survey data does *not* tell us how accurate or inaccurate these beliefs are.

In a research program on moral decision-making spanning three years (2022-24), we placed participants in different situations where they could lie for personal gain, without any

[1] $t(375) = 27.4$, $p < .001$, Cohen's $d = 0.98$

repercussions or detection risk. We also asked all participants to estimate what percentage of other people would lie in a similar situation. The experiments together produced a total of 31 different effects (combined N = 8,094). Meta-analysis of these experiments, including both incentivized choice experiments and hypothetical marketplace scenarios that were initially designed to test a broad collection of different hypotheses, revealed a significant overestimation of others' dishonesty, by an average of 14 percentage points. That is, people are substantially less dishonest than they are thought to be. The findings reveal a pervasive tendency to overestimate the actual rate of dishonesty among other people, suggesting a widespread belief that the world is less moral than it actually is.

1.5. Methodological approach

When I started my PhD, in August 2020, it would be fair to say that doubts started to be cast on some of the underpinnings in moral decision-making research. Whereas popular science books such as *Thinking Fast and Slow*, *Predictably Irrational*, and *Nudge: Improving decisions about health, wealth, and happiness* continued capturing the public imagination, there were big debates in the behavioral science community about a “replicability crisis” (Maxwell et al., 2015; Shrout & Rodgers, 2018) and/or “credibility revolution” (Vazire, 2018). Specifically, a combination of now dubbed “questionable practices” in research, such as having low statistical power of detecting the effect size of interest (Stanley et al., 2018), cherry-picking data points to publish what are essentially “false positives” (Simmons et al., 2011), and only reporting studies that support the proposed hypotheses (Ferguson & Heene, 2012), led to several key findings in social science failing to replicate when the seminal study was rerun.

In the moral domain, seminal findings once published in the PNAS such as how signing at the beginning of a form increases honest subsequent reporting of private information did not replicate in large-scale efforts (Kristal et al., 2020). Reports of null results of nudge interventions to increase honest behavior kept coming from several places and contexts, including field experiments on tax compliance from the UK (John & Blume, 2018) to Guatemala (Kettle et al., 2017) to experiments on reducing insurance fraud in the Nordics (Martuza et al., 2022).

Even outside nudges to promote honest behavior, reports of one of the key findings from moral psychology in the last two decades, moral licensing, started to have failed replications (Rotella & Barclay, 2020) and raised concerns of publication bias (Kuper & Bott, 2019). Further, several propositions of the influential construal level theory were argued as not as robust when a rigorous meta-analysis found robust evidence for publication bias and overestimation of the effect sizes (Maier et al., 2022). Finally, the past five years of controversies surrounding research in the moral domain, with key papers being retracted, and the credibility of leading scholars being investigated, has only exacerbated the uncertainty in what we know with confidence versus what we cannot say for sure.

Taken together, one could conclude that studying moral decisions may be a risky topic for a Ph.D. dissertation- given it is unclear how stable the foundations are. That is, if one aims to build on a particular theory, what if the evidence in support of that theory is not robust? However, I saw this as a potential to *start fresh* with as few priors as possible. So, I often relied on testing basic intuitions empirically before experimentally testing hypotheses.

Allow me to illustrate with an example. One of the first testable research ideas I had was whether people are more likely to cheat big than small businesses. It made intuitive sense, and the literature seemed to suggest several explanations of why people may do so.

That is, one could start with a confirmatory research strategy, testing a specific hypothesis that causally linked the variables of interest (business size and dishonest behavior against them) from the beginning. Nonetheless, I adopted an exploratory approach at first and simply *asked* hundreds of participants across multiple pilot studies if they would be more likely to cheat big than small businesses across several industries. At that time, I thought that was a wasteful thing to do, asking people about something quite obvious, and generating data that perhaps does not add new knowledge to how we see the world. However, after a year and a half of those pilots, I learned that that is one way of stimulus sampling, and both my advisors stated it was a valid approach to start projects.

To that end, my usual "research strategy" involved testing basic intuitions empirically before hypothesizing causal links. This is partially because I do not trust how I may think of relationships (or lack thereof) between variables in my mind to represent how most people may behave in the world. I come from a middle-class family in the "developing world" but had the privilege to live in several OECD countries with scholarships. Currently, I am on the verge of completing a PhD degree. So, my lived experiences, which shape my priors and thinking in ways I may not be aware of, are quite far from the reality of everyday people. So, I may be quite susceptible to making false assumptions about people, reiterating the need for "proof of concept" pilots before experimental studies.

We now turn to the method that has been the workhorse for all articles in this dissertation- experiments. Why? The notion that academics refrain from sweeping statements and hedge their wordings has been around for quite a while. Nonetheless, I want to tell my friends and family that X causes Y. The beauty of experiments is that they bring us as close as possible to being able to make such statements. In an experimental study, the experimenter introduces a controlled variation in X (the independent variable), while holding all else constant, and observes for differences (or lack thereof) in Y (the dependent variable). As a result, experiments enable researchers to test for causal links between variables of interest, and try to advance theory and knowledge based on the presence or absence of those links.

In social science domains spanning psychology, marketing, and even political science, online experiments on platforms with credible pools of participants have been the workhorse in testing theories. Conducting experiments this way allows us to conduct large studies cost-effectively. All studies in my dissertation were conducted on Prolific (prolific.com), an online research participant recruiting platform built specifically for academic research (Buhrmester et al., 2018). Compared to other platforms, Prolific has more stringent pre-screening for participants (Palan & Schitter, 2018), and greater naivety among participants (Peer et al., 2017) than MTurk. Further, Prolific has shown superior performance in terms of participant attention, comprehension, honesty, and reliability compared to MTurk (Peer et al., 2022). Taken together, although conducting experiments online rather than in the lab leads to the loss of some experimenter control, using reputable platforms to recruit participants substantially increases data quality and thereby enhances the reliability and validity of the conclusions drawn.

Regarding reliability and validity, all experiments in the articles that comprise my dissertation have 90% statistical power to detect small-to-medium-sized and/or even small-sized main effects. High statistical power to detect our effect sizes of interest enables us to place more confidence in the results, regardless of detecting the presence or absence of hypothesized effects. Further, for experiments where we failed to find statistically significant support for the hypothesis, we reported equivalence tests to illustrate if and how conclusive the null effect is (Lakens, 2017). Contrary to how traditional hypothesis testing examines significant differences, an equivalence test examines if the difference between two conditions is smaller than a pre-specified smallest effect size of interest (SESOI). In these tests, the null hypothesis is specified as the true difference between the conditions being greater in magnitude (that is either above the upper bound or below the lower bound) than the SESOI. The alternative hypothesis is specified as the true differences lying between zero and the corresponding upper and/or lower bound, hence suggesting an *equivalence* between the groups being compared- in other words, statistical equivalence.

Article 2 in the dissertation perhaps makes a good example of conducting original research *reliably*. The article comprises eight experiments in total- that have been directly and/or conceptually replicated (Schmidt, 2009) at least once, suggesting high internal reliability as similar results were produced each time for the same outcome. Similarly, Article 4, although containing effects from a range of studies conducted for different purposes, the robust meta-analytic effect bolsters the proposition that people tend to think others are more dishonest than they actually are. Regarding cross-cultural generalizability, Article 5 is the only multi-country study- recruiting a balanced sample of participants from nine countries. Although there were no meaningful country-level differences, cell sizes at the country level do not have sufficient power to detect the effect size found in the pooled sample.

Another common methodological thread in my articles (except Article 1) is measuring *actual behavior*. Early in my PhD, I often wondered how much can be said about people's moral decisions by only measuring people's *intentions*. What incentive do people have to report their true intentions? Even more so, a meta-analysis studying the intention-behavior gap found that intentions translate into behavior only about half the time (Sheeran & Webb, 2016). Furthermore, in the moral domain, people have miscalibrated predictions when asked about what they would do in a particular moral decision, which has been conceptualized as a “moral forecasting error” (Teper et al., 2011). This necessitates finding ways to measure actual behavior, when possible, in studying moral decisions.

So, in most of my studies included in the dissertation, I have tried to incentivize decisions. For instance, in the studies in Articles 2, 3, and 5, participants received actual bonuses when they cheated. In several of the studies included in Article 4, participants were incentivized to report their best estimates as we awarded participants with bonuses if their estimates about others’ dishonesty were close to the observed rates of actual dishonesty. Incentive alignment was something I learned about during my master’s thesis on conjoint experiments. In many of the economics circles, *revealed* preferences are usually preferred compared to stated preferences. That is, the expectation is that what people actually *do* says more about their preferences than what they *say* they would.

Although I did not incorporate incentive alignment in Article 1, I borrowed heavily from experimental economics to design incentivized dishonesty measures. To that end, in my projects of *who* is more likely to be cheated (e.g., big vs. small businesses, men vs. women, outgroup vs. ingroup) I used variations of the mind game paradigm (Fischbacher & Föllmi-Heusi, 2013; Jiang, 2013). That is, I tempted participants with real money to cheat a particular victim, anonymously under conditions of no detection, punishments, or reputational consequences. The goal was to set up clear tradeoffs between the financial incentive to cheat and increase material gains vs. the moral motive to behave honestly. Furthermore, honesty in the lab, that is the tasks I use, has also been shown to predict actual cheating behavior (Dai et al., 2018), suggesting some degree of external validity.

Now, if we are to adopt the rational actor perspective (Becker, 1968), we should really not see any differences in cheating. Because if the benefits of cheating are held constant and all costs are eliminated, why would people, on average, cheat X more than Y? Beyond perfectly rational decision-making, are there systematic victim-based differences in which- X or Y- is more acceptable to cheat? Again, the goal of my dissertation has been to conduct foundational research on moral decision-making, testing hypotheses in controlled settings which often may not be possible to test in the real world without big industry collaborations and over a longer timeframe than possible given the time constraints of a PhD. So, Articles 2, 3, and 5 are meant to test basic propositions on how human behavior may vary across contexts, which may give us some insights into possible things at play in the real world.

The final notable methodological common thread across my dissertation has been pre-registering (Nosek et al., 2018) main hypotheses, sample sizes, and planned analyses *prior* to data collection. While these practices of conducting research more rigorously, transparently, and collaboratively sharing data and study materials are becoming increasingly common in social science, I was also lucky to have conducted my PhD as part of a research group that practiced Open Science diligently. To that end, all data, and study materials from the conducted experiments part of my dissertation have already been shared or will be shared soon in corresponding dedicated folders on the Open Science Framework (OSF) website (osf.io). Nonetheless, the degree to which there were deviations in planned analyses, varies across projects. This is because I was "learning by doing", and so the pre-registered analysis for a few of my earlier studies in Article 2 was not always feasible.

Beyond trying to implement Open Science principles as much as possible, this has also helped me bring more discipline to my research: Can I succinctly state my hypotheses and planned analyses before actually running the study? Further, incorporating two registered reports (Chambers & Tzavella, 2021) into my dissertation has enabled me to receive feedback from expert reviewers which substantially helped improve the study and interpret results both at the study level and how those change our broader body of knowledge. In the registered report format, researchers can submit manuscripts without having conducted the empirical study, usually studying research questions that the field benefits from regardless of the results. That is, conducting research in the registered report format research usually involves examining something where finding “no support” for the hypothesis may be just as useful as finding support. Another personal point to acknowledge is that registered reports freed me from financial constraints in designing large-scale studies, knowing my department is more likely to fund “high financial cost” studies *with* an “in-principle acceptance” (IPA) from a leading journal- because of assurance that the research will be published if conducted as planned regardless of the results.

All in all, large samples, incentive compatibility, and open science principles (Open Science Collaboration, 2015) have been the north star of my methods in this dissertation. I have been domain agnostic in getting inspiration for my basic research on how people make moral decisions. From scenario-based experiments in moral psychology to cheating paradigms in decision-making research, I have let the research question be the guide in what methods I use to test my hypotheses. Whereas doing the same thing and expecting different results is widely considered madness, replicating the same study and finding similar results- in my humble opinion, and leading scholars in the field- is a sign of *trying* to do credible science (Zwaan et al., 2018). All in all, to the best of my knowledge and within my constraints, I tried to use as rigorous a method as possible in my dissertation.

1.6. Contributions

This dissertation contributes to advancing knowledge on how the context affects moral decision-making. Specifically, the articles cover how moral decisions may systematically vary across the self, others, and beliefs about others. As a whole, all articles contribute new empirical insights, advancing debates on how the context shapes people's decisions in the moral domain.

In Article 1, we conducted the largest single-lab investigation (N = 5,091) in trying to replicate and extend Conway and Peetz's (2012) seminal hypothesis that recalling moral or immoral deeds from the recent (distant) past should lead to compensatory (consistent) moral behavior. Our results did *not* support conceptual abstraction being a moderator. Instead, we found a robust positive moral consistency effect: Recalling past moral behavior (versus immoral or neutral behavior), regardless of whether from the recent or distant past, increased prosocial intentions. This suggests a silver lining of sorts that "feeling moral" at one point in time may not inadvertently make people behave less prosocially in subsequent decisions. Evidence for a positive moral consistency effect bodes well societally as doing good may lead to virtuous cycles. This may be practically relevant for charitable and volunteering organizations, as individuals consistently doing good can help them and society at large.

The results from Article 1 also challenge the existing theories on compensatory moral behavior and highlight the need for reevaluation in this domain. Given these findings, it may also be important to reconsider the robustness of other hypothesized moderators in sequential moral behavior (e.g., Brown et al., 2011; Griep et al., 2021; Lalot et al., 2022; Susewind & Hoelzl, 2014; Susewind & Walkowitz, 2020). Rigorously testing the factors that may moderate whether doing good or bad leads to similar or opposite behavior afterward is crucial not only in moral psychology but also in domains such as organizational (List & Momeni, 2021) and consumer (Juhl et al., 2017) behavior. To that end, the debate around sequential moral behavior, given its widespread importance, may remain active for the foreseeable future.

In Article 2, we conducted a series of experiments (combined N = 5,670) to study how the mere size of a business may affect dishonest behavior against them. We found that people are more likely to cheat big businesses than small businesses, across both scenario-based and incentivized experiments. Informed by moral typecasting theory (Gray & Wegner, 2009), we also present evidence that one important explanation for this *business-size bias* in dishonesty, is that people perceive large organizations as less vulnerable and less moral than small organizations, which makes it seem more justifiable to cheat them. Although some indications from observational field studies (Mars, 1985), interviews (Smigel, 1956), and surveys (Rotman et al., 2018) exist in the extant literature, Article 2, to our knowledge, is the first to both systematically provide causal evidence, and use actual behavioral measures to examine the business-size bias in consumer dishonest behavior.

Broadly, and in contrast to the general principle of equality under the law, Article 2 suggests that when people make consequential decisions regarding telling the truth or lying, their moral cost of dishonesty is not fixed, but is partly a function of the *size* of the organization their decisions affect. These findings add to the nascent literature on how *victim characteristics* can affect dishonest behavior (Rotman et al., 2018; Soraperra et al., 2019; Yam & Reynolds, 2016). Further, we propose and present evidence of *organizational* moral typecasting, showing how people's tendencies to categorize the social world into agents and patients go beyond individuals, and into the organizational context.

Article 2 also offers practical implications both for organizations and policymakers. Large and growing organizations might be facing a moral "size penalty" in the marketplace, by attracting dishonesty from consumers than they would have as a small business. In the face of mixed findings regarding the effectiveness of interventions to promote honesty (Hertwig & Mazar, 2022), a better understanding of inherent biases in consumer dishonesty can help design more effective strategies. Further, small businesses might actually have an advantage vis-à-vis marketplace morality, as consumers seem to be more reluctant to steal or lie to them for personal gain. This may imply that small businesses may do well to trust their consumers more, and divert resources devoted to surveillance measures to more revenue-generating measures.

Article 3 examined two primary questions using incentivized pre-registered experiments: (1) Are people more likely to cheat outgroups than ingroups for personal gain? (2) Under conditions of no personal gain, are people more likely to cheat when their ingroup than their outgroups receive the gain? In Experiment 1, we found that individuals lied at the same rate for personal gain, irrespective of costing an outgroup or ingroup member. In Experiment 1, where the decision-maker's own outcomes were removed from the equation, individuals lied at a significantly higher rate to benefit their ingroups.

Our main result from Experiment 1, that political group identities did not affect cheating another person for personal gain, is surprising. If there was an intergroup bias in dishonesty for personal gain, we would have been likely to find it a setting with intergroup animosity as high as between Democrats and Republicans (Kranton et al., 2020). Findings from Experiment 2 add to the previously suggestive evidence from prior research in experimental economics (Aksoy & Palma, 2019; Cadsby et al., 2016; Michailidou & Rotondi, 2019) that people may be more likely to cheat to benefit their ingroup than their outgroup. This suggests that when personal benefits are eliminated, the group identity of the beneficiary seems to influence the "pure" moral costs of cheating, adding a group psychology dimension to the literature on altruistic lying (Brocas & Carrillo, 2021; Erat & Gneezy, 2012). Practically, findings from the current research paint a mixed picture. In the current age of challenges for intergroup cooperation and societal polarization, it seems that the group dimension may not affect egotistic lying, but it can exacerbate other-benefitting dishonest behavior. So, although in negotiations, transactions, and other economic interactions, individuals may not cheat someone more just because the potential victim is an outgroup, being willing to actively lie to benefit ingroups merely by symbolism, presents a warning for organizational leaders, regulators, and auditors.

Article 4's main finding, that people systematically overestimate the dishonesty of others, challenges the 'wisdom of crowds' approach (Galton, 1907; Surowiecki, 2005; Van Dolder & Van Den Assem, 2017) to estimate moral behavior from beliefs. Rather, our findings contribute to the negativity bias literature (Baumeister et al., 2001; Rozin & Royzman, 2001) by demonstrating its significant role in shaping social perceptions of dishonesty. Practically, the pervasive overestimation of others' dishonesty may imply the need for trust-building interventions, such as information campaigns or educational programs, aimed at correcting these misperceptions to foster a more trustful society. Organizations may benefit from reconsidering policies enacted from mistrust and excessive surveillance, for example for remote/hybrid work and expense accounts.

Finally, Article 5's robust finding that female decision-makers cheated other females less than they cheated other males, with no such difference among male decision-makers, suggests an asymmetrical gender bias in interpersonal dishonesty. Further, self-reported measures showed that participants paired with a male (vs. female) participant expressed less guilt if they were to cheat their pair, were more likely to think their pair expected to be cheated, and also harbored greater expectation that their pair would have cheated them. Together, this contributes to a gendered view of moral typecasting (Reynolds et al., 2020) and gender dynamics in dishonesty, (Capraro, 2018; Childs, 2012; Grosch & Rau, 2017; Kastlunger et al., 2010; Ward & King, 2018) wherein, people may find it more acceptable to cheat males than females.

The dissertation's key theoretical contribution is increasing the understanding of how victim characteristics can affect dishonest behavior against them. Drawing on the influential theory of moral typecasting (Gray & Wegner, 2009), articles in this dissertation highlight how *who* is being affected may play a role in the moral acceptability of dishonesty. Further, Articles 2-5 used incentivized economic experiments to study psychological phenomena, which makes behavior-based contribution to understanding systematic differences in moral decisions. The tight experimental controls allowed us to precisely test for differences in dishonest behavior in the absence of detection, punishment, or reputational risks- enabling us to attribute differences to differences in the "pure" moral cost of cheating A more than B.

1.7. Limitations and future research

The findings and conclusions drawn in this dissertation must be considered in light of several limitations. First, articles in this dissertation primarily used a confirmatory hypothesis-testing approach, which limits the breadth of findings that could have been attained had a qualitative approach been used. Although research in moral psychology has mostly used experiments (Malle, 2021), coupled with how most people think intuitively about it (Haidt, 2001), it may nonetheless be interesting to interview people with innovative techniques to gain richer insights into people's minds, which may generate a larger number of hypotheses as well.

Second, except for Article 5, our interpretations are based on empirical findings from online experiments conducted mostly in one cultural setting- the USA. This makes cross-cultural generalizability versus variability interesting questions for future research. Specifically, it may be interesting to test if and how the size of the business-size bias and compensatory vs. consistent moral behavior effects may vary across cultures. Even more so, if culture-based *directional* changes are found, it can significantly contribute to our understanding of decisions in the moral domain.

Third, because all experiments were conducted in controlled settings, using online survey experiments, external validity cannot be claimed with confidence. Although we present behavioral evidence and research that shows that moral behavior in the lab predicts real-world behavior (Dai et al., 2018), it remains unclear if *systematic differences* in moral behavior from the lab, that is experimental effects, translate to similar differences in marketplace or organizational behavior. Future work may benefit from also examining real-world base rates by collaborating with firms. For example, examining if large supermarket chains are inflicted with greater retail theft than smaller chains and individual stores can provide robust generalizability of the business-size bias.

Fourth, although our incentivized paradigms (Articles 2, 3, 4, and 5) designed for anonymous decision-making offer insights into how people make moral decisions under conditions of zero detection, punishment, or reputational consequences, external validity may be challenged by the fact these aforementioned factors do <u>not</u> exist in the real-world marketplace and interpersonal decision-making contexts. The specific concern here would be if there are interaction effects that may potentially moderate the effects we found. For example, contrary to our findings in Article 2 of people cheating big businesses more than small businesses in controlled settings, the greater power to enforce litigation by big businesses may influence people to be more fearful of cheating big businesses. On the other hand, small businesses being more woven into the social fabric of decision-makers may actually amplify the size of the effects in Article 2. In the same way, although we found evidence that females are cheated significantly less in Article 5, this may very well reverse in real-world situations where people may be more wary of potential retribution by male than female victims.

Finally, results from mediation and moderation analyses presented across articles in this dissertation have to be interpreted with caution. With regards to mediation analyses in Article 2, although the sample sizes achieved sufficient statistical power to test the statistical significance of potential mediators, our interpretations regarding the process were made without *experimentally* manipulating levels of the mediator, which is increasingly becoming the gold standard (Pirlott & MacKinnon, 2016). This perhaps suggests a meaningful next study for Article 2, which is trying to provide evidence/mechanism of our hypothesized effects experimentally.

A specific limitation to consider for Article 4 is that the effect sizes included in the internal meta-analysis were produced from a range of studies testing a range of hypotheses, and conducted as part of the same research program (my dissertation). Most of the included effects were not pre-registered, and the meta-analysis itself was not pre-registered. In fact, the idea for the project came serendipitously from exploring the properties of different datasets across projects. Nonetheless, follow-up pre-registered studies testing our phenomena of interest bolster our confidence because we found effects in the same direction.

With regards to interpreting moderation analyses, such as in Article 3, power calculations were based on detecting minimum effect sizes of interest vis-à-vis our hypothesized main effects, and so the significance and/or non-significance of moderators we tested and discussed have to be seen in terms of lower statistical power. Nonetheless, because moderators were peripheral to our main hypotheses, future work may find it fruitful to probe moderators with sufficiently powered samples.

However, moderation results from Article 5, which were core to the preregistered hypothesis, had sufficient power to detect a “small” effect size in its respective statistical test. Nonetheless, a set of equivalence tests showed that when the main hypothesis was not supported, the null evidence was nonconclusive- suggesting we cannot rule out the absence of the minimum effect size of interest. One clear observation was that the levels of dishonesty were not particularly high. This may be attributed to our repeated measures paradigm, which although intended to give more precise estimates at the participant level, future research may benefit from one-shot paradigms as there may have been "calibration" by participants to not appear to have cheated "too much".

All in all, how people make moral decisions has grabbed interest from scholars and laypeople alike. In that vein, this dissertation examines systematic differences in people's moral decisions- examining intentions, actual behavior, and beliefs. Informed by psychological theories and using incentivized economic experiments to test theories, articles in this dissertation, despite the limitations, pave the way for future research for interdisciplinary research in moral decisions.

1.8. References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for Truth-Telling. *Econometrica*, *87*(4), 1115–1153. https://doi.org/10.3982/ECTA14673

Ackert, L. F., Church, B. K., Kuang, X. (Jason), & Qi, L. (2011). Lying: An Experimental Investigation of the Role of Situational Factors. *Business Ethics Quarterly*, *21*(4), 605–632. https://doi.org/10.5840/beq201121438

Aksoy, B., & Palma, M. A. (2019). The effects of scarcity on cheating and in-group favoritism. *Journal of Economic Behavior & Organization*, *165*, 100–117. https://doi.org/10.1016/j.jebo.2019.06.024

Alan Fine, G. (2019). Moral Cultures, Reputation Work, and the Politics of Scandal. *Annual Review of Sociology*, *45*(1), 247–264. https://doi.org/10.1146/annurev-soc-073018-022649

Ames, D. R. (2004). Strategies for Social Inference: A Similarity Contingency Model of Projection and Stereotyping in Attribute Prevalence Estimates. *Journal of Personality and Social Psychology*, *87*(5), 573–585. https://doi.org/10.1037/0022-3514.87.5.573

Andre, P., Boneva, T., Chopra, F., & Falk, A. (2024). Globally representative evidence on the actual and perceived support for climate action. *Nature Climate Change*. https://doi.org/10.1038/s41558-024-01925-3

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Ashton, M. C., & Lee, K. (2007). Empirical, Theoretical, and Practical Advantages of the HEXACO Model of Personality Structure. *Personality and Social Psychology Review*, *11*(2), 150–166. https://doi.org/10.1177/1088868306294907

Aue, T., Dricu, M., Moser, D. A., Mayer, B., & Bührer, S. (2021). Comparing personal and social optimism biases: Magnitude, overlap, modifiability, and links with social identification and expertise. *Humanities and Social Sciences Communications*, *8*(1), 233. https://doi.org/10.1057/s41599-021-00913-8

Ayal, S., Celse, J., & Hochman, G. (2019). Crafting messages to fight dishonesty: A field investigation of the effects of social norms and watching eye cues on fare evasion. *Organizational Behavior and Human Decision Processes*, S0749597818305004. https://doi.org/10.1016/j.obhdp.2019.10.003

Balasubramanian, P., Bennett, V. M., & Pierce, L. (2017). The wages of dishonesty: The supply of cheating under high-powered incentives. *Journal of Economic Behavior & Organization*, *137*, 428–444. https://doi.org/10.1016/j.jebo.2017.03.022

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is Stronger than Good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. In N. G. Fielding, A. Clarke, & R. Witt (Eds.), *The Economic Dimensions of Crime* (pp. 13–68). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-62853-7_2

Blackwell, C., & Pickford, R. (2011). The wisdom of the few or the wisdom of the many? An indirect test of the marginal trader hypothesis. *Journal of Economics and Finance*, *35*(2), 164–180. https://doi.org/10.1007/s12197-009-9092-4

Blaine, T., & Boyer, P. (2018). Origins of sinister rumors: A preference for threat-related material in the supply and demand of information. *Evolution and Human Behavior*, *39*(1), 67–75. https://doi.org/10.1016/j.evolhumbehav.2017.10.001

Blanken, I., van de Ven, N., & Zeelenberg, M. (2015). A Meta-Analytic Review of Moral Licensing. *Personality and Social Psychology Bulletin*, *41*(4), 540–558. https://doi.org/10.1177/0146167215572134

Blanken, I., van de Ven, N., Zeelenberg, M., & Meijers, M. H. C. (2014). Three Attempts to Replicate the Moral Licensing Effect. *Social Psychology*, *45*(3), 232–238. https://doi.org/10.1027/1864-9335/a000189

Brocas, I., & Carrillo, J. D. (2021). Self-serving, altruistic and spiteful lying in the schoolyard. *Journal of Economic Behavior & Organization*, *187*, 159–175. https://doi.org/10.1016/j.jebo.2021.04.024

Brown, R. P., Tamborski, M., Wang, X., Barnes, C. D., Mumford, M. D., Connelly, S., & Devenport, L. D. (2011). Moral Credentialing and the Rationalization of Misconduct. *Ethics & Behavior*, *21*(1), 1–12. https://doi.org/10.1080/10508422.2011.537566

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An Evaluation of Amazon's Mechanical Turk, Its Rapid Rise, and Its Effective Use. *Perspectives on Psychological Science*, *13*(2), 149–154. https://doi.org/10.1177/1745691617706516

Bui, Q. (2015). 17,205 People Guessed The Weight Of A Cow. Here's How They Did. *Planet Money*. https://www.npr.org/sections/money/2015/08/07/429720443/17-205-people-guessed-the-weight-of-a-cow-heres-how-they-did

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

Campbell, M. C., & Winterich, K. P. (2018). A Framework for the Consumer Psychology of Morality in the Marketplace. *Journal of Consumer Psychology*, *28*(2), 167–179. https://doi.org/10.1002/jcpy.1038

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., & Tungodden, B. (2007). The Pluralism of Fairness Ideals: An Experimental Approach. *American Economic Review*, *97*(3), 818–827. https://doi.org/10.1257/aer.97.3.818

Capraro, V. (2017). Does the truth come naturally? Time pressure increases honesty in one-shot deception games. *Economics Letters*, *158*, 54–57. https://doi.org/10.1016/j.econlet.2017.06.015

Capraro, V. (2018). Gender differences in lying in sender-receiver games: A meta-analysis. *Judgment and Decision Making*, *13*(4), 345–355. https://doi.org/10.1017/S1930297500009220

Carlsmith, J. M., & Gross, A. E. (1969). Some effects of guilt on compliance. *Journal of Personality and Social Psychology*, *11*(3), 232–239. https://doi.org/10.1037/h0027039

Chambers, C. D., & Tzavella, L. (2021). The past, present and future of Registered Reports. *Nature Human Behaviour*, *6*(1), 29–42. https://doi.org/10.1038/s41562-021-01193-7

Chapman, K. J., Davis, R., Toy, D., & Wright, L. (2004). Academic Integrity in the Business School Environment: I'll Get by with a Little Help from My Friends. *Journal of Marketing Education*, *26*(3), 236–249. https://doi.org/10.1177/0273475304268779

Charness, G., Blanco-Jimenez, C., Ezquerra, L., & Rodriguez-Lara, I. (2019). Cheating, incentives, and money manipulation. *Experimental Economics*, *22*(1), 155–177. https://doi.org/10.1007/s10683-018-9584-1

Chou, E. Y. (2015). Paperless and Soulless: E-signatures Diminish the Signer's Presence and Decrease Acceptance. *Social Psychological and Personality Science*, *6*(3), 343–351. https://doi.org/10.1177/1948550614558841

Christian, A. (2022). The employee surveillance that fuels worker distrust. *BBC Worklife*. https://www.bbc.com/worklife/article/20220621-the-employee-surveillance-that-fuels-worker-distrust

Cohn, A., Gesche, T., & Maréchal, M. A. (2022). Honesty in the Digital Age. *Management Science*, *68*(2), 827–845. https://doi.org/10.1287/mnsc.2021.3985

Cohn, A., Maréchal, M. A., & Noll, T. (2015). Bad Boys: How Criminal Identity Salience Affects Rule Violation. *The Review of Economic Studies*, *82*(4), 1289–1308. https://doi.org/10.1093/restud/rdv025

Cohn, A., Maréchal, M. A., Tannenbaum, D., & Zünd, C. L. (2019). Civic honesty around the globe. *Science*, *365*(6448), 70–73. https://doi.org/10.1126/science.aau8712

Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional Belief. *Annual Review of Psychology*, *62*(1), 271–298. https://doi.org/10.1146/annurev.psych.121208.131622

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

Conway, P., & Peetz, J. (2012). When Does Feeling Moral Actually Make You a Better Person? Conceptual Abstraction Moderates Whether Past Moral Deeds Motivate Consistency or Compensatory Behavior. *Personality and Social Psychology Bulletin*, *38*(7), 907–919. https://doi.org/10.1177/0146167212442394

Cornelissen, G., Bashshur, M. R., Rode, J., & Le Menestrel, M. (2013). Rules or Consequences? The Role of Ethical Mind-Sets in Moral Dynamics. *Psychological Science*, *24*(4), 482–488. https://doi.org/10.1177/0956797612457376

Crockett, M. J. (2016). How Formal Models Can Illuminate Mechanisms of Moral Judgment and Decision Making. *Current Directions in Psychological Science*, *25*(2), 85–90. https://doi.org/10.1177/0963721415624012

Crumpler, H., & Grossman, P. J. (2008). An experimental test of warm glow giving. *Journal of Public Economics*, *92*(5–6), 1011–1021. https://doi.org/10.1016/j.jpubeco.2007.12.014

Curry, O. S. (2016). Morality as Cooperation: A Problem-Centred Approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The Evolution of Morality* (pp. 27–51). Springer International Publishing. https://doi.org/10.1007/978-3-319-19671-8_2

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is It Good to Cooperate? Testing the Theory of Morality-as-Cooperation in 60 Societies. *Current Anthropology*, *60*(1), 47–69. https://doi.org/10.1086/701478

Dai, Z., Galeotti, F., & Villeval, M. C. (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Management Science*, *64*(3), 1081–1100. https://doi.org/10.1287/mnsc.2016.2616

De Dreu, C. K. W., Fariña, A., Gross, J., & Romano, A. (2022). Prosociality as a foundation for intergroup conflict. *Current Opinion in Psychology*, *44*, 112–116. https://doi.org/10.1016/j.copsyc.2021.09.002

Dijker, A. J. M. (2010). Perceived vulnerability as a common basis of moral emotions. *British Journal of Social Psychology*, *49*(2), 415–423. https://doi.org/10.1348/014466609X482668

Dricu, M., Moser, D. A., & Aue, T. (2022). Optimism bias and its relation to scenario valence, gender, sociality, and insecure attachment. *Scientific Reports*, *12*(1), 18534. https://doi.org/10.1038/s41598-022-22031-4

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., … Nosek, B. A. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, *67*, 68–82. https://doi.org/10.1016/j.jesp.2015.10.012

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites. *Journal of Experimental Social Psychology*, *45*(3), 590–593. https://doi.org/10.1016/j.jesp.2009.02.001

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology*, *6*, 32–35. https://doi.org/10.1016/j.copsyc.2015.03.017

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

Eskine, K. J. (2013). Wholesome Foods and Wholesome Morals?: Organic Foods Reduce Prosocial Behavior and Harshen Moral Judgments. *Social Psychological and Personality Science*, *4*(2), 251–254. https://doi.org/10.1177/1948550612447114

Evans, A. D., & Lee, K. (2014). The relation between 8- to 17-year-olds' judgments of other's honesty and their own past honest behaviors. *International Journal of Behavioral Development*, *38*(3), 277–281. https://doi.org/10.1177/0165025413517580

Feess, E., Feld, J., & Noy, S. (2021). People Judge Discrimination Against Women More Harshly Than Discrimination Against Men – Does Statistical Fairness Discrimination Explain Why? *Frontiers in Psychology*, *12*, 675776. https://doi.org/10.3389/fpsyg.2021.675776

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785–791. https://doi.org/10.1038/nature02043

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Festinger, L. (1957). *A theory of cognitive dissonance* (Reissued by Stanford Univ. Press in 1962, renewed 1985 by author, [Nachdr.]). Stanford Univ. Press.

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

Fischer, I. (2009). Friend or foe: Subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General*, *138*(3), 341–350. https://doi.org/10.1037/a0016073

Forbes, R. C., & Stellar, J. E. (2022). When the ones we love misbehave: Exploring moral processes within intimate bonds. *Journal of Personality and Social Psychology*, *122*(1), 16–33. https://doi.org/10.1037/pspa0000272

Fox, E. (2012). *Rainy brain, sunny brain: The new science of optimism and pessimism*. William Heinemann.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*(2), 195–202. https://doi.org/10.1037/h0023552

Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, *531*(7595), 496–499. https://doi.org/10.1038/nature17160

Gaissmaier, W., & Marewski, J. N. (2011). Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls. *Judgment and Decision Making*, *6*(1), 73–88. https://doi.org/10.1017/S1930297500002102

Galton, F. (1907). Vox Populi. *Nature*, *75*(1949), 450–451. https://doi.org/10.1038/075450a0

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, *145*(1), 1–44. https://doi.org/10.1037/bul0000174

Gibson, R., Tanner, C., & Wagner, A. F. (2013). Preferences for Truthfulness: Heterogeneity among and within Individuals. *American Economic Review*, *103*(1), 532–548. https://doi.org/10.1257/aer.103.1.532

Gigerenzer, G. (2010). Moral Satisficing: Rethinking Moral Behavior as Bounded Rationality. *Topics in Cognitive Science*, *2*(3), 528–554. https://doi.org/10.1111/j.1756-8765.2010.01094.x

Gneezy, U., Imas, A., & Madarász, K. (2014). Conscience Accounting: Emotion Dynamics and Social Behavior. *Management Science*, *60*(11), 2645–2658. https://doi.org/10.1287/mnsc.2014.1942

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520. https://doi.org/10.1037/a0013748

Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. The Penguin Press.

Greene, J. D. (2015). The rise of moral cognition. *Cognition*, *135*, 39–42. https://doi.org/10.1016/j.cognition.2014.11.018

Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523. https://doi.org/10.1016/S1364-6613(02)02011-9

Griep, Y., Germeys, L., & Kraak, J. M. (2021). Unpacking the Relationship Between Organizational Citizenship Behavior and Counterproductive Work Behavior: Moral Licensing and Temporal Focus. *Group & Organization Management*, *46*(5), 819–856. https://doi.org/10.1177/1059601121995366

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, *316*(5827), 998–1002. https://doi.org/10.1126/science.1137651

Halevy, N., Bornstein, G., & Sagiv, L. (2008). "In-Group Love" and "Out-Group Hate" as Motives for Individual Participation in Intergroup Conflict: A New Game Paradigm. *Psychological Science*, *19*(4), 405–411. https://doi.org/10.1111/j.1467-9280.2008.02100.x

Hare, B. (2017). Survival of the Friendliest: *Homo sapiens* Evolved via Selection for Prosociality. *Annual Review of Psychology*, *68*(1), 155–186. https://doi.org/10.1146/annurev-psych-010416-044201

Heider, F. (1946). Attitudes and Cognitive Organization. *The Journal of Psychology*, *21*(1), 107–112. https://doi.org/10.1080/00223980.1946.9917275

Hertwig, R., & Mazar, N. (2022). Toward a taxonomy and review of honesty interventions. *Current Opinion in Psychology*, *47*, 101410. https://doi.org/10.1016/j.copsyc.2022.101410

Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology*, *20*(4), 311–335. https://doi.org/10.1002/ejsp.2420200404

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

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

Huber, J., Goldsmith, K., & Mogilner, C. (2008). Reinforcement versus balance response in sequential choice. *Marketing Letters*, *19*(3–4), 229–239. https://doi.org/10.1007/s11002-008-9042-5

Jervis, R. (2006). Understanding Beliefs. *Political Psychology*, *27*(5), 641–663. https://doi.org/10.1111/j.1467-9221.2006.00527.x

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, *93*, 328–336. https://doi.org/10.1016/j.jebo.2013.04.003

Jihwan Chae, Kunil Kim, Yuri Kim, Gahyun Lim, Daeeun Kim, & Hackjin Kim. (2022). Ingroup favoritism overrides fairness when resources are limited. *Scientific Reports*. https://doi.org/10.1038/s41598-022-08460-1

John, P., & Blume, T. (2018). How best to nudge taxpayers? The impact of message simplification and descriptive social norms on payment rates in a central London local authority. *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.10

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the Moral Self: The Effects of Recalling Past Moral Actions on Future Moral Behavior. *Personality and Social Psychology Bulletin*, *37*(5), 701–713. https://doi.org/10.1177/0146167211400208

Juhl, H. J., Fenger, M. H. J., & Thøgersen, J. (2017). Will the Consistent Organic Food Consumer Step Forward? An Empirical Analysis. *Journal of Consumer Research*, *44*(3), 519–535. https://doi.org/10.1093/jcr/ucx052

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Keltner, D., Kogan, A., Piff, P. K., & Saturn, S. R. (2014). The Sociocultural Appraisals, Values, and Emotions (SAVE) Framework of Prosociality: Core Processes from Gene to Meme. *Annual Review of Psychology*, *65*(1), 425–460. https://doi.org/10.1146/annurev-psych-010213-115054

Kettle, S., Hernandez, M., Sanders, M., Hauser, O., & Ruda, S. (2017). Failure to CAPTCHA Attention: Null Results from an Honesty Priming Experiment in Guatemala. *Behavioral Sciences*, *7*(4), 28. https://doi.org/10.3390/bs7020028

Klass, E. T. (1978). Psychological effects of immoral actions: The experimental evidence. *Psychological Bulletin*, *85*(4), 756–771. https://doi.org/10.1037/0033-2909.85.4.756

Kluegel, J. R., & Smith, E. R. (1981). Beliefs about Stratification. *Annual Review of Sociology*, *7*(1), 29–56. https://doi.org/10.1146/annurev.so.07.080181.000333

Kocher, M. G., Schudy, S., & Spantig, L. (2018). I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. *Management Science*, *64*(9), 3995–4008. https://doi.org/10.1287/mnsc.2017.2800

Kranton, R., Pease, M., Sanders, S., & Huettel, S. (2020). Deconstructing bias in social preferences reveals groupy and not-groupy behavior. *Proceedings of the National Academy of Sciences*, *117*(35), 21185–21193. https://doi.org/10.1073/pnas.1918952117

Kristal, A. S., Whillans, A. V., Bazerman, M. H., Gino, F., Shu, L. L., Mazar, N., & Ariely, D. (2020). Signing at the beginning versus at the end does not decrease dishonesty. *Proceedings of the National Academy of Sciences*, *117*(13), 7103–7107. https://doi.org/10.1073/pnas.1911695117

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology*, *18*(1), 1–35. https://doi.org/10.1080/10463280701284645

Krueger, J. I. (2013). Social Projection as a Source of Cooperation. *Current Directions in Psychological Science*, *22*(4), 289–294. https://doi.org/10.1177/0963721413481352

Kuper, N., & Bott, A. (2019). Has the evidence for moral licensing been inflated by publication bias? *Meta-Psychology*, *3*. https://doi.org/10.15626/MP.2018.878

Kurzban, R., Burton-Chellew, M. N., & West, S. A. (2015). The Evolution of Altruism in Humans. *Annual Review of Psychology*, *66*(1), 575–599. https://doi.org/10.1146/annurev-psych-010814-015355

Lakens, D. (2017). Equivalence Tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Lalot, F., Falomir-Pichastor, J. M., & Quiamzade, A. (2022). Regulatory focus and self-licensing dynamics: A motivational account of behavioural consistency and balancing. *Journal of Environmental Psychology*, *79*, 101731. https://doi.org/10.1016/j.jenvp.2021.101731

Leib, M., Köbis, N., Soraperra, I., Weisel, O., & Shalvi, S. (2021). Collaborative dishonesty: A meta-analytic review. *Psychological Bulletin*, *147*(12), 1241–1268. https://doi.org/10.1037/bul0000349

Levi, M. (2022). The Power of Beliefs. *Annual Review of Political Science*, *25*(1), 1–19. https://doi.org/10.1146/annurev-polisci-051120-013517

Levine, T. R. (2014). Truth-Default Theory (TDT): A Theory of Human Deception and Deception Detection. *Journal of Language and Social Psychology*, *33*(4), 378–392. https://doi.org/10.1177/0261927X14535916

Levine, T. R. (2022). Truth-default theory and the psychology of lying and deception detection. *Current Opinion in Psychology*, *47*, 101380. https://doi.org/10.1016/j.copsyc.2022.101380

List, J. A., & Momeni, F. (2021). When Corporate Social Responsibility Backfires: Evidence from a Natural Field Experiment. *Management Science*, *67*(1), 8–21. https://doi.org/10.1287/mnsc.2019.3540

Maier, M., Bartoš, F., Oh, M., Wagenmakers, E.-J., Shanks, D., & Harris, A. J. L. (2022). *Adjusting for Publication Bias Reveals That Evidence for and Size of Construal Level Theory Effects is Substantially Overestimated* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/r8nyu

Malle, B. F. (2021). Moral Judgments. *Annual Review of Psychology*, *72*(1), 293–318. https://doi.org/10.1146/annurev-psych-072220-104358

Markowitz, D. M., & Hancock, J. T. (2018). Deception in Mobile Dating Conversations. *Journal of Communication*, *68*(3), 547–569. https://doi.org/10.1093/joc/jqy019

Mars, G. (1985). Cheats at Work: The Anthropology of Workplace Crime. *Anthropology of Work Review*, *6*(4), 44–47. https://doi.org/10.1525/awr.1985.6.4.44

Martuza, J. B., Skard, S. R., Løvlie, L., & Thorbjørnsen, H. (2022). Do honesty-nudges really work? A large-scale field experiment in an insurance context. *Journal of Consumer Behaviour*, *21*(4), 927–951. https://doi.org/10.1002/cb.2049

Mastroianni, A. M., & Gilbert, D. T. (2023). The illusion of moral decline. *Nature*, *618*(7966), 782–789. https://doi.org/10.1038/s41586-023-06137-x

Mattingly, C., & Throop, J. (2018). The Anthropology of Ethics and Morality. *Annual Review of Anthropology*, *47*(1), 475–492. https://doi.org/10.1146/annurev-anthro-102317-050129

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist*, *70*(6), 487–498. https://doi.org/10.1037/a0039400

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

Mazar, N., & Zhong, C.-B. (2010). Do Green Products Make Us Better People? *Psychological Science*, *21*(4), 494–498. https://doi.org/10.1177/0956797610363538

Meyersohn, N. (2022). Why Old Spice, Colgate and Dawn are locked up at drug stores. *CNN Business*. https://edition.cnn.com/2022/07/30/business/drug-stores-locked-products/index.html

Michailidou, G., & Rotondi, V. (2019). I'd lie for you. *European Economic Review*, *118*, 181–192. https://doi.org/10.1016/j.euroecorev.2019.05.014

Moery, E., & Calin-Jageman, R. J. (2016). Direct and Conceptual Replications of Eskine (2013): Organic Food Exposure Has Little to No Effect on Moral Judgments and Prosocial Behavior. *Social Psychological and Personality Science*, *7*(4), 312–319. https://doi.org/10.1177/1948550616639649

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*(1), 33–43. https://doi.org/10.1037/0022-3514.81.1.33

Mota, P. I. S. (2023). *Essays on Unethical Behaviour* [Norwegian School of Economics]. https://hdl.handle.net/11250/3090298

Mullen, E., & Monin, B. (2016). Consistency Versus Licensing Effects of Past Moral Behavior. *Annual Review of Psychology*, *67*(1), 363–385. https://doi.org/10.1146/annurev-psych-010213-115120

Nisbett, R. E., & Kunda, Z. (1985). Perception of social distributions. *Journal of Personality and Social Psychology*, *48*(2), 297–311. https://doi.org/10.1037/0022-3514.48.2.297

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Oeberst, A., & Imhoff, R. (2023). Toward Parsimony in Bias Research: A Proposed Common Framework of Belief-Consistent Information Processing for a Set of Biases. *Perspectives on Psychological Science*, 174569162211481. https://doi.org/10.1177/17456916221148147

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Paharia, N., Keinan, A., Avery, J., & Schor, J. B. (2011). The Underdog Effect: The Marketing of Disadvantage and Determination through Brand Biography. *Journal of Consumer Research*, *37*(5), 775–790. https://doi.org/10.1086/656219

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Pascual-Ezama, D., Prelec, D., Muñoz, A., & Gil-Gómez de Liaño, B. (2020). Cheaters, Liars, or Both? A New Classification of Dishonesty Profiles. *Psychological Science*, *31*(9), 1097–1106. https://doi.org/10.1177/0956797620929634

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Pfattheicher, S., & Böhm, R. (2018). Honesty-humility under threat: Self-uncertainty destroys trust among the nice guys. *Journal of Personality and Social Psychology*, *114*(1), 179–194. https://doi.org/10.1037/pspp0000144

Pirlott, A. G., & MacKinnon, D. P. (2016). Design approaches to experimental mediation. *Journal of Experimental Social Psychology*, *66*, 29–38. https://doi.org/10.1016/j.jesp.2015.09.012

Pizarro, D. A., & Tannenbaum, D. (2012). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil.* (pp. 91–108). American Psychological Association. https://doi.org/10.1037/13091-005

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57–75. https://doi.org/10.1037/a0021867

Reynolds, T., Howard, C., Sjåstad, H., Zhu, L., Okimoto, T. G., Baumeister, R. F., Aquino, K., & Kim, J. (2020). Man up and take it: Gender bias in moral typecasting. *Organizational Behavior and Human Decision Processes*, *161*, 120–141. https://doi.org/10.1016/j.obhdp.2020.05.002

Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3

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

Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

Rotella, A., & Barclay, P. (2020). Failure to replicate moral licensing and moral cleansing in an online experiment. *Personality and Individual Differences*, *161*, 109967. https://doi.org/10.1016/j.paid.2020.109967

Rotman, J. D., Khamitov, M., & Connors, S. (2018). Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically. *Journal of Consumer Psychology*, *28*(2), 353–361. https://doi.org/10.1002/jcpy.1002

Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning Saints and Saintly Sinners: The Paradox of Moral Self-Regulation. *Psychological Science*, *20*(4), 523–528. https://doi.org/10.1111/j.1467-9280.2009.02326.x

Sagarin, B. J., Rhoads, K. V. L., & Cialdini, R. B. (1998). Deceiver’s Distrust: Denigration as a Consequence of Undiscovered Deception. *Personality and Social Psychology Bulletin*, *24*(11), 1167–1176. https://doi.org/10.1177/01461672982411004

Schein, C., & Gray, K. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Schmidt, S. (2009). Shall we Really do it Again? The Powerful Concept of Replication is Neglected in the Social Sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Schotter, A., & Trevino, I. (2014). Belief Elicitation in the Laboratory. *Annual Review of Economics*, *6*(1), 103–128. https://doi.org/10.1146/annurev-economics-080213-040927

Schwabe, M., Dose, D. B., & Walsh, G. (2018). Every Saint has a Past, and Every Sinner has a Future: Influences of Regulatory Focus on Consumers' Moral Self-Regulation. *Journal of Consumer Psychology*, *28*(2), 234–252. https://doi.org/10.1002/jcpy.1025

Sen, A., Deaton, A., & Besley, T. (2020). Economics with a Moral Compass? Welfare Economics: Past, Present, and Future. *Annual Review of Economics*, *12*(1), 1–21. https://doi.org/10.1146/annurev-economics-020520-020136

Septianto, F., & Kwon, J. (2021). Too cute to be bad? Cute brand logo reduces consumer punishment following brand transgressions. *International Journal of Research in Marketing*, S016781162100118X. https://doi.org/10.1016/j.ijresmar.2021.12.006

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

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. https://doi.org/10.1038/nn.2949

Sheeran, P., & Webb, T. L. (2016). The Intention–Behavior Gap. *Social and Personality Psychology Compass*, *10*(9), 503–518. https://doi.org/10.1111/spc3.12265

Shepherd, S., Kay, A. C., & Gray, K. (2019). Military veterans are morally typecast as agentic but unfeeling: Implications for veteran employment. *Organizational Behavior and Human Decision Processes*, *153*, 75–88. https://doi.org/10.1016/j.obhdp.2019.06.003

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, *69*(1), 487–510. https://doi.org/10.1146/annurev-psych-122216-011845

Simbrunner, P., & Schlegelmilch, B. B. (2017). Moral licensing: A culture-moderated meta-analysis. *Management Review Quarterly*, *67*(4), 201–225. https://doi.org/10.1007/s11301-017-0128-0

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simon, H. A. (1990). Invariants of Human Behavior. *Annual Review of Psychology*, *41*(1), 1–20. https://doi.org/10.1146/annurev.ps.41.020190.000245

Skowronek, S. E. (2022). DENIAL: A Conceptual Framework to Improve Honesty Nudges. *Current Opinion in Psychology*, 101456. https://doi.org/10.1016/j.copsyc.2022.101456

Smigel, E. O. (1956). Public Attitudes Toward Stealing as Related to the Size of the Victim Organization. *American Sociological Review*, *21*(3), 320. https://doi.org/10.2307/2089287

Soraperra, I., Weisel, O., & Ploner, M. (2019). Is the victim *Max (Planck)* or *Moritz* ? How victim type and social value orientation affect dishonest behavior. *Journal of Behavioral Decision Making*, *32*(2), 168–178. https://doi.org/10.1002/bdm.2104

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin*, *144*(12), 1325–1346. https://doi.org/10.1037/bul0000169

Stone, J., Wiegand, A. W., Cooper, J., & Aronson, E. (1997). When exemplification fails: Hypocrisy and the motive for self-integrity. *Journal of Personality and Social Psychology*, *72*(1), 54–65. https://doi.org/10.1037/0022-3514.72.1.54

Surowiecki, J. (2005). *The wisdom of crowds* (Nachdr.). Anchor Books.

Susewind, M., & Hoelzl, E. (2014). A matter of perspective: Why past moral behavior can sometimes encourage and other times discourage future moral striving: A matter of perspective. *Journal of Applied Social Psychology*, *44*(3), 201–209. https://doi.org/10.1111/jasp.12214

Susewind, M., & Walkowitz, G. (2020). Symbolic Moral Self-Completion – Social Recognition of Prosocial Behavior Reduces Subsequent Moral Striving. *Frontiers in Psychology*, *11*, 560188. https://doi.org/10.3389/fpsyg.2020.560188

Tarrant, M., Branscombe, N. R., Warner, R. H., & Weston, D. (2012). Social identity and perceptions of torture: It’s moral when we do it. *Journal of Experimental Social Psychology*, *48*(2), 513–518. https://doi.org/10.1016/j.jesp.2011.10.017

Tenbrunsel, A. E., & Smith-Crowe, K. (2008). 13 Ethical Decision Making: Where We’ve Been and Where We’re Going. *Academy of Management Annals*, *2*(1), 545–607. https://doi.org/10.5465/19416520802211677

Tennyson, S. (1997). Economic institutions and individual ethics: A study of consumer attitudes toward insurance fraud. *Journal of Economic Behavior & Organization*, *32*(2), 247–265. https://doi.org/10.1016/S0167-2681(96)00904-3

Teper, R., Inzlicht, M., & Page-Gould, E. (2011). Are We More Moral Than We Think?: Exploring the Role of Affect in Moral Behavior and Moral Forecasting. *Psychological Science*, *22*(4), 553–558. https://doi.org/10.1177/0956797611402513

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, *78*(5), 853–870. https://doi.org/10.1037/0022-3514.78.5.853

Thielmann, I., & Hilbig, B. E. (2014). Trust in me, trust in you: A social projection account of the link between personality, cooperativeness, and trustworthiness expectations. *Journal of Research in Personality*, *50*, 61–65. https://doi.org/10.1016/j.jrp.2014.03.006

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual Review of Psychology*, *64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A Person-Centered Approach to Moral Judgment. *Perspectives on Psychological Science*, *10*(1), 72–81. https://doi.org/10.1177/1745691614556679

Urban, J., Bahník, Š., & Kohlová, M. B. (2019). Green consumption does not make people cheat: Three attempts to replicate moral licensing effect due to pro-environmental behavior. *Journal of Environmental Psychology*, *63*, 139–147. https://doi.org/10.1016/j.jenvp.2019.01.011

Utikal, V., & Fischbacher, U. (2013). Disadvantageous lies in individual decisions. *Journal of Economic Behavior & Organization*, *85*, 108–111. https://doi.org/10.1016/j.jebo.2012.11.011

Van Dolder, D., & Van Den Assem, M. J. (2017). The wisdom of the inner crowd in three large natural experiments. *Nature Human Behaviour*, *2*(1), 21–26. https://doi.org/10.1038/s41562-017-0247-6

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

Vazire, S. (2018). Implications of the Credibility Revolution for Productivity, Creativity, and Progress. *Perspectives on Psychological Science*, *13*(4), 411–417. https://doi.org/10.1177/1745691617751884

Vincent, L. C., Emich, K. J., & Goncalo, J. A. (2013). Stretching the Moral Gray Zone: Positive Affect, Moral Disengagement, and Dishonesty. *Psychological Science*, *24*(4), 595–599. https://doi.org/10.1177/0956797612458806

Wang, L., & Murnighan, J. K. (2017). How Much Does Honesty Cost? Small Bonuses Can Motivate Ethical Behavior. *Management Science*, *63*(9), 2903–2914. https://doi.org/10.1287/mnsc.2016.2480

Weber, M. (1958). *The protestant ethic and the spirit of capitalism, translated by Talcott Parsons*. Scribner.

Weidman, A. C., Sowden, W. J., Berg, M. K., & Kross, E. (2020). Punish or Protect? How Close Relationships Shape Responses to Moral Violations. *Personality and Social Psychology Bulletin*, *46*(5), 693–708. https://doi.org/10.1177/0146167219873485

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*(5), 806–820. https://doi.org/10.1037/0022-3514.39.5.806

West, C., & Zhong, C.-B. (2015). Moral cleansing. *Current Opinion in Psychology*, *6*, 221–225. https://doi.org/10.1016/j.copsyc.2015.09.022

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

Wolf, M., Krause, J., Carney, P. A., Bogart, A., & Kurvers, R. H. J. M. (2015). Collective Intelligence Meets Medical Decision-Making: The Collective Outperforms the Best Radiologist. *PLOS ONE*, *10*(8), e0134269. https://doi.org/10.1371/journal.pone.0134269

Yam, K. C., & Reynolds, S. J. (2016). The Effects of Victim Anonymity on Unethical Behavior. *Journal of Business Ethics*, *136*(1), 13–22. https://doi.org/10.1007/s10551-014-2367-5

Yamagishi, T. (2001). Trust as a form of social intelligence. In *Trust in society.* (pp. 121–147). Russell Sage Foundation.

Zhao, X., & Epley, N. (2022). Surprisingly Happy to Have Helped: Underestimating Prosociality Creates a Misplaced Barrier to Asking for Help. *Psychological Science*, *33*(10), 1708–1731. https://doi.org/10.1177/09567976221097615

Zhong, C.-B., Ku, G., Lount, R. B., & Murnighan, J. K. (2010). Compensatory Ethics. *Journal of Business Ethics*, *92*(3), 323–339. https://doi.org/10.1007/s10551-009-0161-6

Zhong, C.-B., Liljenquist, K., & Cain, D. M. (2009). Moral self-regulation: Licensing and compensation. In *Psychological perspectives on ethical behavior and decision making.* (pp. 75–89). Information Age Publishing, Inc.

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*, e120. https://doi.org/10.1017/S0140525X17001972

Chapter 2. Article 1: Does Conceptual Abstraction Moderate Whether Past Moral Deeds Motivate Consistency or Compensatory Behavior? A Registered Replication and Extension of Conway and Peetz (2012)

Jareef Martuza[1], Olivia Kim[1]

[1]Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, Bergen, Norway

Author notes

Correspondence concerning this article should be sent to Jareef Martuza, Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, Bergen 5045. Email: jareef.martuza@nhh.no. Phone: +47 9251 3344

We (the authors) have no conflicts of interest to disclose. This manuscript has not been published elsewhere and is not under consideration by another journal. Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. As Norwegian laws and regulations do not require review by an institutional review board for anonymous, non-medical, low-risk research with human participants, we did not submit the project to such a review. All codes, data, and study materials from the two planned experiments are publicly available on the Open Science Framework (OSF) repository here.

We thank the editor Michael D. Robinson and three reviewers for detailed feedback during both Stage 1 and Stage 2. The comments and suggestions helped us a lot to improve the manuscript both before and after we conducted the study. We also thank Paul Conway for sharing materials from the original study which enabled us to conduct a close replication.

We received funding for the experiment from the Distribusjonsøkonomiske forskningsprosjekter (Distribution economics research project) thanks to Professor Sven Arne Haugland. The funding did not influence the study or our interpretations of the results. We also thank Professor Helge Thorbjørnsen for his feedback during the early stages of the project, and participants of the 2023 Johan Arndt Conference.

This is a forthcoming paper that has been accepted by the Personality and Social Psychology Bulletin. It is not the "copy of record", and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without the authors' permission. The final article will be available, upon publication.

Abstract

A long-standing debate in psychology concerns whether doing something good or bad leads to more of the same or the opposite. Conway and Peetz (2012) proposed that *conceptual abstraction* moderates if past moral deeds lead to consistent or compensatory behavior. Although cited 384 times across disciplines, we did not find any direct replications. It was also unclear how increases or decreases from one's baseline prosociality might underlie the effect. A large-scale experiment (N=5091) in the registered report format tested Conway and Peetz’s original hypothesis. The hypothesized interaction was *not* replicated: conceptual abstraction did not moderate the effect of recalling moral vs. immoral behavior on prosocial intentions. Our results show that recalling moral behavior led to higher prosocial intentions than recalling either immoral or neutral behavior, irrespective of recalling from the recent or distant past. Thus, the current research found no evidence for compensatory moral behavior, only for positive moral consistency.

Keywords: conceptual abstraction; moral consistency; moral licensing; moral cleansing.

Does Conceptual Abstraction Moderate Whether Past Moral Deeds Motivate Consistency or Compensatory Behavior? A Registered Replication and Extension of Conway and Peetz (2012)

Does doing something good or bad lead to more of the same or the opposite? Behavioral consistency theories propose that individuals tend to maintain a steady course of action (Festinger, 1954; Heider, 1946). For example, individuals who heavily invested in failing projects doubled down and invested more (Arkes & Blumer, 1985), people who had previously agreed to a small request were more likely to agree to a larger one later (Freedman & Fraser, 1966), and shoppers who brought their own bags to the store purchased more organic products (Karmarkar & Bollinger, 2015). Conversely, the *moral licensing* literature proposes that engaging in ethical behavior can give individuals a perceived license to act less ethically afterward (Effron et al., 2009; Effron & Conway, 2015; Monin & Miller, 2001; Mullen & Monin, 2016). For example, when individuals expressed disagreement with racist statements (Monin & Miller, 2001) or endorsed Barack Obama, they subsequently displayed a preference for hiring a white person (Effron et al., 2009), individuals recalling past moral actions showed higher future prosocial intentions (Jordan et al., 2011), and individuals who purchased environmentally friendly products subsequently acted less altruistically and were more likely to cheat or steal (Mazar & Zhong, 2010).

To reconcile the debate between sequential moral consistency and compensatory behavior, several studies have examined potential moderating factors determining *when* each is more likely to occur. Research shows compensatory (consistency) moral behavior is more likely when people adopt a concrete (abstract) mental construal (Brown et al., 2011; Conway & Peetz, 2012), are oriented toward an outcome-based (rule-based) ethical mindset (Cornelissen et al., 2013), have a goal progress (commitment) mindset (Susewind & Hoelzl, 2014), (do not) feel social recognition of the initial good behavior (Susewind & Walkowitz, 2020), focus on the past (future) organizational citizenship behavior (Griep et al., 2021), or have prevention (promotion) regulatory focus (Lalot et al., 2022). This stream of research suggests that the moral consistency vs. compensatory behavior debate may not be about *which* is more likely but *when* is each more likely.

In that vein, Conway and Peetz (2012) hypothesized that *conceptual abstraction* moderates whether past moral or immoral behavior leads to consistency or compensatory behavior. Specifically, Conway and Peetz's (2012) seminal Study 1 (hereafter C&P) found that recalling moral/immoral behavior in the distant past leads to consistency, whereas recalling moral/immoral behavior in the recent past leads to compensatory behavior. C&P informed later studies across several areas: including prosocial time giving (Reed et al., 2016), energy conservation (Tiefenbeck et al., 2013), workplace incivility (Rosen et al., 2016), CSR and employee misconduct (List & Momeni, 2021), and pro-environmental behavior (van der Werff et al., 2014). While C&P's Studies 2 and 3 also supported their main hypothesis, most other researchers who investigated moderators refer to the findings of C&P's Study 1- recalling distant vs. recent past behavior- when contextualizing their own results (Gholamzadehmir et al., 2019; Lalot et al., 2022; Susewind & Hoelzl, 2014).

Although cited 384 times across disciplines since publication in the *Personality and Social Psychology Bulletin*, to the best of our knowledge, no direct replications of C&P's hypothesized moderator of conceptual abstraction existed. This proposed moderator was based on construal level theory (Trope & Liberman, 2003, 2010). Considering the recent findings that highlight publication bias and the tendency to overestimate the effects of construal level theory, as well as calls for high-powered registered replications (Maier et al., 2022), it is crucial to also revisit the role of conceptual abstraction (based on temporal construal) as a *moderator*. Moreover, given the lack of control conditions in C&P, in which participants would have recalled *neutral* behavior, it remains unclear how the differences after recalling moral or immoral behavior can be attributed to an increase and/or decrease from one's *baseline* prosocial intentions, a common shortcoming of existing literature on sequential moral behavior (Mullen & Monin, 2016).

Given the general need to independently revisit key psychological findings using current standards (Open Science Collaboration, 2015) and improve understanding of sequential moral behavior with recommended best practices (Blanken et al., 2015; Effron & Conway, 2015; Mullen & Monin, 2016), we conducted a registered replication and extension of Conway and Peetz's (2012) Study 1. We followed the general advice for using larger samples (Chambers & Tzavella, 2021) in replications, resulting in, to our knowledge, the largest single-lab study in the sequential moral behavior literature. Furthermore, we added two control conditions to delineate how the effects of recalling past moral and immoral behavior can increase and/or decrease prosocial intentions from baselines.

Replicating and Extending Conway and Peetz's (2012) Study 1

Conway and Peetz (2012) proposed that level of *conceptual abstraction* can moderate whether past moral behavior leads to moral consistency or compensatory behavior. They posited that moral behaviors from the distant past may be construed in more abstract terms, motivating individuals to act consistently with their salient moral identity (Blasi, 1980; Reed et al., 2007). Conversely, moral behaviors from the recent past, being more concretely construed, might lead individuals to feel they have made sufficient progress toward their moral goals, leading them to engage in compensatory behavior. Indeed, Conway and Peetz's (2012) findings revealed that participants who recalled moral (immoral) behavior from the *recent* past reported lower (higher) prosocial intentions, suggesting compensatory moral behavior. Conversely, participants who recalled moral (immoral) behavior from the *distant* past reported higher (lower) prosocial intentions, suggesting consistent moral behavior.

Our main motivations for the current replication and extension were two-fold. First, despite the widespread influence of Conway and Peetz's Study 1, we found no independent direct replications in the literature. While related studies exist, such as Rotella and Barclay (2020) and Griep et al. (2021), their methods varied significantly from C&P. Rotella and Barclay (2020) conducted an online experiment that did not replicate either moral licensing or moral cleansing. This is not a direct replication of C&P because it had a three-condition design *without* any temporal specifications (recent vs. distant past) to the behaviors participants were asked to recall. Nonetheless, Rotella and Barclay's (2020) inclusion of a control condition (recalling neutral behavior) inspired our design and instructions. Additionally, although Griep et al. (2021) examined the moderating role of *temporal focus*, they measured individual differences in how people focus on their past, present, and future. So, their investigation cannot be directly compared to *experimentally* manipulated levels of conceptual abstraction as in C&P.

Further, recent null findings cast doubt on the generalizability and robustness of compensatory moral behavior altogether. For example, writing about positive traits (Sachdeva et al., 2009) did not lead to lower donations (Blanken et al., 2014), green consumption (Mazar & Zhong, 2010) did not increase subsequent cheating (Urban et al., 2019), and exposure to organic food (Eskine, 2013) did not reduce altruistic intentions (Moery & Calin-Jageman, 2016). However, another way of interpreting these null findings may be that compensatory (consistent) moral behavior may be more likely in *some* conditions than others. As C&P found, compensatory (consistent) moral behavior was more likely to manifest under concrete (abstract) conceptualization of the initial behavior. So, a large-scale replication can bolster our confidence in this key moderator.

Second, it is an open question if the compensatory moral behavior found by C&P was driven by moral licensing, cleansing, or both. In this paper, compensatory moral behavior refers to when an initial moral or immoral act leads to a subsequent behavior of the opposite moral valence (Mullen & Monin, 2016). Within this framework, moral licensing refers to a decrease in subsequent moral behavior following an initial moral behavior (Merritt et al., 2010), while moral cleansing refers to an increase in subsequent moral behavior following an initial immoral behavior (Zhong & Liljenquist, 2006).

In their study, C&P asked participants to recall and write about past moral or immoral behavior from the recent or distant past, and then measured prosocial intentions. C&P's main hypothesis was that recalling behaviors from the recent (distant) past should lead to compensatory (consistent) moral behavior. However, without any baseline (recalling neutral behavior) in the recent past conditions, it is unclear if prosocial intentions distinctly *decreased* after recalling moral behavior (licensing) or *increased* after past recalling immoral behavior (cleansing). As moral licensing and cleansing can rely on different psychological processes, neutral conditions can avoid *conflating* one with the other (Mullen & Monin, 2016), and so give us clearer insights into compensatory moral behaviors. Additionally, without any baseline in the distant past conditions, we cannot say if prosocial intentions distinctly increased after recalling moral behavior (positive consistency) or decreased after recalling immoral behavior (negative consistency). Therefore, extending C&P by adding two baseline conditions, that is also asking participants to recall neutral behavior from both the recent and distant past, can substantially improve our understanding of sequential moral behavior.

The Current Research

We conducted a large-scale (N = 5091) replication and extension of Conway and Peetz's (2012) original Study 1. We used the original materials and measures and closely followed C&P's procedures and materials for an independent direct replication. In line with Conway and Peetz's (2012) findings, we primarily hypothesized that recalling behaviors from the recent (distant) past should lead to compensatory (consistent) moral behavior.

All our hypotheses and analyses were pre-registered and can be accessed at: https://aspredicted.org/it26q.pdf. The study was programmed using Qualtrics (*Qualtrics*, 2022). All analyses were done using the statistical software jamovi (The jamovi project, 2022) and R (R Core Team, 2013). All codes, data, and study materials from the two planned experiments are publicly available on the Open Science Framework (OSF) repository here.

Method

Recruitment and Data Quality. We recruited participants from Prolific Academic, diverging from the original study which employed Amazon's Mechanical Turk (MTurk). This decision was based on Prolific having functions built specifically for academic research (Buhrmester et al., 2018), more stringent pre-screening for participants (Palan & Schitter, 2018), and greater naivety among participants (Peer et al., 2017) than MTurk. Moreover, Prolific has shown superior performance in terms of participant attention, comprehension, honesty, and reliability compared to MTurk (Peer et al., 2022). Given that both MTurk and Prolific are online crowdsourcing platforms with many shared characteristics (Goodman & Paolacci, 2017), choosing Prolific over Mturk should not bias results.

Participants. To minimize attrition, we stated in the recruitment post that the survey involved written communication, requesting participation only from those comfortable with typing one or two paragraphs (Zhou & Fishbach, 2016). Out of 5713 participants who began the study, 5165 participants completed the study (548 participants started but did not complete it). The attrition was slightly higher in the moral (205, 10.7%) and immoral (213, 11.2%) than in the neutral (130, 6.8%) conditions. To test for differences in attrition, we conducted a binary logistic regression with Attrition (1 = did not complete, 0 = completed) as the outcome variables, and Event Valence, Event Distance, and the interaction between Event Valence and Event Distance as predictors. The results indicated that compared to the neutral conditions, attrition rates were significantly higher in both the moral (estimate = .032, p = .035; OR = 1.033, 95% CI [1.048, 1.080]) and immoral (estimate = .042, p = .006; OR = 1.042, 95% CI [1.002, 1.064]) conditions. Neither Event Distance nor any of the interaction terms were significant (estimates < .024, ps > .117).

Further, to rigorously assess the impact of missing data on our findings, we conducted sensitivity analyses by condition (Valence: Moral, Immoral, Neutral) using Manski bounds. Manski bounds allow us to estimate the range of possible outcomes for our dependent variables—Willingness to Volunteer (WTV) and Willingness to Help (WTH)—under two contrasting scenarios. The 'lower bound' represents a conservative 'worst case' scenario, assuming that all missing data would have resulted in the lowest possible outcomes. Conversely, the 'upper bound' reflects an optimistic 'best case' scenario, assuming missing data would have resulted in the highest possible outcomes. The calculated bounds for both WTV (Moral [4.20, 4.77], Immoral [3.94, 4.55], Neutral [4.06, 4.43]) and WTH (Moral [5.22, 5.81], Immoral [5.12, 5.76], Neutral [5.20, 5.59]) across the three valence conditions were within narrow ranges. This indicates that irrespective of the assumptions made about the nature of the missing data, the observed differences in attrition rates have minimal impact on our study’s conclusions.

We now turn to exclusions from the 5165 complete responses. As per our pre-registered exclusion criteria, we excluded 59 responses that were flagged by Qualtrics’ fraud detection measure (ReCaptchaScore <= .5), and an additional 15 responses where participants failed two out of three attention checks. The final dataset for analyses comprised responses from 5091 participants (50.87% female; M_{age}= 39.3, SD = 14.0). For the direct replication analyses, a subset of N = 3339 participants was used (51.4% female; M_{age}= 39.2, SD = 14.1).

Due to a technical glitch in Qualtrics, there were unequal cell sizes with nearly three times as many participants in the “distant” conditions as in the “recent” conditions (as embedded in Qualtrics). Also, we mislabeled conditions such that neutral/distant and neutral/recent were reversed when programming Qualtrics, and so we relabeled those before data analyses. Although the stimuli in terms of descriptions shown to participants were unaffected, our coding error led to three times more participants in the neutral/recent than neutral/distant. Exact cell sizes in the final dataset are reported in Table 1.

Table 1. Achieved cell sizes per condition (exclusions in parentheses) in the final dataset.

Participants/ condition	Recent	Distant
Moral	447 (4)	1226 (23)
Immoral	455 (4)	1221 (19)
Neutral	1316 (15)	436 (9)

Unequal cell sizes may raise questions about whether homogeneity of variances assumptions hold. For the 2 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent X Distant) ANOVA for the replication analyses, the assumptions of equal variances were not violated for either of the dependent variables: Willingness-to-volunteer: $F(3, 3335) = 1.63$, $p = .180$, or Willingness-to-help: $F(3, 3335) = 2.50$, $p = .058$. For the 3 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent X Distant) ANOVA for the extension analyses, the assumptions of equal variances were violated for both of the dependent variables: willingness-to-volunteer: $F(5, 5085) = 2.40$, $p = .035$), and WTV: $F(3, 3335) = 2.30$, p

= .042). Nonetheless, Levene's test is quite sensitive in large samples because even small deviations from the homogeneity of variances assumption would be statistically significant, even though not meaningful in terms of the effect. Indeed, the ratios between the largest and smallest variation across cells for both dependent variables, WTV (ratio = 1.24) and WTH (ratio =1.44), were under 1.5, and therefore, unproblematic (Please see Table 4 to compare standard deviations). So, a statistically significant result in a large sample like ours may not imply a practically significant difference in variances. We also conducted robust linear regressions (non-preregistered) on both outcome variables. The results did not change in either the direction of effects or their significance levels. Please see Table S5C in section 6.4 in the Online Supplement for exact estimates.

Detectable effect sizes. Sensitivity analyses using G*Power (Faul et al., 2007) showed that the original C&P study, with 90% power, could detect an effect size of $f = .326$ or greater at $p < .05$ for the Event Valence X Event Distance interaction term in a 2 X 2 ANOVA, and an effect size of $d = .651$ or greater at $p < .05$ (two-tailed) for the planned contrast in a two-tailed t-test. At 90% power and $p < .05$ (two-tailed), our conducted replication can detect effect sizes more than five times smaller ($f = .056$ and $d = .112$ respectively) than the original C&P study can detect.

Considering the recent null results and possible publication bias inflating the moral licensing effect observed in meta-analyses ($d = .31$: Blanken et al., 2014; $d = .32$: Simbrunner & Schlegelmilch, 2017), the classical moral licensing effect may have been overestimated (Kuper & Bott, 2019). A "many-labs" study suggests a more modest effect size of $d = .14$ (Ebersole et al., 2016). Our sample size ($N = 3339$), which is more than 30 times that of C&P's original study ($N = 101$), allows us to detect effect sizes as small as $d = .112$ for C&P's specific a priori contrast (H1a), as well as for contrasts with neutral baselines (H2a and H2c). Please see Table 2 for a summary of minimal detectable effect sizes. Please see Figures S1.1 to S8 in the Online Supplement for power calculations.

Table 2. Summary of hypotheses and minimal detectable effect sizes (MDES) at 90% power for two-tailed tests at $p < .05$.

Effect	#	Hypothesis	MDES
Original Interaction		Event Valence X Event Distance Interaction (2 X 2 ANOVA)	$f = .056$ (Figure S1)
Contrast	H1	Participants in the moral/distant and immoral/recent conditions will exhibit higher prosocial intentions than participants in the moral/recent and immoral/distant conditions.	$d = .112$ (Figure S2)
	H2a	Participants in the moral/distant and immoral/recent conditions will exhibit higher prosocial intentions than participants in the neutral/recent and neutral/distant conditions.	$d = .111$ (Figure S3)
	H2b	Participants in the moral/recent and immoral/distant conditions will exhibit lower[2] prosocial intentions than participants in the neutral/recent and neutral/distant conditions.	$d = .111$ (Figure S4)

[2] In our original pre-registration on Aspredicted, we made a typo, whee we wrote “higher” instead of “lower”. H2a and H2b proposes to test if prosocial intentions increase and/or decrease from the baseline, with H2a and H2b mirroring each other such that H2a hypothesizes higher (increase) and H2b hypothesizes lower (decrease).

Licensing	H3a	Participants in the moral/recent condition will exhibit lower prosocial intentions than participants in the neutral/recent condition.	d = .200 (Figure S5)
Cleansing	H3b	Participants in the immoral/recent condition will exhibit higher prosocial intentions than participants in the neutral/recent condition.	d = .196 (Figure S6)
Positive consistency	H3c	Participants in the moral/distant condition will exhibit higher prosocial intentions than participants in the neutral/distant condition.	d = .201 (Figure S7
Negative consistency	H3d	Participants in the immoral/distant condition will exhibit lower prosocial intentions than participants in the neutral/distant condition.	d = .201 (Figure S8)

Procedure. In a 3 (Event Valence: Moral vs. Immoral vs. Neutral) X 2 (Event Distance: Recent vs. Distant) between-participants design, participants were randomly assigned to one of six conditions. They were instructed to recall and describe a moral, immoral, or neutral behavior from either the recent or distant past. Subsequently, participants responded to outcome measures of willingness-to-help (WTH) and willingness-to-volunteer (WTV), consistent with the original study.

Behavior Recall manipulation. We manipulated Event Valence by instructing participants to recall a moral, immoral, or neutral event. Participants in the moral condition were asked to recall a time when they acted in such a way that they felt righteous or honorable. Participants in the immoral condition were asked to recall a time when they acted in such a way that they felt guilty or ashamed. Participants in the neutral condition were asked to recall a time when they went shopping by themselves.

Event Distance was manipulated by asking participants to describe an event that occurred either in the past week (recent conditions) or over one year ago (distant conditions). For example, the prompt in the moral/recent condition was “Please recall a time within the past week when you acted in such a way that you felt righteous or honorable. Perhaps you were loyal to a friend, were generous when you could have been selfish, were kind to someone for no particular reason, or caring toward someone who needed you.” To elicit elaborate responses and strengthen the manipulations, participants were also told, “Please provide as much detail as you can, and write at least a paragraph with complete sentences.” on the same page as the response box in all conditions. Please see Table S2 in the Online Supplement for the recall prompts across conditions in exact words.

Prosocial Intention Measures. As dependent variables, we measured willingness to volunteer and help others right after participants completed the behavioral recall task. First, participants completed a 5-item willingness-to-volunteer measure (DeVoe & Pfeffer, 2007; $\alpha = .88$) on a 7-point scale anchored at 1 (completely disagree) to 7 (completely agree) comprising a randomized order of items such as “Volunteering is a worthwhile use of my time even if I do not get paid”.

Then, participants read four vignettes in a randomized order, each depicting other people needing small everyday help (e.g., paying a few extra cents for someone else’s restaurant bill). Participants indicated their willingness to help on a 7-point scale anchored at 1 (very unlikely) to 7 (very likely) for each scenario. These responses were aggregated into an index of willingness-to-help ($\alpha = .64$).

Manipulation checks. After the main outcome measures, participants rated event positivity, “The event I wrote about made me feel good about myself”, and perceived temporal distance, “The event I wrote about happened a long time ago”) on a 7-point scale anchored at 1 (completely agree) to 7 (completely disagree). At the very end of the study, participants were also asked to respond to an additional direct manipulation check of whether the event they recalled happened (a) within the last week, (b) more than a year ago, or (c) in between a week and a year.

PANAS. Participants completed the 20-item positive and negative affect schedule (PANAS; Watson et al., 1988) after the main measures. Participants indicated the extent to which they currently felt ten positive (e.g., interested, excited) and ten negative (e.g., depressed, upset) emotions on a 5-point scale anchored at 1 (not at all) to 5 (very strongly). Respective items were averaged into a positive emotion subscale (α = .93) and a negative emotion subscale (α = .92). Finally, participants responded to individual-level questions including indicating their age, gender, education, and income. Please see Table S1 in the Online Supplement for a list of all measures and items of the survey.

Differences between C&P and the current research

In Table 3, we list the known differences between the original and the replication study concerning design, materials, measures, and psychometric properties. Two notable differences may stand out. First, event positivity ratings in the moral conditions are much higher in our study. Second, perceived event distance ratings in the distant conditions are also higher in our replication. However, we argue that this perhaps suggests greater strengths in our manipulations and should not matter for replicating C&P's interaction effect.

Table 3. Comparisons of design, sample characteristics, and psychometric properties of key variables between Conway and Peetz (2012) and the current study.

	C&P	Current Study
Design	4 conditions	6 conditions

	2 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent vs. Distant) between-participants.	3 (Event Valence: Moral vs. Immoral vs. Neutral) X 2 (Event Distance: Recent vs. Distant) between-participants.
Materials: Recall prompt format (Illustration with Moral/Recent condition)	Please recall a time within the last week when you acted in such a way that you felt righteous or honorable. Perhaps you were loyal to a friend, were generous when you could have been selfish, were kind to someone for no particular reason, or caring toward someone who needed you.	Please recall a time within the last week when you acted in such a way that you felt righteous or honorable. Perhaps you were loyal to a friend, were generous when you could have been selfish, were kind to someone for no particular reason, or caring toward someone who needed you. Please provide as much detail as you can, and write at least a paragraph with complete sentences.
Outcome measures	willingness-to-volunteer, willingness-to-help	willingness-to-volunteer, willingness-to-help
Manipulation check measures	perceived event positivity, perceived event distance	perceived event positivity, perceived event distance, specific time of recalled event.
Sample Characteristics	Mturk participants N = 101 Female: 68% Age: M = 43.91, SD = 14.19	Prolific participants N = 5091 Female: 50.9% Age: M = 39.3, SD = 14.0

Cronbach's alpha (a) of measured constructs	Willingness-to-volunteer: α = .92 Willingness-to-help: α = .56 Negative emotion: α = .88 Positive emotion: α = .88	Willingness-to-volunteer: α = .89 Willingness-to-help: α = .64 Negative emotion: α = .92 Positive emotion: α = .93
Relevant means (M) and standard deviations (SD) of manipulation checks	Event positivity: Moral: M = 4.25, SD = .84 Immoral: M = 1.70, SD = .69. Event distance: Recent: M = 1.40, SD = .54 Distant: M = 3.02, SD = 1.09	Event positivity: Moral: M = 5.81, SD = 1.30 Immoral: M = 1.81, SD = 1.64 Neutral: M = 4.0, SD = 1.64. Event distance Recent: M = 1.54, SD = 1.17 Distant: M = 4.07, SD = 1.86

Results

First, we compare descriptive statistics between the original and current study for the two dependent variables listed in Table 4 while also detailing results from manipulation checks. Then, we replicate C&P's exact analyses on a subset of the data (N = 3339) comprising the four original conditions of C&P. Following that, we report analyses based on the full dataset (N = 5091) with all six conditions. Although our current sample reported lower willingness-to-volunteer than C&P *on average*, this may merely be due to differences in possible sampling error between C&P (N = 101) and the current research (N = 5091).

Table 4. Comparing means (M) and standard deviations (SD) of dependent variables (D) between C&P and current study across conditions.

DV	Condition	C&P	Current Study

Willingness-to-volunteer	Moral/Recent	M = 5.42, SD = 1.37	M = 4.51, SD = 1.39
	Immoral/Recent	M = 5.97, SD = 1.30	M = 4.29, SD = 1.38
	Moral/Distant	M = 5.94, SD = 1.02	M = 4.63, SD = 1.32
	Immoral/Distant	M = 5.26, SD = 1.49	M = 4.37, SD = 1.39
	Neutral/Recent	NA	M = 4.32, SD = 1.42
	Neutral/Distant	NA	M = 4.26, SD = 1.47
Willingness-to-help	Moral/Recent	M = 5.81, SD = .70	M = 5.69, SD = 0.89
	Immoral/Recent	M = 6.00, SD = .73	M = 5.56, SD = 0.99
	Moral/Distant	M = 6.01, SD = .78	M = 5.68, SD = 1.01
	Immoral/Distant	M = 5.56, SD = .78	M = 5.63, SD = 1.02
	Neutral/Recent	NA	M = 5.48. SD = 1.07
	Neutral/Distant	NA	M = 5.53, SD = 1.07

Manipulation checks

With respect to the full dataset consisting of all six conditions, a 3 (Event Valence: Moral vs. Immoral vs. Neutral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on recalled event positivity revealed the expected effect of Event Valence on event positivity, F (2, 5085) = 2568.82, p < .001, ${\eta^2}_p = .502$. Event Distance did not significantly moderate the effect of valence, F (2, 5085) = .102, p = .903, ${\eta^2}_p < .001$. Participants instructed to recall and write about a moral event (M = 5.81, SD = 1.30) reported their event to be significantly more positive than those in the immoral conditions (M = 1.81, SD = 1.31), t (5086) = 71.6, p_{Tukey} < 0.001, d = 2.79, 95% CI [2.70, 2.88]. Participants instructed to recall and write about a neutral event (M = 4.00, SD = 1.64) reported their event to be significantly *less* positive than participants in the moral conditions (M = 5.81, SD = 1.30), t (5085) = -33.3, p_{Tukey} < .001, d = -1.30, 95% CI [-1.38, 1.-22], and significantly *more* positive than participants in the immoral conditions (M = 1.81, SD = 1.31), t (5085) = 38.2, p_{Tukey} < .001, d = 1.49, 95% CI [1.41, 1.57].

A 2 (Event Distance: Recent vs. Distant) X 2 (Event Valence: Moral vs. Immoral) between-participants ANOVA on perceived distance of the recalled event revealed the expected effect of Event Distance on perceived distance, F (1, 5085) = 2472.46, p < .001, ${\eta^2}_p = .327$. Event Valence did not moderate the effect of Event Distance, F (2, 5085) = .249, p = .780, ${\eta^2}_p < .001$. Participants instructed to recall and write about a distant event (M = 3.17, SD = 1.84) perceived the event to be significantly more distant than participants in the recent conditions (M = 2.39, SD = 2.05), t (5085) = 49.72, p_{Tukey} < .001, d = 1.58, 95% CI [1.51, 1.65]. Please see Tables S4A and S4B in the Online Supplement for comparisons of test statistics of manipulation checks between the original and replication.

Our direct manipulation check measure for Event Distance showed that a significant majority of participants in the 'Recent' condition perceived the events as occurring 'within the last week' (90.7%) than otherwise (9.03%), Chi-squared statistic = 882.91, p < .001. Moreover, a significant proportion of participants in the 'Distant' condition perceived the events as occurring 'more than a year ago' (63.8%) than otherwise (36.2%), Chi-squared statistic = 203.37, p < .001. Including or excluding participants who answered incorrectly with respect to this check did not materially change results. Please see Tables S5A and S5B in the Online Supplement for an overview and comparisons of test statistics between datasets.

Replicating C&P's analyses

Here, we report the same analyses as in C&P to test if support for their original hypothesis (H1) can be replicated. For comparisons of the test statistics between C&P and our study, please see Tables S4C and S4D in the Online Supplement.

Willingness-to-volunteer (WTV). A 2 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on willingness to volunteer (WTV) revealed a significant effect of Event Valence, F (1, 3335) = 20.332, p <.001, ${\eta^2}_p = .006$; and a non-significant effect of Event Distance, F (1, 3335) = 3.817, p = .051, ${\eta^2}_p = .001$. Crucially, the Event Valence X Event Distance interaction effect on WTV was **not** significant, F (1, 3335) = .176, p = .675, ${\eta^2}_p < .001$, rendering the planned comparison of H1 irrelevant.

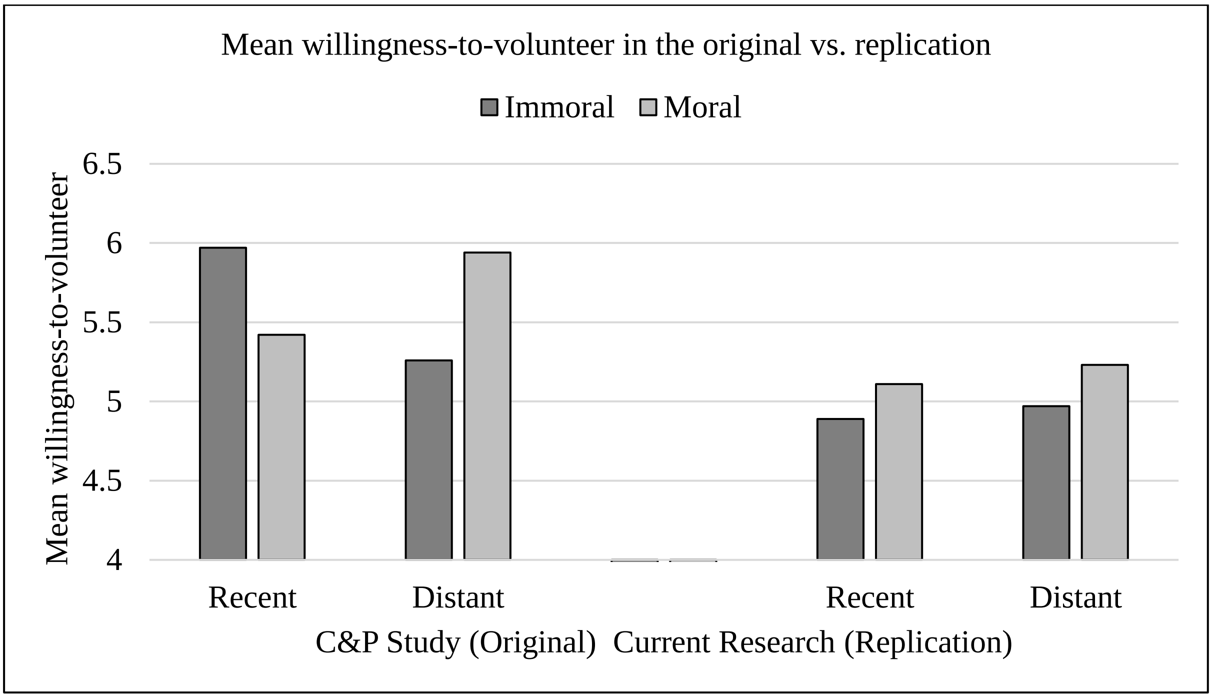


Figure 1A. Mean willingness-to-volunteer across conditions in the original versus replication. Responses were measured on 7-point scales anchored at 1 and 7.

Tukey's post hoc comparisons showed that WTV was higher in the moral (M = 4.60, SD = 1.34) than immoral (M = 4.35, SD = 1.39) conditions, t (3335) = 4.51, $p_{\mathrm{Tukey}} < .001$, d = .176, 95% CI [.099, .252]. A closer look also revealed that WTV was non-significantly higher in the moral/recent (M = 5.11, SD = 1.39) than immoral/recent (M = 4.89, SD = 1.38) condition, t (3335) = 2.39, $p_{\mathrm{Tukey}} = .079$, d = .159, 95% CI [.029, .290]). Similarly, WTV was higher in the moral/distant (M = 5.23, SD = 1.32) than immoral/distant (M = 4.97, SD = 1.39) condition, t (3335) = 4.74, p_{Tukey} < .001, d = .192, 95% CI [.113, .272]).

So, contrary to C&P, our results suggest an overall presence of moral consistency in both recent and distant conditions, and the absence of compensatory moral behavior in the recent conditions, together suggesting that the original interaction hypothesis (H1) is not supported.

Willingness-to-help. A 2 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on willingness to help (WTH) revealed a significant effect of Event Valence, F (1, 3335) = 5.643, p = .017, $\eta^2{}_p = .002$; and a non-significant effect of Event Distance, F (1, 3335) = 0.519, p = .47, $\eta^2{}_p < .001$. Crucially, the Event Valence X Event Distance interaction effect on WTH was **not** significant, F (1, 3335) = 1.052, p = .305, $\eta^2{}_p < .001$, rendering the planned comparisons with respect to H1 superfluous.

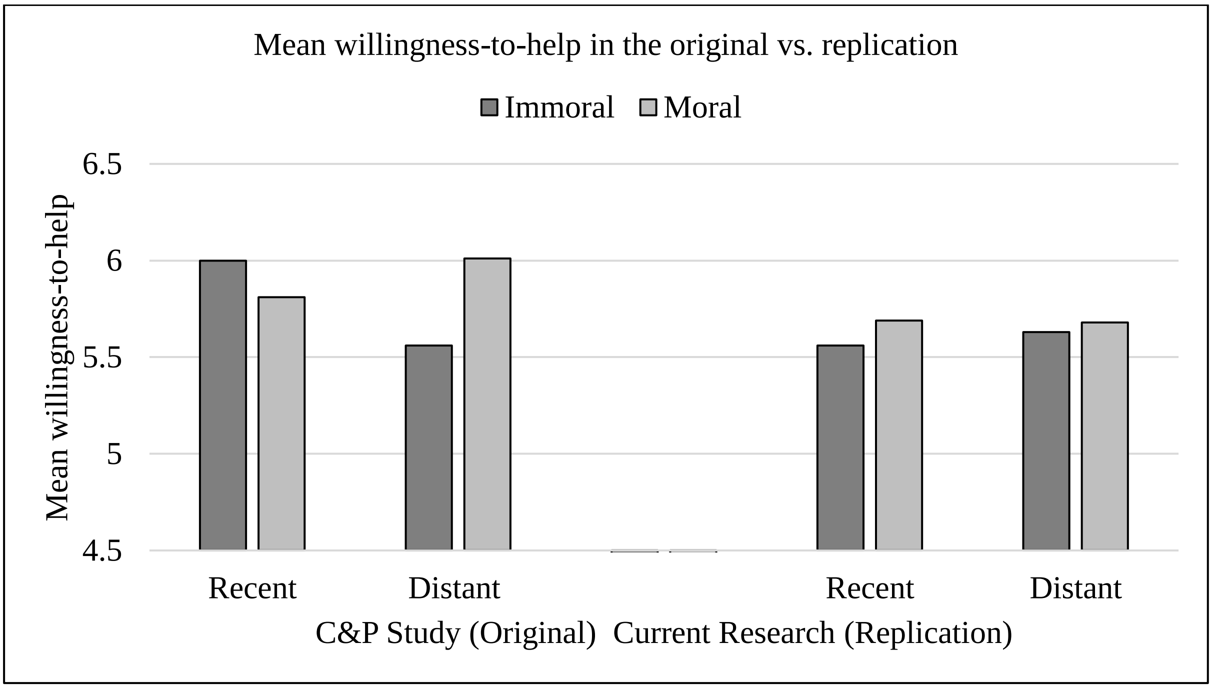


Figure 1B. Mean willingness-to-help across conditions in the original versus replication. Responses were measured on 7-point scales anchored at 1 and 7.

Tukey's post hoc comparisons showed that WTH was higher in the moral (M = 5.68, SD = 0.98) than immoral (M = 5.61, SD = 1.01) conditions, t (3335) = 2.38, p_{Tukey} = .017, d = .093, 95% CI [.016, .169]. A closer look also revealed that WTH was non-significantly higher in the moral/recent (M = 5.69, SD = .089) than immoral/recent (M = 5.56, SD = 0.99) condition, t (3335) = 1.99, p_{Tukey} = .190, d = .133, 95% CI [.002, .263]). However, WTH was non-significantly different in the moral/distant (5.68, SD = 1.01) and immoral/distant (M = 5.63, SD = 1.02) conditions, t (3335) = 1.31, p_{Tukey} = .560, d = .068, 95% CI [-.040, .176]).

So, contrary to C&P, our results suggest a minimal effect of moral consistency (d = .093) when the recent and distant conditions are combined, and the absence of compensatory moral behavior in the recent conditions, again suggesting that the original interaction hypothesis (H1) is not supported.

Analyses with the full dataset

We now turn to our analyses with the full dataset to test hypotheses H2a-b and H3a-d. For all non-significant planned comparisons, we conducted equivalence tests (Lakens, 2017) to examine if the null was conclusive or inconclusive. These were not pre-registered and should be seen as exploratory when interpreting null results. As opposed to how traditional hypothesis testing looks for significant differences, an equivalence test examines if the difference between two conditions is smaller than a pre-specified smallest effect size of interest (SESOI). Here, the null hypothesis posits that the true difference between the conditions is greater in magnitude (that is either above the upper bound or below the lower bound). The alternative hypothesis posits that the true differences lie within the bound, hence suggesting equivalence between the groups, and so statistical equivalence. For each TOST analysis, we reported the p-value against the bound (lower or upper) depending on the direction of the hypothesis. For a primer and tutorial, please see Lakens (2017) and Lakens et al. (2018).

We determined our smallest effect sizes of interest (SESOI) to d = .14, informed by Ebersole et al.'s (2016) multi-lab study finding the average effect of moral licensing effect to be d = -.14. Accordingly, we set our upper bounds as d = .14 when equivalence testing for moral cleansing and positive moral consistency, and lower bounds as d = -.14 when equivalence testing for moral licensing and negative moral consistency.

Willingness-to-volunteer. A 3 (Event Valence: Moral vs. Immoral vs. Neutral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on willingness-to-volunteer (WTV) revealed a significant effect of Event Valence, F (2, 5085) = 15.54, p < .001, ${\eta^2}_p = .006$; but a non-significant effect of Event Distance, F (1, 5085) = 1.16, p = .281, ${\eta^2}_p < .001$. The Event Valence X Event Distance interaction effect on WTV was **not** significant, F (2, 5085) = 1.69, p = .184, ${\eta^2}_p < .001$, rendering planned comparisons for H2a-b superfluous.

Tukey's post hoc pairwise comparisons revealed that although WTV was lower in the neutral (M = 4.31, SD = 1.43) than in the moral (M = 4.60, SD = 1.34) conditions, t (5085) = -4.36, p_{Tukey} < .001, d = -.170, 95% CI [-.247, -.094], WTV in the neutral conditions was not significantly different than in the immoral (M = 4.35, SD = 1.39) conditions, t (5085) = -.73, p_{Tukey} = .746, d = -.028, 95% CI [-.105, .048].

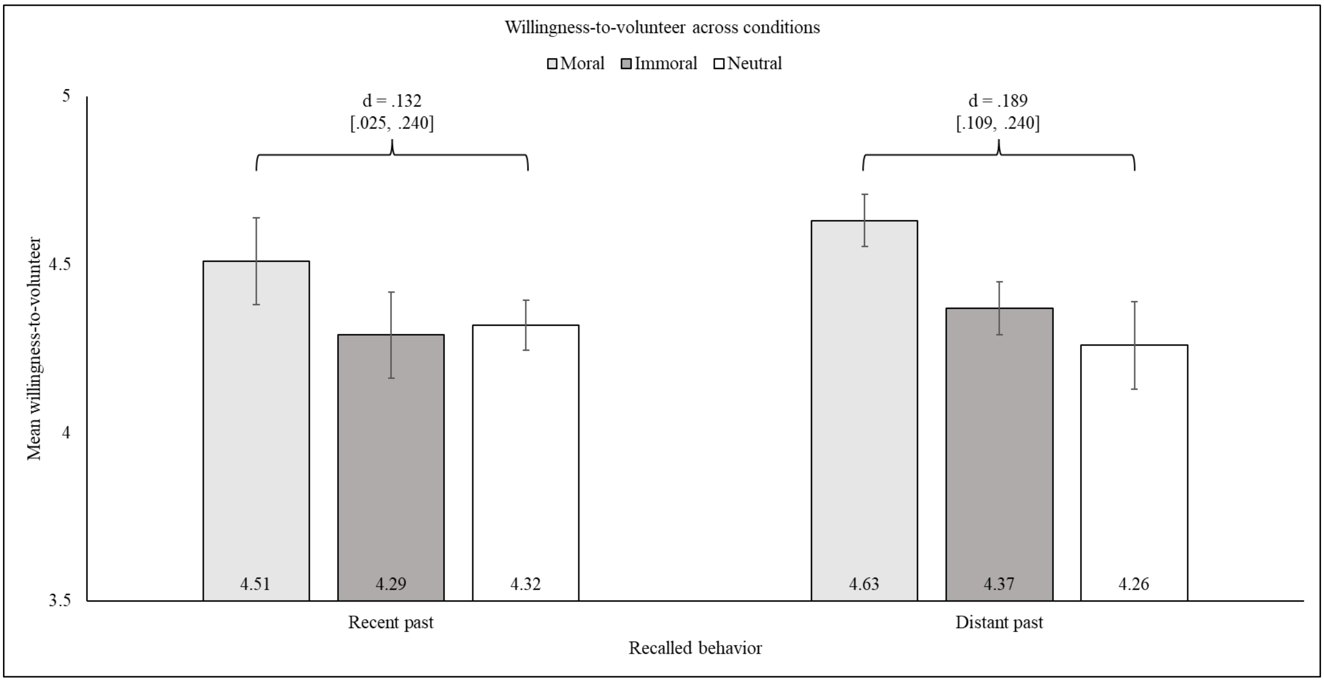


Figure 2A. Mean willingness-to-volunteer (WTV) across conditions. Error bars refer to 95% confidence intervals.

To test H3a (moral licensing) and H3b (moral cleansing), we turn to planned comparisons to test for the presence of *compensatory* moral behavior in the "recent" event conditions. WTV was non-significantly higher in the moral/recent (M = 4.51, SD = 1.39) than neutral/recent (M = 4.32, SD = 1.42) condition, t (5085) = -2.41, p_{Tukey} = .150, d = .132, 95% CI [.025, .240]. The opposite direction of the significant effect suggests strong evidence *against* a moral licensing effect (H3a) in the recent conditions.

However, WTV was not significantly different in the immoral/recent (M = 4.29, SD = 1.38) than in the neutral/recent (M = 4.32, SD = 1.42) condition, t (5085) = -.44, p_{Tukey} = .998, d = .024, 95% CI [-.082, .131]. An equivalence test with an upper bound of d = .14 was significant, t (809) = 3.04, p = .001. So, the standardized mean difference between the immoral and neutral conditions lies between 0 and .14, a magnitude smaller than our minimum effect size of interest, and so suggesting conclusive evidence of a null effect of moral cleansing (H3b).

To test H3c (positive moral consistency) and H3d (negative moral consistency), we turn to planned comparisons to test for *consistent* moral behavior in the "distant" event conditions. WTV was significantly higher in the moral/distant (4.63, SD = 1.32) than neutral/distant (M = 4.26, SD = 1.47) condition, t (5085) = 4.84, p_{Tukey} < .001, d = .189, 95% CI [.109, .268], suggesting a positive moral consistency effect in the distant conditions.

However, WTV was not significantly different in the immoral/distant (M = 4.37, SD = 1.39) than in neutral/distant (M = 4.26, SD = 1.47) condition, t (5085) = 1.45, p_{Tukey} = .695, d = .081, 95% CI [.028, -.191]. An equivalence test with a lower bound of d = -.14 was significant, t (733) = -3.86, p < .001. So, the standardized mean difference between the immoral and neutral conditions lies between 0 and .14, suggesting conclusive evidence of a null effect of negative moral consistency (H3d).

Willingness-to-help. A 3 (Event Valence: Moral vs. Immoral vs. Neutral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on willingness-to-help (WTH) revealed a significant effect of Event Valence, F (2, 5085) = 9.53, p < .001, ${\eta^2}_p$ = .004, and non-significant effect of Event Distance, F (1, 5085) = 1.20, p = .273, ${\eta^2}_p < .001$. The Event Valence X Event Distance interaction effect on WTH was **not** significant, F (2, 5085) = .555, p = .574, ${\eta^2}_p < .001$, making the planned comparisons of H2a-b redundant.

Tukey's post hoc pairwise comparison tests revealed that WTH was significantly lower in the neutral (M = 5.50, SD = 1.07) than moral (M = 5.68, SD = .98) conditions, t (5085) = -4.36, p < .001, d = -.17, 95% CI [-.247, -.094], and not significantly different than immoral (M = 5.61, SD = 1.01) condition, t (5085) = -2.05, p_{Tukey} = .10, d = -.080, 95% CI [-.003, -.156]. This again hints at the presence of an overall positive moral consistency effect.

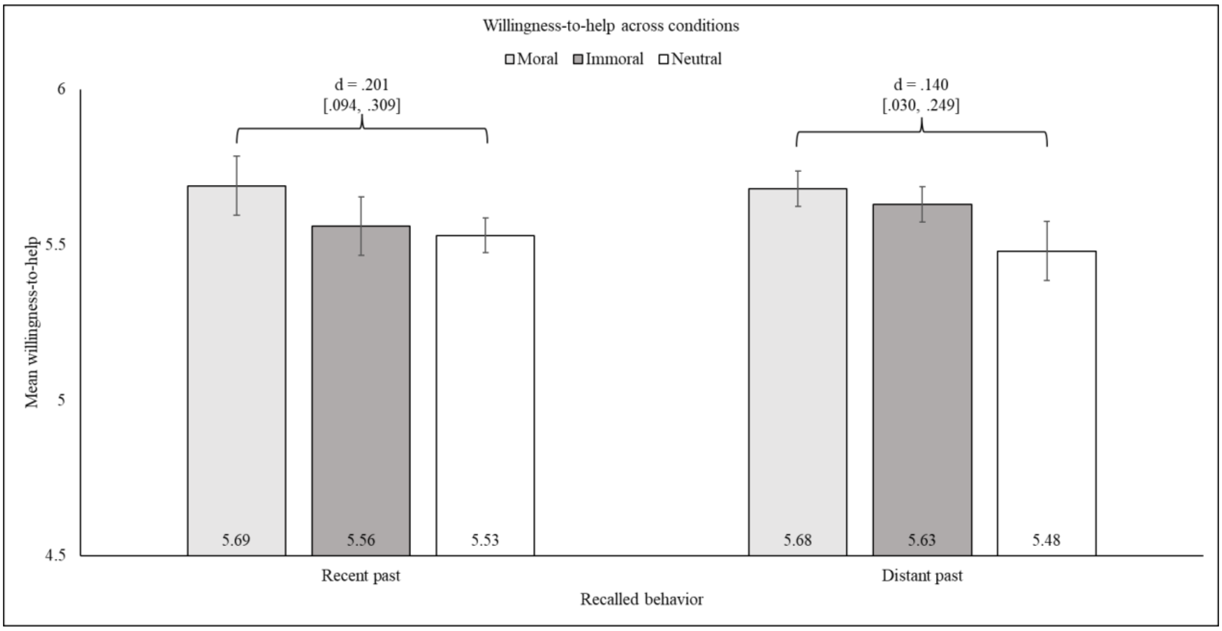


Figure 2B. Mean willingness-to-help (WTH) across conditions. Error bars refer to 95% confidence intervals.

To test H3a (moral licensing) and H3b (moral cleansing), we turn to planned comparisons to test for the presence of *compensatory* moral behavior in the "recent" distant conditions. WTH was significantly higher in the moral/recent (M = 5.69, SD = .89) than neutral/recent (M = 5.53, SD = 1.07) condition, t (5085) = -3.67, p_{Tukey} = .003, d = .201, 95% CI [.094, .309]. The opposite direction of the significant effect provides strong evidence *against* moral licensing (H3a).

However, WTH was not significantly different in the immoral/recent (M = 5.56, SD = .99) than neutral/recent (M = 5.53, SD = 1.07) condition, t (5085) = -.44, p_{Tukey} = .998, d = .024, 95% CI [-.082, .131]. An equivalence test with an upper bound of d = .14 was not significant, t (841) = 1.29, p = .098, thereby suggesting the standardized mean difference between the immoral and neutral condition may lie outside 0 and -.14, potentially having a magnitude bigger than our minimum effect size of interest. This presents inconclusive evidence for a null effect of moral cleansing (H3b) when comparing the "recent" conditions.

To test H3c (positive moral consistency) and H3d (negative moral consistency), we turn to planned comparisons to test for *consistent* moral behavior in the "distant" event conditions. WTH was not significantly higher in the moral/distant (5.68, SD = .98) than in the neutral/distant (M = 5.48, SD = 1.07) condition, $t(5085) = 2.51$, $p_{\mathrm{Tukey}} = .122$, $d = .140$, 95% CI [.030, .249]. Note that the p-value *without* correcting for multiple post hoc comparisons, $p = .012$, is below the conventional significance level. Furthermore, an equivalence test with an upper bound of $d = .14$ was not significant, $t(731) = .052$, $p = .479$, thereby suggesting the standardized mean difference between the immoral and neutral condition may lie outside 0 and .14, a magnitude greater than our minimum effect size of interest. This presents inconclusive evidence for a null effect of positive moral consistency (H3c) the distant conditions.

However, WTH was not significantly different in the immoral/distant (M = 5.63, SD = 1.02) than in neutral/distant (M = 5.48, SD = 1.07) condition, $t(5085) = 1.58$, $p_{\mathrm{Tukey}} = .612$, $d = .088$, 95% CI [.021, -.198]. An equivalence test with a lower bound of $d = -.14$ was significant, $t(733) = -4.01$, $p < .001$. So, the standardized mean difference between the immoral and neutral conditions may lie between 0 and -.14, a magnitude smaller than our minimum effect size of interest, suggesting conclusive evidence of a null effect of negative moral consistency (H3d).

Discussion

Our study, as the first independent replication of C&P and the largest single-lab investigation in this domain, tested the hypothesis proposed by Conway and Peetz (2012) that recalling moral or immoral deeds from the recent (distant) past should lead to compensatory (consistent) moral behavior. The results did *not* support conceptual abstraction being a moderator. Instead, the results suggest a positive moral consistency effect. Specifically, recalling past moral behavior (versus immoral or neutral behavior), regardless of whether from the recent or distant past, increased prosocial intentions. These results challenge the prevailing theories on compensatory moral behavior and highlight the need for reevaluation in this domain.

Given that we could not replicate C&P's proposed moderation by conceptual abstraction, it may also be important to reconsider other hypothesized moderators in sequential moral behavior (e.g., Brown et al., 2011; Griep et al., 2021; Lalot et al., 2022; Susewind & Hoelzl, 2014; Susewind & Walkowitz, 2020). Understanding the factors that regulate whether doing good or bad leads to similar or opposite behavior afterward is crucial not only in moral psychology but also in domains such as organizational (List & Momeni, 2021) and consumer (Juhl et al., 2017) behavior.

Building on these broader implications, our findings of a robust positive moral consistency effect resonate with some existing literature in the field. For instance, a study tracking purchases of 8704 randomly selected Danish retail consumers over 20 months found that those who bought organic products at one point in time were more likely to buy *more* organic products (Juhl et al., 2017), supporting the positive moral consistency perspective. Moreover, Rotella and Barclay (2020) found that participants donated significantly more after recalling past moral behavior than neutral behavior, and *marginally* after recalling past moral than immoral behavior. While we find effects in the same direction as Rotella and Barclay (2020), ours are more robust given the much larger sample size (N = 519 vs. N = 5091). Together with other null effects of compensatory moral behavior (e.g., Blanken et al., 2014; Eskine, 2013), our study supports the relevance of classical behavioral consistency theories (Festinger, 1954; Heider, 1946).

Our findings should be interpreted with some caveats. First, empirical findings from a single-lab online experiment conducted on a single population (participants based in the USA) limit cross-cultural generalizability. Interestingly, a culture-moderated meta-analysis examining the moral licensing effect (Simbrunner & Schlegelmilch, 2017) found moral licensing to be stronger among North American than Western European and Asian participants. Finding conclusive null effects with U.S. American participants, as we do, may suggest a need to rethink how we understand compensatory moral behavior because this was a cultural sample where the effects are supposedly stronger (Simbrunner & Schlegelmilch, 2017).

Second, we do not claim to rule out compensatory moral behavior altogether. A meta-analysis on interpersonal and intrapsychic mechanisms finds strong evidence in favor of compensatory behavior only when individuals are *observed* (Rotella et al., 2023). Because our study was conducted online and anonymously, recalling past behavior may not have granted a license or need to behave oppositely after (Rotella & Barclay, 2020). This may explain why we did not observe a robust licensing or cleansing effect.

Third, because our replication was conducted at least 11 years after Conway and Peetz (2012), experimental participants may have become less sensitive to recall-and-write tasks. Nonetheless, our manipulations were successful in that there were clear differences with respect to both manipulation check measures and automated text analyses (Berger et al., 2020) also indicated clear differences by condition.

Future research could explore whether this positive moral consistency is observed in settings with aligned incentives, such as donation or dictator games, or in the context of dishonesty measures, to see if recalling moral/immoral deeds influences cheating behavior. Although we find robust evidence of positive moral consistency across one prior and one sequential behavior, it can be interesting to test how consistent behaviors are across *multiple* subsequent behaviors.

Conclusion

Our large-scale replication and extension of Conway and Peetz’s (2012) study challenge longstanding assumptions in the field of moral psychology, particularly regarding the role of conceptual abstraction in moderating moral behavior. Our findings, which did not support the hypothesized interaction between moral behavior recollection and its temporal distance, suggest a more prevalent influence of positive moral consistency across different temporal contexts. This highlights a potential need to reevaluate the theoretical frameworks underpinning compensatory moral behavior. By revealing the robustness of moral consistency, irrespective of whether past behavior is recalled from the recent or distant past, our study supports the classical consistency perspective in moral decision-making and underscores the importance of registered replications in psychology.

References

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*, 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Berger, J., Sherman, G., & Ungar, L. (2020). *TextAnalyzer* [Computer software]. http://textanalyzer.org

Blanken, I., van de Ven, N., & Zeelenberg, M. (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin*, *41*, 540–558. https://doi.org/10.1177/0146167215572134

Blanken, I., van de Ven, N., Zeelenberg, M., & Meijers, M. H. C. (2014). Three attempts to replicate the moral licensing effect. *Social Psychology*, *45*, 232–238. https://doi.org/10.1027/1864-9335/a000189

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*, 1–45. https://doi.org/10.1037/0033-2909.88.1.1

Brown, R. P., Tamborski, M., Wang, X., Barnes, C. D., Mumford, M. D., Connelly, S., & Devenport, L. D. (2011). Moral credentialing and the rationalization of misconduct. *Ethics & Behavior*, *21*, 1–12. https://doi.org/10.1080/10508422.2011.537566

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science*, *13*, 149–154. https://doi.org/10.1177/1745691617706516

Chambers, C. D., & Tzavella, L. (2021). The past, present and future of Registered Reports. *Nature Human Behaviour*, *6*, 29–42. https://doi.org/10.1038/s41562-021-01193-7

Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, *38*, 907–919. https://doi.org/10.1177/0146167212442394

Cornelissen, G., Bashshur, M. R., Rode, J., & Le Menestrel, M. (2013). Rules or consequences? The role of ethical mind-sets in moral dynamics. *Psychological Science*, *24*, 482–488. https://doi.org/10.1177/0956797612457376

DeVoe, S. E., & Pfeffer, J. (2007). When time is money: The effect of hourly payment on the evaluation of time. *Organizational Behavior and Human Decision Processes*, *104*, 1–13. https://doi.org/10.1016/j.obhdp.2006.05.003

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., … Nosek, B. A. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, *67*, 68–82. https://doi.org/10.1016/j.jesp.2015.10.012

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites. *Journal of Experimental Social Psychology*, *45*, 590–593. https://doi.org/10.1016/j.jesp.2009.02.001

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology*, *6*, 32–35. https://doi.org/10.1016/j.copsyc.2015.03.017

Eskine, K. J. (2013). Wholesome foods and wholesome morals?: Organic foods reduce prosocial behavior and harshen moral judgments. *Social Psychological and Personality Science*, *4*, 251–254. https://doi.org/10.1177/1948550612447114

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. https://doi.org/10.3758/BF03193146

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140. https://doi.org/10.1177/001872675400700202

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*, 195–202. https://doi.org/10.1037/h0023552

Gholamzadehmir, M., Sparks, P., & Farsides, T. (2019). Moral licensing, moral cleansing and pro-environmental behaviour: The moderating role of pro-environmental attitudes. *Journal of Environmental Psychology*, *65*, 101334. https://doi.org/10.1016/j.jenvp.2019.101334

Goodman, J. K., & Paolacci, G. (2017). Crowdsourcing consumer research. *Journal of Consumer Research, 44*, 196–210. https://doi.org/10.1093/jcr/ucx047

Griep, Y., Germeys, L., & Kraak, J. M. (2021). Unpacking the relationship between organizational citizenship behavior and counterproductive work behavior: Moral licensing and temporal focus. *Group & Organization Management, 46*, 819–856. https://doi.org/10.1177/1059601121995366

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology*, *21*, 107–112. https://doi.org/10.1080/00223980.1946.9917275

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, *37*, 701–713. https://doi.org/10.1177/0146167211400208

Juhl, H. J., Fenger, M. H. J., & Thøgersen, J. (2017). Will the consistent organic food consumer step forward? An empirical analysis. *Journal of Consumer Research*, *44*, 519–535. https://doi.org/10.1093/jcr/ucx052

Karmarkar, U. R., & Bollinger, B. (2015). BYOB: How bringing your own shopping bags leads to treating yourself and the environment. *Journal of Marketing*, *79*, 1–15. https://doi.org/10.1509/jm.13.0228

Kuper, N., & Bott, A. (2019). Has the evidence for moral licensing been inflated by publication bias? *Meta-Psychology*, *3*. https://doi.org/10.15626/MP.2018.878

Lalot, F., Falomir-Pichastor, J. M., & Quiamzade, A. (2022). Regulatory focus and self-licensing dynamics: A motivational account of behavioural consistency and balancing. *Journal of Environmental Psychology*, *79*, 101731. https://doi.org/10.1016/j.jenvp.2021.101731

List, J. A., & Momeni, F. (2021). When corporate social responsibility backfires: Evidence from a natural field experiment. *Management Science*, *67*, 8–21. https://doi.org/10.1287/mnsc.2019.3540

Maier, M., Bartoš, F., Oh, M., Wagenmakers, E.-J., Shanks, D., & Harris, A. J. L. (2022). *Adjusting for publication bias reveals that evidence for and size of construal level theory effects is substantially overestimated* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/r8nyu

Mazar, N., & Zhong, C.-B. (2010). Do green products make us better people? *Psychological Science*, *21*, 494–498. https://doi.org/10.1177/0956797610363538

Moery, E., & Calin-Jageman, R. J. (2016). Direct and conceptual replications of Eskine (2013): Organic food exposure has little to no effect on moral judgments and prosocial behavior. *Social Psychological and Personality Science*, *7*, 312–319. https://doi.org/10.1177/1948550616639649

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*, 33–43. https://doi.org/10.1037/0022-3514.81.1.33

Mullen, E., & Monin, B. (2016). Consistency versus licensing effects of past moral behavior. *Annual Review of Psychology*, *67*, 363–385. https://doi.org/10.1146/annurev-psych-010213-115120

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716. https://doi.org/10.1126/science.aac4716

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*, 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Qualtrics (Version 2022). (2022). [Computer software]. Qualtrics. https://www.qualtrics.com

R Core Team. (2013). *R* [Computer software]. R Foundation for Statistical Computing. www.R-project.org/

Reed, A., Aquino, K., & Levy, E. (2007). Moral identity and judgments of charitable behaviors. *Journal of Marketing*, *71*, 178–193. https://doi.org/10.1509/jmkg.71.1.178

Reed, A., Kay, A., Finnel, S., Aquino, K., & Levy, E. (2016). I don’t want the money, I just want your time: How moral identity overcomes the aversion to giving time to prosocial causes. *Journal of Personality and Social Psychology*, *110*, 435–457. https://doi.org/10.1037/pspp0000058

Rosen, C. C., Koopman, J., Gabriel, A. S., & Johnson, R. E. (2016). Who strikes back? A daily investigation of when and why incivility begets incivility. *Journal of Applied Psychology*, *101*, 1620–1634. https://doi.org/10.1037/apl0000140

Rotella, A., & Barclay, P. (2020). Failure to replicate moral licensing and moral cleansing in an online experiment. *Personality and Individual Differences*, *161*, 109967. https://doi.org/10.1016/j.paid.2020.109967

Rotella, A., Jung, J., Chinn, C., & Barclay, P. (2023). *Observation moderates the moral licensing effect: A meta-analytic test of interpersonal and intrapsychic mechanisms* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/tmhe9

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, *20*, 523–528. https://doi.org/10.1111/j.1467-9280.2009.02326.x

Simbrunner, P., & Schlegelmilch, B. B. (2017). Moral licensing: A culture-moderated meta-analysis. *Management Review Quarterly*, *67*, 201–225. https://doi.org/10.1007/s11301-017-0128-0

Susewind, M., & Hoelzl, E. (2014). A matter of perspective: Why past moral behavior can sometimes encourage and other times discourage future moral striving. *Journal of Applied Social Psychology*, *44*, 201–209. https://doi.org/10.1111/jasp.12214

Susewind, M., & Walkowitz, G. (2020). Symbolic moral self-completion – social recognition of prosocial behavior reduces subsequent moral striving. *Frontiers in Psychology*, *11*, 560188. https://doi.org/10.3389/fpsyg.2020.560188

The jamovi project. (2022). *Jamovi.* (2.3) [Computer software]. https://www.jamovi.org

Tiefenbeck, V., Staake, T., Roth, K., & Sachs, O. (2013). For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy, 57*, 160–171. https://doi.org/10.1016/j.enpol.2013.01.021

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*, 403–421. https://doi.org/10.1037/0033-295X.110.3.403

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*, 440–463. https://doi.org/10.1037/a0018963

Urban, J., Bahník, Š., & Kohlová, M. B. (2019). Green consumption does not make people cheat: Three attempts to replicate moral licensing effect due to pro-environmental behavior. *Journal of Environmental Psychology*, *63*, 139–147. https://doi.org/10.1016/j.jenvp.2019.01.011

van der Werff, E., Steg, L., & Keizer, K. (2014). Follow the signal: When past pro-environmental actions signal who you are. *Journal of Environmental Psychology*, *40*, 273–282. https://doi.org/10.1016/j.jenvp.2014.07.004

Wang, Y., Rodríguez De Gil, P., Chen, Y.-H., Kromrey, J. D., Kim, E. S., Pham, T., Nguyen, D., & Romano, J. L. (2017). Comparing the performance of approaches for testing the homogeneity of variance assumption in one-factor ANOVA Models. *Educational and Psychological Measurement*, *77*, 305–329. https://doi.org/10.1177/0013164416645162

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of*

Personality and Social Psychology, *111*, 493–504.

https://doi.org/10.1037/pspa0000056

Chapter 3. Article 2: Business-Size Bias in Moral Concern: People are More Dishonest Against Big than Small Organizations

PREPRINT

Working paper part of a doctoral dissertation (March 2024). All conclusions are preliminary. Please do not copy, share, or cite without the authors' permission.

Jareef Martuza[1], Hallgeir Sjåstad[1], Helge Thorbjørnsen[1,2]

[1]Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, Bergen 5045, Norway

[2] SNF, Center for Applied Research at NHH, Helleveien 30, Bergen 5045, Norway

Author notes

The authors declare to have no known conflicts of interest. Correspondence about this research should be addressed to Jareef Martuza (jareef.martuza@nhh.no) PhD student at the Norwegian School of Economics.

Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. As Norwegian laws and regulations do not require review by an institutional review board for anonymous, non-medical, low-risk research with human participants, this project was not submitted to such review. Informed consent was collected from all participants. We received partial funding for our studies from the Center of Applied Research (SNF) at NHH to conduct our experiments.

Abstract

Despite the potential for market domination and power abuse, it is also true that big businesses can enhance consumer choice, job benefits, and welfare. Yet, big businesses are sometimes portrayed as inherently bad. Why? And does it matter for how we treat these organizations as decision-makers? Informed by moral typecasting theory, we suggest people are less likely to perceive big businesses as a *vulnerable victim* than small businesses, making it seem more acceptable to cheat them for selfish gain. We studied this “business-size bias” across eight experiments (N = 5,670). Experiments 1a-c found that people intended to be more dishonest against big than small businesses, mediated by lower vulnerability perceptions. Experiments 2a-c found that people also *behaved* more dishonestly toward big businesses. Experiments 3a-b replicated the business-size bias in dishonesty, and found that it was mediated by perceptions of big businesses as less vulnerable and less moral than small businesses.

Public Significance Statement

We present robust evidence from a series of scenario-based and incentivized experiments (N =5,670) that people are more dishonest toward big than small businesses. One important explanation for this *business-size bias* in dishonesty, is that people perceive large organizations as less vulnerable and less moral than small organizations, which makes it seem more justifiable to cheat them. Unlike the general principle of equality under the law, our research suggests that people operate with different moral standards depending on the size of the organization they are interacting with.

Keywords: Business size; big business; moral typecasting; dishonesty

Business-Size Bias in Moral Concern:

People are More Dishonest Against Big than Small Organizations

Equality under the law is a legal principle that is meant to ensure all people and organizations equal protection against theft, violence, and other forms of crime. If it is wrong to steal $100 from business A on the left side of the street, it is equally wrong to steal $100 from business B on the right side of the street. But what if business A is a large organization with thousands of employees, whereas business B is a local store with only five workers? The principle of equality under the law would give both businesses equal protection. Would you do the same?

When moving from legal principles to decision-making in everyday life, the intuitive nature of human psychology is often less impartial, systematically favoring close others over strangers (McManus et al., 2020) – which is probably why well-functioning societies need universal laws and rights to begin with (Acemoglu & Robinson, 2019; Hobbes, 1651; Pinker, 2012). That is, when ordinary people make moral judgments of right and wrong, that judgment is not only focused on the specific act itself (e.g., stealing or not stealing, how much is being stolen, etc.). It is also shaped by a context of social relationships between the individuals involved (Earp et al., 2021; John et al., 2014), and related judgments of harm and deservingness through the lens of mind perception (Gray & Wegner, 2009; Greenberg, 1993; Schein & Gray, 2018). *Who* is doing what to *whom?* For these reasons, asymmetries might occur when people interact with organizations of different types: Some organizations might elicit greater moral concern than others. It is therefore not obvious at all that the average person will treat business A and business B the same way in the introducing scenario above, despite the formal principle of legal impartiality.

In the current research, we will examine one overarching question across eight experiments: Can the mere *size* of an organization affect people's dishonesty toward them? Informed by previous research, we define big and small organizations based on salient size metrics such as number of employees and operating locations (Sung et al., 2022; Wirtz & McColl-Kennedy, 2010; Woolley et al., 2022). Informed by moral typecasting theory (Gray et al., 2012; Gray & Wegner, 2009), our primary hypothesis predicts a 'business-size bias' in moral concern: Due to systematic differences in the perceived vulnerability and morality of the organization, individuals will be more likely to cheat a big business than a small business for selfish gain.

Perceptions of Big Business

At the beginning of the twentieth century, a few big businesses started to dominate major industries in the USA (Chandler, 1959). During the decades that followed, big businesses have contributed positively to innovation (Acs & Audretsch, 1987), worker pay (Idson & Oi, 1999), and economic growth (Lee et al., 2013) – and in consequence, improved human welfare. More recently, big technology firms, popularly called the "FAANG" (Facebook, Amazon, Apple, Netflix, and Google), have grown exceptionally in size and impact. In the early stage of the COVID-19 pandemic, big pharmaceutical companies collaborated with leading university scientists and were able to develop safe and effective vaccines in less than a year – saving millions of lives around the world (Okereke, 2021). Thus, in modern society, big businesses can play an important role in many domains of life.

However, in the absence of government regulation, there are also risks of power abuse and market monopolization by large corporations. Businesses can become too big and too influential, leading them to pursue corporate incentives in ways that no longer align with the best interests of citizens (Thiel, 2014). In such cases, big businesses might no longer contribute to better products and technology, lower prices, better working conditions, or other benefits to society. Thus, in terms of societal impact, big businesses can be both good and bad.

Despite the mixed consequences of big businesses in society, in the public eye, there is evidence that people dislike large corporations in a more one-sided way, and even stereotype big businesses as less moral in general (Freund et al., 2023). Moreover, in a national survey in the United States, the latest Gallup Poll found that almost everyone in the representative sample (97%) reported a positive perception of small businesses, whereas a slight majority (53%) reported a *negative* view of big businesses (53%) (Saad, 2022). Why is this the case? And does it matter for how we treat these organizations as decision-makers?

Victim characteristics and dishonesty

Despite the severe economic costs, factors that drive dishonesty in the marketplace remain underexamined (Fombelle et al., 2020). The Federal Bureau of Investigation (2020) estimates the cost of non-health insurance fraud in the United States at over 40 billion USD. This costs the average American family between $400 and $700 in increased premiums. Merchandise returns fraud cost US retailers $23.2 billion in 2021 (NRF, Appriss Retail, 2022). Across the board, interventions to fight dishonesty have shown limited success (Skowronek, 2022). Extant research has tested field interventions *without* unpacking the psychological mechanisms (Hertwig & Mazar, 2022). This underscores the need to understand what drives dishonesty in the first place.

Although research in moral decision-making has examined incentives (Wang & Murnighan, 2017; Wiltermuth, 2011), the choice environment (Ayal et al., 2019; Cohn et al., 2022; Kocher et al., 2018), and the characteristics of the perpetrator (Pascual-Ezama et al., 2020; Vincent et al., 2013), the fundamental characteristics of the *victim* have been mostly overlooked. In our view, neglecting the basic characteristics of the victim of dishonesty and economic crime is akin to ignoring half the story, because acts of harm tend to involve separate categorization of moral agents and patients (Gray & Wegner, 2009). Nonetheless, recent research shows that when tempted with identical risk-free gains through dishonest behavior, individuals are more likely to cheat another person than an organization (Soraperra et al., 2019), an identifiable rather than non-identifiable victim (Yam & Reynolds, 2016), and a harmless corporation than a harmful corporation (Rotman et al., 2018). Together, these suggest that the *characteristics* of the potential victim, even in the face of equal material prospects, can play an important role in the decision to tell the truth or not.

How might dishonest behaviors of everyday people differ against big versus small businesses? Anecdotally, about 20% of members part of the anti-Walmart brand community stated that stealing from Walmart is acceptable because 'Walmart is evil' (Kucuk, 2019, p.98 in (Septianto & Kwon, 2021). However, is being evil the only reason people may cheat big businesses? In the current research, we suggest that people tend to perceive big businesses as powerful agents that are *less vulnerable* and *less moral* than small businesses, and in turn, that this perception makes it seem more acceptable to lie for personal gain when interacting with a big business. If true, this research would provide new insights into the moral psychology involved when people interact with organizational structures and decide to tell the truth or not, which is common in everyday life, and also add to recent literature suggesting that people do in fact perceive organizations as having "human-like" properties, at least to some extent (Strohminger & Jordan, 2022).

Although observational field studies (Mars, 1985) and interviews (Smigel, 1956) suggest that employees deem it more acceptable to steal from larger than smaller organizations (Greenberg, 2002), it remains unclear how pervasive this bias is in the *general* population, and whether there is a *causal* effect of organizational size on behavioral dishonesty. So, as a pre-study of this basic idea, we asked 297 American participants on Prolific Academic how important (11-point scale: 0 = Not at all, 10 = Very much) it is to be truthful towards big and small banks, tech firms, insurance companies, and retailers. On average, participants reported that it was less important to be truthful toward big ($M = 7.50$, $SD = 2.40$) than small ($M = 8.25$, $SD = 1.81$) businesses, paired sample t (296) = -9.01, $p < .001$, $d = -.52$. We also asked a nationally representative sample (USA) of 498 Prolific participants whether is more acceptable to conceal/falsify information towards big or small banks and tech firms. Although 73% of participants stated cheating neither was acceptable, 26% of people stated it was acceptable to cheat a big business vs. only 1% to a small business. Together, these present preliminary evidence of a pervasive business-size bias in moral concern that may exist across several industries, warranting a causal investigation using experiments. (Please see section "1. Pilots" in the Supplementary Online Materials for details).

Theoretical framework

In traditional economics, Becker's (1968) theory of rational crime predicts that people will behave dishonestly if the expected material benefits exceed the costs. Although dishonesty undoubtedly can be affected by incentives (Kajackaite & Gneezy, 2017), it is now clear that expected material gain is only one part of the equation. During the last 50 years or so, it has been well documented in psychology and behavioral economics that people are motivated by both economic self-interest and their moral standards, and usually try to find a balance between the two motives when making decisions (Fehr & Fischbacher, 2003; Gibson et al., 2013; Kahneman et al., 1986).

Although people sometimes lie to increase their material pay-offs, a meta-analysis of incentivized behavioral experiments shows that people actually lie "surprisingly little" compared to predictions from the traditional model in economics (Abeler et al., 2019). Moreover, people are rarely willing to maximize their dishonesty to maximize economic rewards (Abeler et al., 2019), basically leaving money on the table even when their decisions are completely anonymous and without any repercussions (Gerlach et al., 2019; Rosenbaum et al., 2014). These findings support the view that people have an internal “moral cost” of lying (Kajackaite & Gneezy, 2017), which regulates and restrains the self-interested pursuit of material gains through dishonest behavior. In other words, this moral cost of dishonesty promotes honest behavior even when acts of dishonesty are undetectable and without negative consequences for the self.

Building on this modern framework of decision-making, we suggest that people may have a moral cost of lying that also extends to *organizations*, not just other people. Second, and crucially, we suggest that the moral cost of lying is not fixed, but that it varies as a function of what type of organization the decision-maker is interacting with. If true, the question becomes what organizations people are most and least likely to cheat for personal gain, and what psychological processes can account for the difference.

Business-size bias in moral concern

When people interact with businesses, signals about the size of the business (e.g., the number of employees, branches, and revenue) are often salient (Sung et al., 2022; Thompson & Arsel, 2004; Wallach & Popovich, 2023; Woolley et al., 2022). Indeed, recent research shows that 95% of Fortune 500 companies mentioned the company’s size in their marketing communications, and 81% of consumers are somewhat aware of company sizes (Woolley et al., 2022). Extant research shows that an organization’s size signal can shape its competence vs. warmth perceptions (Yang & Aggarwal, 2019), quality inferences (Woolley et al., 2022), evaluations of corporate social responsibility (Sung et al., 2022; Wallach & Popovich, 2023) and even moral judgments (Freund et al., 2023). Taken together, this shows that size is an aspect that businesses signal frequently in the marketplace, and that people often consider organizational size metrics in consequential ways.

The primary question in the current research, is whether and how the size of an organization can affect our moral concern for them. According to the influential framework of moral typecasting theory in psychology (Gray et al., 2012; Gray & Wegner, 2009), people tend to categorize the social world as dyads of mutually exclusive units, broadly described as *agents* or *patients*. The agent is the capable performer of a specific action, whereas the patient is the experiencing receiver. In the moral domain of inflicted harm, these categories will typically take the form of an agentic perpetrator and a vulnerable victim, in which people feel more sympathy and willingness to help the vulnerable victim who is assumed to have a high sensitivity for pain and other emotional experiences but little agency to act on them (Reynolds et al., 2020; Shepherd et al., 2019). This theory of dyadic morality (Schein & Gray, 2018) has been previously found to explain differences in condemnation of the same harm inflicted on victims differing in their moral agency and patiency (e.g., men versus women; Reynolds et al., 2020). In the current research, we propose that people may think about organizations in a similar way: Based on clear differences in perceived size, decision-makers may typecast the organization they are interacting with as either an agentic perpetrator or a vulnerable victim.

Using the mere size of a business as a proxy for their intuitive suitability as moral agents or victims, we suggest that big businesses are typically categorized as powerful agents in a way that makes them seem less vulnerable as potential victims of dishonesty, although in real life theft and crime is a serious problem for big businesses too (National Retail Federation, 2022). Second, we suggest that this initial categorization process makes big businesses seem less morally deserving of our help and care than small businesses, in stark contrast to the impartial principle of equal rights.

Based on this general theoretical framework, integrating the moral typecasting of organizations with a modern perspective on decision-making, our primary hypothesis predicts a ‘business-size bias’ in moral concern: On average, people will be more willing to cheat a big business than a small business for selfish gain.

The Current Research

In the current research, the goal is to examine whether people have a different moral cost of lying to different types of organizations, expressed as greater dishonesty toward big than small businesses. To test this hypothesis, we conducted eight experiments (total N = 5,670) using both scenario-based hypothetical choices and incentivized behavioral measures. Experiments 1a-c tested if people express greater dishonesty intentions against a big than small business in marketplace decisions, and explored the potential role of perceived vulnerability of the business. Experiments 2a-c tested if people actually *behaved* more dishonestly to increase their selfish payoffs at the cost of big than small businesses. Finally, Experiments 3a-c explored size-dependent changes in perceived vulnerability and morality as a potential psychological process underlying the business-size bias in moral concern, and situated the dishonesty toward big vs. small in the context of dishonesty toward medium-sized and no-size signaling businesses.

Transparency and Openness

In these studies, we report all measures, manipulations, and exclusions. Please see Table S1C in the Supplemental Materials for a complete list of all items. Sample sizes were determined before any data analysis. The primary hypothesis and statistical analyses were pre-registered for seven out of eight experiments.

The studies were programmed using Qualtrics (*Qualtrics*, 2022). All analyses were done using the statistical software jamovi (The jamovi project, 2022) and R (R Core Team, 2013). The data, analysis scripts, and study materials are openly available on the OSF platform (anonymized link: here).

Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. Informed consent was obtained from all participants. As Norwegian laws and regulations do not require review by an institutional review board for non-medical, low-risk research with human participants, we did not submit the project to such a review.

Experiment 1a

Experiments 1a-c provided the first test of the general hypothesis of a business-size bias in moral concern, predicting greater dishonesty against big than small organizations. In a between-subjects design with hypothetical choices, participants in Experiments 1a-c were asked to imagine a scenario where they had to buy a pair of dress shoes from an apparel store for a one-off usage, and then considered the possibility of returning the shoes to get their full payment reimbursed by claiming that they were dissatisfied with the shoes. The only difference between the two conditions, was whether the apparel store was described as a big or small organization. The main outcome variable was a measure of how likely the participant would be to return their shoes by falsely claiming dissatisfaction. This experiment was pre-registered (open PDF: https://aspredicted.org/4JP_HF2).

Method

Participants

In Experiment 1a, we recruited 401 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. At the start of the study, participants were asked to indicate that they were paying attention to instructions[3] by selecting both "A great deal" and "A lot" to the question “How much do you like sports?”. We removed responses from participants who chose otherwise, that is failed the attention check (N = 8). The final dataset had 393 participants (M_{age}=34.6, SD = 12.7; 199 female, 186 male, 8 missing; 298 White, 21 Black, 27 Asian, 23 Mixed, 12 missing). A sensitivity power analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen’s d = .33 or larger in an independent t-test ($p < .05$, two-tailed). All participants were paid £0.25 for participation.

Procedure and measures

Reading a hypothetical scenario text, participants were asked to imagine attending an out-of-town wedding for which they had forgotten to bring their dress shoes. They find a store called Nimbus. Crucially, the store is *either* described as “Nimbus- a big multinational fashion company that sells all kinds of apparel. Nimbus has 1000+ stores and the company owns many popular brands” (big business condition), *or*, “Nimbus- a local store that sells clothes, shoes, and accessories. Nimbus is family-run from a small shop downtown and they sell unique fashion items” (small business condition).

This business-size manipulation was adapted from a previous study that tested opportunistic consumer behavior against a big vs. small caterer in case of service failure (Wirtz & McColl-Kennedy, 2010). Then, participants were asked to imagine that they decided to buy a pair of dress shoes and wear those at the wedding. Participants were told that

[3] We used this same attention check for Experiments 1a-c, and 2a.

the shoes fit okay, and they did not experience much discomfort. However, in this specific scenario, they already have dress shoes at home, and the business would accept product returns from dissatisfied customers within 30 days of purchase. To increase survey engagement (Strohminger & Jordan, 2022), participants were also asked what real business came to mind similar to the described business as an open-ended question[4].

Then, as the primary outcome measure, the participants were asked how likely it would be for them to return the shoes by falsely expressing dissatisfaction. Specifically, participants were asked "How likely would you be to return the shoes by expressing dissatisfaction?" on a slider scale from -100 ("No way") to 100 ("For sure").

Results and discussion

An independent sample Welch's t-test[5] showed that the likelihood of returning the shoes under false premises to a big business ($M = -7.4$, $SD = 66.6$) versus small business ($M = -18.2$, $SD = 66.2$), Welch's $t(390) = 1.61$, $p = .108$, $d = .16$, $g = .16$, 95% CI [-.04, .36] was not significantly different at the pre-specified alpha level of 5%. As a follow-up analysis, we then performed an equivalence test (Lakens, 2017) for a two-sample Welch's t-test, with equivalence bounds of $d = -.2$ and $d = .2$. The two one-sided-tests gave t-values and p-values both for the test against the upper bound, $t(390) = -.37.$, $p = .356$, and the lower bound, $t(390) = 3.59$, $p < .001$. This implies that while we can reject differences equal to or lower than $d = -.2$, we cannot reject differences equal to or higher than $d = .2$. In our view, this suggests that although we can reject the hypothesis that people are more dishonest against

[4] This was implemented the same way, before the primary outcome measure, across Experiments 1a-c.

[5] We include both Cohen's d and g throughout Experiments 1a-c because of the violation of the equal variances assumption, as indicated by Levene's being significant at $p < .05$. The estimated effect sizes are almost identical.

small businesses than big business, we cannot reject the hypothesis that people may be more dishonest against big businesses than small businesses from Experiment 1a alone.

Experiment 1b

Experiment 1a offered inconclusive evidence to support or reject our primary hypothesis of a business-size bias in moral concern. In Experiment 1b, we replicated the study using a much larger sample, to provide sufficient statistical power to detect or reject a broader range of effect sizes than what was possible in Experiment 1a. Using an identical between-subjects design, participants in Experiment 1b were asked to imagine a scenario where they had to buy a pair of shoes for a one-off usage, and then report the likelihood of returning the shoes under false premises to get their money back. The experiment was pre-registered (open PDF: https://aspredicted.org/blind.php?x=JCY_QVB).

Method

Participants

In Experiment 1b, we recruited 995 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. After removing those who failed the attention check[6] (N = 23), the final dataset had 972 participants (M_{age}=34.8, SD = 12.5; 477 female, 475 male, 1 prefer not to say, 8 missing; 758 White, 61 Black, 48 Asian, 55 Mixed, 28 Other, 22 missing). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's d = .21 or larger in an independent t-test (p < .05, two-tailed). All participants were paid £0.25 for participation. A power analysis using WebPower (Zhang et al., 2023) showed

[6] The same attention check as in Experiment 1a.

that for mediation tests, the final sample size could detect a and b paths with effect sizes (standardized estimated coefficients) as small as .15 with 91% power at two-tailed alpha = .05. Further, our final sample (N = 972) is larger than the sample size needed (N = 462) to detect small-small effects (a-path = .14, b-path = .14, direct effect = .14) with 80% power in bias-corrected bootstrap mediation models (Fritz & MacKinnon, 2007).

Procedure and measures

We used the same hypothetical scenario of committing returns fraud, and experimental manipulation of business size as in Experiment 1a. As the primary outcome measure, the participants were asked how likely it would be for them to return the shoes by falsely expressing dissatisfaction. Specifically, participants were asked "How likely would you be to return the shoes by expressing dissatisfaction?" on a slider scale from -100 ("No way") to 100 ("For sure") on an 11-point scale increasing by a value of 20 for each scale point. To explore potential mediator variables, we included a set of new measures of organizational perception: perceived likability- "How likable does Nimbus seem?" on a scale from -100 ("Not at all likable") to 100 ("Very likable"), and perceived vulnerability- "How vulnerable does Nimbus seem?" on an 11-point scale from -100 ("Not at all vulnerable") to 100 ("Very vulnerable"), increasing by a value of 20 for each scale point. These measures were placed after the main outcome measure.

Results and discussion

Consistent with our primary hypothesis of business-size bias in moral concern, an independent sample Welch's t-test showed people were more likely to return the shoes under false premises to a big business (M = 21, SD = 64.2) than a small business (M = -.94, SD =

66.0), Welch's t (969) = 5.4, $p < .001$, d = .34, g = .34, 95% CI [.21, .46]. The effect size was moderately strong, and it was highly significant.

As an exploratory analysis, we also examined whether people's organizational perception of the business can explain part of the business-size effect on dishonest intentions. When conducting a series of Welch's t-tests, these results showed that people rated the small business as more likable (M = 57.1, SD = 30.7 vs. M = 49.2, SD = 38.9, $p < .001$, d = .23, g = .23, 95%, CI [.10, .35]) and more vulnerable (M = 16.2, SD = 46.1 vs. M = -10.7, SD = 55.5, $p < .001$, d = .53, g = .53, 95% CI [.40, .66]) than the big business. When including perceived vulnerability and likeability as potential process variables in a parallel mediation test using Process model 4 (Hayes, 2018), we found suggestive evidence that the effect of greater dishonesty intentions toward big than small businesses was partly driven by reduced vulnerability (indirect effect = 4.60, BootSE = 1.29, 95% CI [2.16, 7.30]. Statistically significant support was not found for likability (indirect effect = -.39, BootSE = .51, 95% CI [-1.39, .67] perceptions of big businesses as a mediator. This suggests that perceived vulnerability may be a more important part of the explanation than likability.

Experiment 1c

We conducted Experiment 1c as a replication of the main effect observed in Experiment 1b, and to explore additional possible mediators. The only design difference in Experiment 1c was the addition of extended multi-item measures of organizational perception. These included perceived vulnerability, likability, morality, and humanness. The vulnerability and likability items were developed by the authors, whereas the morality and humanness items were adapted from Strohminger and Jordan (2022). Similar to Experiments 1a and 1b, the primary hypothesis predicted that on average, people would be more likely to falsely return merchandise to a big than a small business. As a secondary analysis, we wanted to examine different dimensions of organizational perception as potential mediators for the business-size effect on dishonesty intentions. The experiment was pre-registered (open PDF: https://aspredicted.org/XHJ_9QH).

Method

Participants

In Experiment 1c, we recruited 998 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. After removing those who failed the attention check[7] (N = 15), the final dataset had 983 participants (M_{age}= 39.8, SD = 13.6; 480 female, 489 male, 14 missing; 774 White, 69 Black, 54 Asian, 45 Mixed, 23 Other, 18 missing). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's d = .21 or larger in an independent t-test ($p < .05$, two-tailed). All participants were paid £0.25 for participation. A power analysis using WebPower (Zhang et al., 2023) showed that for mediation tests, the final sample size could detect a and b paths with effect sizes (standardized estimated coefficients) as small as .15 with 91% power at two-tailed alpha = .05. Further, our final sample (N = 983) is larger than the sample size needed (N = 462) to detect small-small effects (a-path = .14, b-path = .14, direct effect = .14) with 80% power in bias-corrected bootstrap mediation models (Fritz & MacKinnon, 2007).

Procedure and measures

We used the same hypothetical scenario of committing returns fraud, experimental manipulation of business size, and outcome measure of dishonesty intention as in Experiments 1a and 1b. To explore potential mediator variables, we included extended

[7] The same attention check as in Experiment 1a-b

measures of organizational perception after the main outcome measure, by asking people how vulnerable, powerful, profitable, likable, moral, human, and societally harmful vs. beneficial the business seemed (Example item, vulnerability: “How vulnerable does Nimbus seem?”, using 11-point rating scales from -100 (“Not at all” to 100 (“Very”). These measures were placed after the main outcome measure and the order was randomized. Informed by our theoretical framework, we combined three items (vulnerable, powerful, and profitable) into an index of perceived vulnerability (Cronbach’s alpha = .65), after reverse-coding the ratings of power and profitability perceptions.

Results and discussion

Once again consistent with the primary hypothesis of a business-size bias in moral concern, an independent sample Welch’s t-test showed that people were more likely to return the shoes under false premises to a big business ($M = 12.8$, $SD = 65.8$) than a small business ($M = .57$, $SD = 69.4$), Welch’s $t(978) = 2.82$, $p = .005$, $d = .18$, $g = .18$, 95% CI [.05, .31]. The effect size was modest, but highly significant.

We then examined to what extent the extended measures of organizational perception can explain part of the business-size effect on dishonest intentions. When conducting a series of Welch’s t-tests, the results showed that people rated the small business as more vulnerable than the big business (vulnerability index $M = -11.7$, $SD = 27.7$ vs. $M = -40.8$, $SD = 29.3$, $t(978) = 16.0$, $p < .001$, $d = 1.02$, $g = 1.02$, 95% CI [.88, .1.16]). The small business was also rated as more likable ($M = 50.7$, $SD = 35.2$ vs. $M = 38.04$, $SD = 39.8$, $t(966) = 5.3$, $p < .001$, $d = .34$, $g = .34$, 95% CI [.21, .46]), moral ($M = 35.6$, $SD = 38.2$ vs. $M = 22.0$, $SD = 39.8$, $t(979) = 5.5$, $p < .001$, $d = .35$, $g = .35$, 95% CI [.52, .78]), human ($M = 17.32$, $SD = 56.9$ vs. $M = -8.6$, $SD = 58.2$, $t(966) = 7.1$, $p < .001$, $d = .65$, $g = .65$, 95% CI [.52, .78]), and societally beneficial ($M = 34.3$, $SD = 38.3$ vs. $M = 22.7$, $SD = 41.0$, $t(980) = 4.6$, $p < .001$, d

= .29, g = .29, 95% CI [.17, .42]) than the big business. These ratings suggest that business size can reliably affect a wide range of organizational perceptions.

When including organizational perception measures as potential process variables in a parallel mediation test, the results suggest that the effect of greater dishonesty toward big than small businesses was driven by reduced vulnerability perceptions (indirect effect = 9.04, BootSE = 2.63, 95% CI [4.29, 14.62]. Statistically significant support was not found for likability (indirect effect = 1.03, BootSE = 1.08, 95% CI [-1.04, 3.21]), morality (indirect effect = -.86, BootSE = 1.16, 95% CI [-3.29, 1.25]), or humanness ([indirect effect = 1.38, BootSE = .99, 95% CI [-.29, 3.57]) perceptions of the big business as mediators. These results suggest that perceived vulnerability may be a central mechanism of the business-size bias in moral concern, consistent with an underlying process of size-based moral typecasting of organizations.

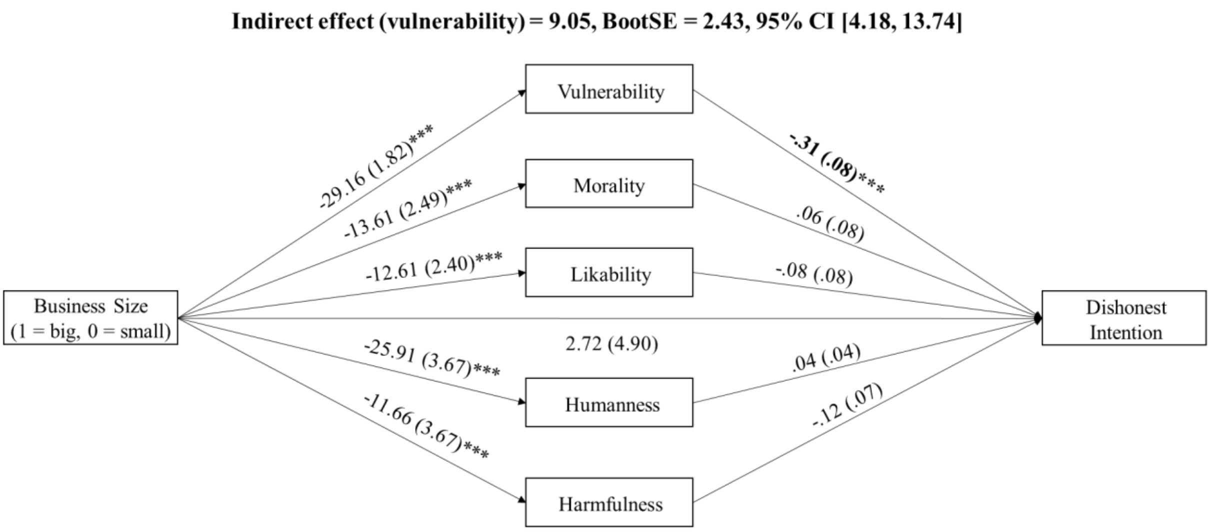


Figure 2. Mediation analysis (Experiment 1c). The big (vs. small) business organization was perceived as less vulnerable, and low (vs. high) vulnerability perception was associated with greater dishonesty intentions toward the apparel store. Path coefficients are unstandardized betas with standard errors in parentheses. *p < .05, ***p < .01, ***p < .001.

Experiment 2a

Although Experiments 1a-c provided evidence of a moderate but robust business-size bias in dishonest marketplace intentions, it remains an open question whether the size of the organization can also affect actual dishonest *behavior*. For that reason, Experiments 2a-c will provide a test of the same hypothesis using an incentivized outcome measure, predicting higher levels of behavioral dishonesty against big than small organizations. Across Experiments 2a-c, participants were presented with a decision where they could lie to increase their payoffs at the cost of either a big or small business, in a risk-free task without any reputational concerns. Here, the goal was to move beyond mere self-reports (see Baumeister et al. 2007) and provide behavioral evidence of potential differences in dishonesty against big vs. small businesses. Behavioral dishonesty in such experimental tasks has been found to reliably predict dishonesty in the field (Dai et al., 2018), suggesting that it provides a realistic indication of the basic willingness to lie or tell the truth.

In Experiment 2a, participants were asked to evaluate sample website content of a fictitious big (EveryDayMart) or small (Mike's Store) retailer in a between-subjects design. At the end of the experiment, participants were given the opportunity to increase their earnings in an incentivized die-roll guessing task, as an adapted "mind-game paradigm" (Jiang, 2013; Kajackaite & Gneezy, 2017), that was supposedly sponsored by a big (EveryDayMart) or small (Mike's Store) retailer. The main outcome variable was the proportion of participants falsely reporting having correctly guessed the score of two future die-rolls to increase their payoffs at the expense of EveryDayMart (big business condition) *or* Mike's Store (small business condition).

Because participants guess in their minds and only self-report whether they guessed correctly, we do not observe actual dishonesty at the individual level. This provides participants with the opportunity to increase their potential earnings by behaving dishonestly, without any detection risk or reputational concerns. Since we do know the statistically expected rate of correct guesses (1/6 per die-roll), however, this task enables a behavioral estimate of actual dishonesty at the group level, within each condition in the experiment. The only difference between the two conditions is who the participant thinks is paying for the additional payoffs of the game: A big (EveryDayMart) or small (Mike's Store) business. Importantly, this design creates a direct trade-off between the financial incentive to lie versus the moral costs of lying, where a difference in average dishonesty between conditions would suggest that people have a different moral cost of lying. Accordingly, we predicted that a greater proportion of participants would report a "correct guess" when interacting with a big than a small business, reflecting a lower moral cost of lying to large organizations. The experiment was pre-registered (open PDF: https://aspredicted.org/blind.php?x=54G_3SP).

Method

Participants

In Experiment 2a, we recruited 399 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. After removing those who failed the attention check[8] (N = 14), the final dataset had 385 Prolific participants (M_{age}= 27.5, SD = 4.8; 191 female, 189 male, 5 missing; 276 White, 29 Black, 31 Asian, 29 Mixed, 10 Other, 10 missing). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's w =.19 or larger in a Chi-square test ($p < .05$, two-tailed). All participants were paid £0.50 for participation. Participants could also increase their payoffs by up to an additional £6 depending on the outcome of the die score, their reporting if they guessed correctly, and being chosen in the random lottery.

[8] The same attention check as in Experiment 1a-c

Procedure and measures

We used the cover of a website content evaluation task to manipulate the size perception of the target business. In a deception-based design, participants were asked to evaluate short texts supposedly from the "About Us" section of either EveryDayMart (big business condition) or Mike's Store (small business condition). The About Us section described EveryDayMart (big business condition) as: "one of the biggest retailers in the country. The company has over 5,000 stores across all states, and employs more than 1.4 million workers.", whereas Mike's Store (small business condition) was described as: "a small corner store. The store operates on a vibrant street downtown and is run by a local family with the help of part-time workers." To increase survey engagement (Strohminger & Jordan, 2022), participants were asked an open-ended question about what real business comes to mind similar to the described business[9].

After exposure to either the big or small business description, participants responded to some filler items and rating questions about the story description, before they reported their organizational perceptions of the target business (please see Table S2A in the Supplemental Materials for a complete overview of all measures).

Getting to the behavioral measure of dishonesty, which was the primary dependent variable at the end of the experiment, participants in both conditions were told that the store whose website content they evaluated (EveryDayMart or Mike's Store) was sponsoring a bonus game where ten lucky winners could win up to £6 depending on the outcome. This bonus game was an adapted "mind-game paradigm" (Jiang, 2013), a well-established incentive-aligned guessing task. In short, participants were asked to guess the score of a future die-roll by picking a number from 1-6 *in their minds,* and only self-reporting after the

[9] This was implemented the say way, right after the description, across Experiments 2a-c.

die-roll if their guess was correct or not. If they report a correct guess, they can win a reward corresponding to the score of the die (1 = £1, 2 = £2, and so on). Participants were presented with the bonus task instructions and explicitly told that the target business would pay the bonus amount to 10 participants randomly chosen by lottery. Although the description of the source of the reward was deception-based, the payout of the reward by having 10 participants randomly chosen to actually receive their bonus award was truthful.

Note that the only systematic difference in the scenarios across the two conditions is the description of the target business as either big or small. As there is a 1 in 6 statistical chance to correctly guess the result of a fair die-roll, the proportion of reported correct guesses above 16.7% is treated as the estimated dishonesty level in each condition. If there should be a significant difference between reported correct guesses in the two conditions (above the expected rate of actual correct guesses of 16.7%), that would indicate a systematic difference in behavioral dishonesty.

Results and discussion

Consistent with our primary hypothesis of a business-size bias in moral concern, a Chi-square test revealed that the proportion of reported correct guesses was higher when the bonus task was sponsored by a big than a small business, $\chi^2(1, 385) = 7.24$, $p = .007$, odds ratio (OR) = 1.74, 95% CI [1.16, 2.61]. Compared to the statistically expected proportion of 16.7 % correct guesses (1/6), 53.8% of all participants reported a correct guess. Crucially, 60.5% of participants reported a correct guess when the game was sponsored by a big business (EveryDayMart) whereas 46.8% of participants reported a correct guess when the game was sponsored by a small business (Mike's Store). Subtracting the 16.7% of participants expected to have actually guessed the correct die-roll, we can infer that 43.8% of participants were dishonest against the big business, compared to only 30.1% of participants in the small-business condition, $\chi^2(1, 385) = 7.63$, p = .006, OR = 1.80, 95% CI [1.18, 2.74]. This business-size effect on behavioral dishonesty corresponds to a treatment difference of 14 percentage points, and is highly significant.

For the second bonus task, the proportion of reported correct guesses was also higher when the bonus task was sponsored by a big versus small business, $\chi^2(1, 385) = 3.99$, p = .046, OR = 1.51, 95% CI [1.01, 2.25]. 48.3 % of all participants reported a correct guess in the second round. Crucially, 53.3% of participants reported a correct guess when the game was sponsored by a big business, whereas 43.2% of participants reported a correct guess when the game was sponsored by a small business. Here, we can infer that about 31.6% of participants were dishonest against the big business, compared to 26.5% of participants in the small-business condition, $\chi^2(1, 385) = 4.55$, p = .043, OR = 1.60, 95% CI [1.04, 2.48]. The effect was weaker than it was for the first die-roll, corresponding to a treatment difference of 6 percentage points, but it was still statistically significant at the conventional alpha level of 5%.

Experiment 2b

Experiment 2a showed that people actually behaved more dishonestly against a big than a small business. However, it is possible that the name “Mike’s Store”, although it follows the common nomenclature of small-business names, might have also made the business seem more “human” and/or made the victim more identifiable. Therefore, we conducted Experiment 2b as a conceptual replication of the main effect observed in Experiment 2a using more similar organization names between conditions, and also adapted the scenario to the new context of hotels.

In a between-subjects design, participants were asked to evaluate the website content of a fictitious big (Indigo Hotels and Resorts) or small (Indigo Hotel) hotel. Then, participants completed the same honesty task as Experiment 2a, where they could lie to increase their payoffs at the expense of either the big or small business (depending on the condition). Like before, we predicted that a greater proportion of participants would behave dishonestly to increase their payoffs against a big than a small business. The experiment was pre-registered (open PDF: https://aspredicted.org/blind.php?x=WVY_GV8).

Method

Participants

In Experiment 2b, we recruited 600 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. At the start of the study, participants were asked to what extent they agreed with the statement "I swim across the Atlantic Ocean to get to work every day." We removed responses from participants (N = 12) who indicated they agreed or strongly agreed with the statement, as swimming across the Atlantic Ocean to get to work every day is an impossible feat. After removing those who failed the attention check; the final dataset had 588 Prolific participants (M_{age}= 27.6, SD = 4.76; 283 female, 283 male, 17 other, 1 preferred not to say; 407 White, 35 Black, 61 Asian, 29 Mixed, 17 Other, 12 missing). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's w = .16 or larger in a Chi-square test ($p < .05$, two-tailed). All participants were paid £0.40 for participation and could increase their payoffs by up to an additional £6 depending on the outcome and their reporting of the die-roll.

Procedure and measures

Similar to Experiment 2a, participants read short texts supposedly from the "About Us" section of either Indigo Hotels and Resorts (big business condition) or Indigo Hotel (small business condition). The About Us section described Indigo Hotels and Resorts (big business condition) as: “With more than 8,000 hotels and over 14,000 rooms, Indigo Hotels & Resorts is one of the world’s largest hotel chains.”, whereas Indigo Hotels (small business condition) was described as: “With 42 rooms in a convenient downtown location, Indigo Hotel caters to travelers of different needs.”

After exposure to the big or small description as per their randomly assigned condition, participants responded to the same measures as in Experiment 2a. This includes filler tasks and organizational rating measures (please see Table S2B in the Supplemental Materials for a complete overview of all measures), and a behavioral dishonesty task as the primary outcome measure: Participants could misreport their actual guess in a die-roll task to win a bonus award that was either covered by the big *or* the small business, depending on the condition. We do not report any mediation analyses of experiments 2a-c, since the incentivized guessing task does not provide an individual measure of actual dishonesty (but instead, an estimate of behavioral dishonesty at the group level).

Results and discussion

A Chi-square test found that the proportion of reported correct guesses was not statistically different when sponsored by a big (Indigo Hotels & Resorts) rather than a small business (Indigo Hotel), $\chi^2(1, 588) = .003$, p = .954. Overall, 52.6 % of all participants reported correct guesses. Divided by condition, 52.7% of participants reported correct guesses when the game was sponsored by Indigo Hotels and Resorts (big business condition), whereas 52.4 % of participants reported correct guesses when the game was sponsored by Indigo Hotel (small business condition). An equivalence test of the two proportions showed that the estimated 90% equivalence bounds [-.07, .07] were significantly different from the

TOST lower (Z = -2.37, p = .06, test bound = -.1) and TOST upper (Z = 2.48, p = .009; test bound = .1), suggesting evidence in favor of the null hypothesis.

The lack of any meaningful difference between conditions in Experiment 2b might suggest that the business-size bias in behavioral dishonesty is less robust than indicated by Experiment 2a. However, a competing explanation for the current null result, is that the size perception of the hotel might not have been sufficiently clear or *salient* for the participants during the decision-making part of the honesty task[10]. So, we conducted a binary logistic regression with reported correct guess (0 = no, 1 = yes) as the outcome and predictors of experimental condition (0 = small, 1 = big), size perception manipulation check measure, and their interaction. The estimate of both the experimental condition (Z = 2.08, p = .037, OR = 2.50, 95% CI [1.06, 5.91]) and its interaction term with size perceptions (Z = -2.73, p = .006, OR = .74, 95% CI [.60, .92]) were significant, whereas size perception in itself was not (Z = 1.46, p = .143, OR = 1.07, 95% CI [.98, 1.17]). This suggests that the differences between the conditions may become nullified for participants who perceived our manipulated small business to be also big in size like the big business.

Experiment 2c

In our third and final behavioral experiment, we used a clearer and presumably stronger manipulation of business size by varying the specific *name* of the hotels as well, to ensure a more sensitive test of the primary hypothesis. In a between-subjects design, participants were asked to evaluate the website content of a fictitious big (Grand Hotels and Resorts) or small (Little Hotel Parkside) hotel. Keeping the descriptions of the big and small hotels the same as in 2b, the only difference in materials was the names of the hotels to

[10] See Figure S2B.2 in Supplemental Materials for an illustration of the heterogeneity of perceptions of the small business manipulation.

clearly signal their size. Keeping the same primary hypothesis as in Experiments 2a-b, in Experiment 2c we predicted that more people would behave dishonestly to increase their payoffs against a big than a small business. The experiment was pre-registered (open PDF: https://aspredicted.org/blind.php?x=WWX_L7P).

Method

Participants

We recruited 599 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. After removing those who failed the attention check[11] (N = 8), the final dataset had 591 Prolific participants (M_{age}= 38.7, SD = 13.8; 355 female, 218 male, 17 other; 450 White, 40 Black, 30 Asian, 43 Mixed, 15 Other, 12 missing). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's w =.15 or larger in a Chi-square test ($p < .05$, two-tailed). All participants were paid £0.50 for participation.

Procedure and measures

Just like in Experiments 2a-b, participants were asked to evaluate short texts supposedly from the "About Us" section of either Grand Hotels and Resorts (big business condition) or Little Hotel Parkside (small business condition). The About Us section described Grand Hotels and Resorts (big business condition) as: "With more than 8,000 hotels and over 14,000 rooms, Grand Hotels & Resorts is one of the world's largest hotel chains.", whereas Little Hotel Parkside (small business condition) was described as: "With 42 rooms in a convenient downtown location, Indigo Hotel caters to travelers of different needs.".

[11] The same attention check as in Experiment 2b.

After exposure to the big or small description as per their randomly assigned condition, participants responded to the same set of measures as in Experiment 2a-b (please see Table S2C in the Supplemental Materials for a complete overview of all measures). As the primary dependent variable, the experiment ended with a behavioral measure of dishonesty using an adapted version of the mind-game paradigm (an incentivized die-roll guessing task).

Results and discussion

Consistent with our primary hypothesis of a business-size bias in moral concern, a Chi-square test revealed that the proportion of reported correct guesses was higher when the bonus task was sponsored by a big versus small business, $\chi^2(1, 592) = 15.4$, $p < .001$, odds ratio = 1.92, 95% CI [1.38, 2.66]. Compared to the statistical expectation of 16.7 % of participants reporting a correct guess (1/6), 49.4% of all participants reported a correct guess. Divided by condition, 57.1% of participants reported a correct guess when the game was sponsored by a big business (Grand Hotels & Resorts), whereas 41.0% of participants reported a correct guess when the game was sponsored by a small business (Little Hotel Parkside). By subtracting the 16.7% of participants who are expected to actually guess correctly, we can infer that about 40.4% of participants were dishonest against the big business compared to only 24.3% of participants in the small-business condition, $\chi^2(1, 592) = 17.56$, $p < .001$, odds ratio = 2.12, 95% CI [1.49, 3.02]. This business-size effect on behavioral dishonesty corresponds to a treatment difference of 16 percentage points, and was highly significant. See Figure 3 for an illustration of the differences in dishonesty across Experiments 2a-c.

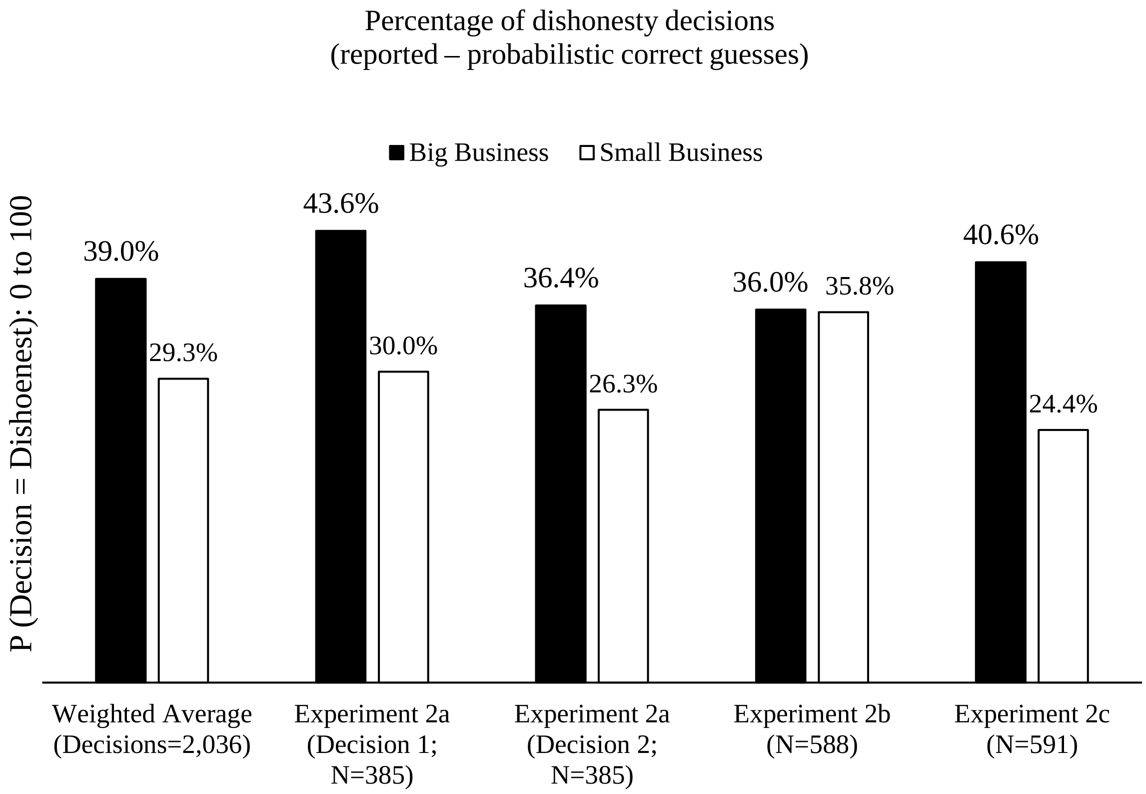


Figure 3. Proportions of participants behaving dishonestly against big vs. small businesses (Experiments 2a-c; N = 1651, 2,306 decisions). On average across the three experiments, behavioral dishonesty was almost 10 percentage points higher when people thought they were interacting with a big (39.0%) than a small (29.3%) business. Dishonesty was calculated by subtracting 16.7% (probabilistic proportion of correctly guessing the future roll of a six-sided die) from the actual proportion of self-reported correct guesses in each condition.

Experiment 3a

In Experiment 3a, we aimed to replicate the business-size bias in moral concern in a new setting, and to further explore the psychological processes of size-dependent differences in perceived vulnerability and morality. Whereas Experiments 1a-c (dishonesty intentions) and 2a-c (behavioral dishonesty) provided robust evidence for the main effect of organizational size, the identification of the underlying process has so far remained suggestive. In Experiments 1b-c, we observed a mediational role of perceived vulnerability, but the multi-item scale used in Experiment 1c showed only moderate reliability.

In a between-participants design, participants in Experiment 3a were asked to imagine a hypothetical scenario where they spotted an underbilling error at a self-service checkout of a store (for a similar method, see Giroux et al. 2022). We manipulated store size experimentally between two conditions, describing it either as big (Colossal Mart- a big chain retailer) or small (Lil Mart- an independent small retailer). The main outcome variable was a binary choice between reporting the underbilling error to the store (i.e., being honest at a financial cost) or leaving without reporting (i.e., saving money by being dishonest). We also measured perceived vulnerability (index a = .86) and morality (index a = .87) of the organization using more conventional rating scales (0-100) than in our first experiments, as potential mediator mechanisms that may identify an underlying process of *organizational* moral typecasting. Like before, our primary hypothesis predicted that a greater proportion of participants would choose to be dishonest (i.e., leave without reporting a self-serving underbilling error) when interacting with a big business than a small business.

Method

Participants

We recruited 600 participants (approval rate of 95% and above for at least 50 task completions) based in the USA from Prolific Academic. After presenting participants with the scenario to imagine passing through a town, they were asked to choose which activity they were thinking of between “Imagining passing through a town” and “Swimming in the Arctic Ocean”. Choosing the first option indicated attentive following of instructions in the survey as all participants were told beforehand to imagine passing through a town in the scenario presented. We removed responses from participants who chose otherwise (N = 1), that is failed the attention check. Further, we removed responses from participants who stated they had never used a self-service checkout (N = 14). Keeping or removing these responses did not change results in any meaningful way.

The final dataset had 585 participants (M_{age}= 29.6, SD = 6.0; 284 female, 289 male, 2 other, 4 prefer not to say; 400 White, 61 Black, 49 Asian, 50 Mixed, 23 Other). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's w = .16 or larger in a Chi-square test ($p < .05$, two-tailed). All participants were paid £0.40 for participation. A power analysis using WebPower (Zhang et al., 2023) showed that for mediation tests, the final sample size could detect a and b paths with effect sizes (standardized estimated coefficients) as small as .2 with 93% power at two-tailed alpha = .05. Further, our final sample (N = 585) is larger than the sample size needed (N = 462) to detect small-small effects (a-path = .14, b-path = .14, direct effect = .14) with 80% power in bias-corrected bootstrap mediation models (Fritz & MacKinnon, 2007).

Procedure and measures

Participants were asked to imagine they were passing through a town they had not been to before and would probably never return to in the future, and that they went out to shop for some items. This description was included to minimize perceived social pressure and reputational consequences in this scenario (e.g., as opposed to attending a grocery store in one's own neighborhood or the canteen at work). After searching on Google for stores nearby, they came across either Colossal Mart or Lil Mart, depending on the experimental condition. Colossal Mart (big business condition) was described like this: "Colossal Mart seems like a large chain retailer that sells thousands of products. It has 1000+ stores". Lil

Mart (small business condition) was described like this: “Lil Mart seems like a small independent retailer that sells several necessary products.” To increase survey engagement (Strohminger & Jordan 2022), participants were asked an open-ended question about what real business comes to mind similar to the described business.

In both conditions, participants read that when they were completing their shopping trip using a cashier-less self-service checkout terminal, they noticed a billing error and realized that they had been billed only $1.89 for an item usually priced at $18.90. As the primary outcome measure, participants were then asked to make a choice in this hypothetical scenario: “What would you do?”, from (1) “report the billing error to a Colossal / Lil Mart employee”, or (2) “leave Colossal / Lil Mart without reporting the billing error”. On the next page, this choice was followed by a question of how sure they were that they would report (leave) with the measure “How sure are you that you would ____”, with the participant’s choice from the previous page piped in to complete the question. This two-step procedure is adapted from previous research on hypothetical dishonesty choice tasks (LaMothe & Bobek, 2020).

Right after the decision section, participants were asked to respond to four statements inspired by the DeNiAL framework (Skowronek, 2022) to explore potential moral rationalization by participants (see Table S3A in the Supplementary Materials for details). But importantly, as potential mediators, participants were asked about their perception of the target business. Here, they responded to the measure “To what extent does Colossal / Lil Mart seem …” with specific rating items: rich, powerful, and vulnerable to form the vulnerability index (Cronbach’s a = .862); and virtuous, prosocial, and ethical to form the morality index (Cronbach’s a = .875), using slider scales from 0 (“Not at all”) to 100 (“Completely”). The order of the six items was randomized. We note that unlike our first experiment using multi-item mediator measures of this kind, the reliability was good for both of these mediators measured in Experiment 3a, and the scale range was also more conventional and intuitive (0-100 increasing with 10 for each scale point, rather than -100 to +100 increasing with 20 for each scale point).

Results and discussion

Consistent with our primary hypothesis of a business-size bias, a Chi-square test revealed that participants were more likely to make the dishonest choice of leaving the store without correcting their bill when the scenario placed them in a big-business context ("Colossal Mart": 78.6%), rather than a small-business context ("Lil Mart": 55.7 %), χ^2(1, 584) = 34.9, p < .001, odds ratio = 2.93, 95% CI = 2.04, .4.21]. The difference corresponds to a treatment effect of 23 percentage points, and is highly significant.

Regarding the self-reported certainty of their decision (leave or report), participants felt significantly more certain about their decision toward Colossal Mart (M = 86.13, SD = 16.9) than Lil Mart (M = 82.4, SD = 18.8), Welch's t (574) = 2.52, p = .01, d = .21, g = .21, 95% CI [.05, .37]. This result might indicate that people experienced more ambivalence when interacting with a small than a large business, possibly reflecting a higher moral cost of lying. However, this exploratory and non-predicted result should be interpreted as suggestive.

To examine whether differences in organizational perception can explain the business-size effect on dishonest intentions, separate Welch's t-tests show that Lil Mart was perceived as both more vulnerable, (index M = 61.0, SD = 18.2 vs. M = 19.9, SD = 16.5, t (574) = 28.59, p < .001, d = 2.37, g = 2.36, 95 % CI [2.15, 2.58]) and more moral (index M = 53.9, SD = 16.5 vs. M = 40.5, SD = 22.6, t (538.9) = 8.17, p < .001, d = .68, g = .67, 95% CI [.51, .84] than Colossal Mart. These results provide strong evidence for size-dependent differences in the perceived vulnerability and morality of big versus small organizations. Interestingly, the perceptions of vulnerability and morality were strongly correlated

(Pearson's r (df = 582) = .470, p < .001, which might suggest a potential psychological process that could be studied as a serial mediation model.

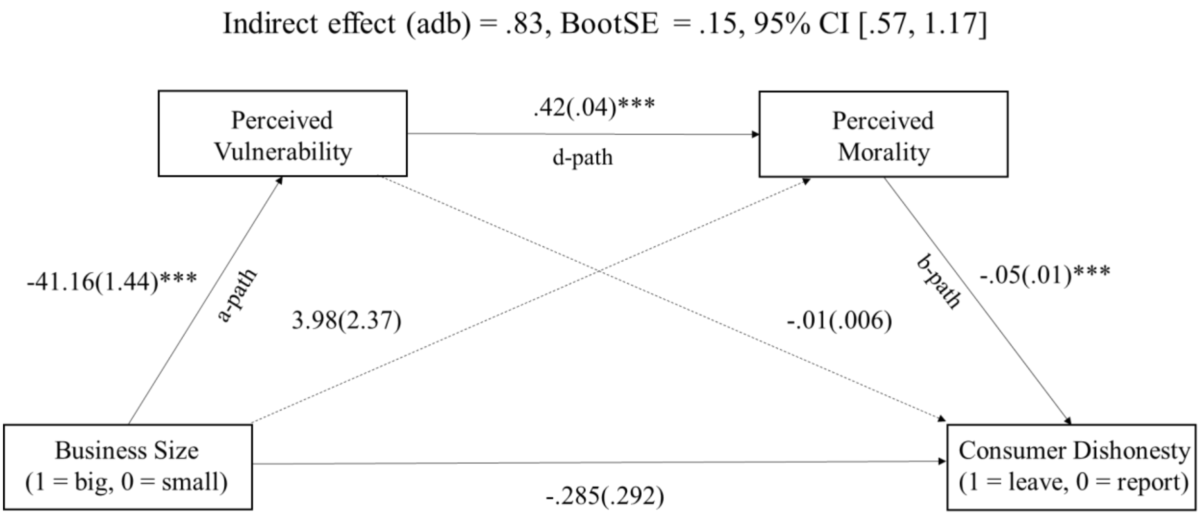


Figure 4. Mediation analysis (Experiment 3a). The big (vs. small) business organization was perceived as less vulnerable and less moral, and these perceptions were associated with greater dishonesty intentions toward big businesses when modeled as a serial mediation process. Path coefficients are unstandardized betas with standard errors in parentheses. *p < .05, ***p < .01, ***p < .001.

Informed by moral typecasting theory and our broader framework, we conducted a serial mediation analysis to examine the interplay of perceived vulnerability and morality, modeled as potential drivers of the business-size effect on dishonesty intentions. The results were consistent with this model, as there was a significant indirect effect of that kind, Ind3 (adb) = .83, bootstrapped SE = .15, 95% CI [.57, 1.17]; Process model 6: Hayes 2018). Specifically, the big (vs. small) organization was perceived as less vulnerable, low vulnerability perception was associated with low morality perception, and low perception of the morality of the organization was associated with more dishonesty directed toward the organization. Interestingly, there was no direct effect from vulnerability to dishonesty when morality perception was controlled for in this serial mediation analysis (see Figure 4 for illustration). Seen as a whole, the mediation results provide suggestive evidence that greater dishonesty against big businesses is driven by a combined process of perceiving these organizations as less vulnerable *and* less moral.

Experiment 3b

Experiment 3b was a pre-registered replication of the main effect and serial mediation of business-size bias in moral concern through vulnerability and moral perceptions as revealed in Experiment 3a, with two additional conditions. In a between-participants design, participants were asked to imagine the same hypothetical scenario as 3a. We experimentally manipulated store size between the four conditions, describing it as big (Colossal Mart- a big chain retailer), small (Lil Mart- an independent small retailer), medium-sized (Mid Mart- a medium-sized retailer), or without any overt size signal (Z Mart). The main outcome variable was a binary choice between reporting the underbilling error to the store (i.e., being honest at a financial cost) or leaving without reporting (i.e., saving money by being dishonest). We also measured perceived vulnerability. Like before, our primary hypothesis predicted that a greater proportion of participants would choose to be dishonest (i.e., leave without reporting a self-serving underbilling error) when interacting with a big business than a small business, and that the effects will be serially mediated through perceived vulnerability and morality. The experiment was pre-registered (open PDF: https://aspredicted.org/8PR_MW9).

Method

Participants

We recruited 1215 participants (approval rate of 95% and above for at least 100 task completions) based in the USA from Prolific Academic. After removing those who failed the attention check[12] (N = 3) and those who have never used a self-service checkout (N = 33), the final dataset had 1178 participants (M_{age}= 37.76, SD = 14.27; 579 female, 572 male, 2 other; 798 White, 161 Black, 75 Asian, 161 Black, 89 Mixed, 45 Other, 10 missing). The removal of these responses did not change results in any meaningful way.

A sensitivity analysis in G*Power showed that the final sample provided a 90% chance of detecting a simple main effect of Cohen's w = .13 or larger in a Chi-square test (p < .05, two-tailed). All participants were paid £0.45 for participation. A power analysis using WebPower (Zhang et al., 2023) showed that for mediation tests, the smallest two cells as a pair (N = 585) could detect a and b paths with effect sizes (standardized estimated coefficients) as small as .2 with 93% power at two-tailed alpha = .05. Further, our smallest two cells as a pair (N = 585) is larger than the sample size needed (N = 462) to detect small-small effects (a-path = .14, b-path = .14, direct effect = .14) with 80% power in bias-corrected bootstrap mediation models (Fritz & MacKinnon, 2007).

Procedure and measures

Participants were asked to imagine the same scenario as in S3b of whether they would report an error of being billed $1.89 for an $18.90 item at the self-service checkout of a store. The big and small size signals were exactly the same as in Experiment 3a. The store in the medium-sized condition was described as "Mid Mart seems like a medium-sized retailer that sells a wide range of necessary products.", and in the no-size condition as "Z Mart seems like a retailer that sells a variety of necessary products."

The same as in Experiment 3a, participants were then asked to make a choice in this hypothetical scenario: "What would you do?", from (1) "report the billing error to a Colossal/ Lil/ Mid/ Z Mart employee", or (2) "leave Colossal/ Lil/ Mid/ Z Mart without reporting the

[12] The same attention check as in Experiment 3a.

billing error”, and then how sure they were that they would report (leave) with the measure “How sure are you that you would ____”, with the participant’s choice from the previous page piped in to complete the question.

As potential mediators, right after completing the choice section, participants were asked about their perception of the target business. Here, they responded to the measure “To what extent does Colossal/ Lil Mart seem …” with specific rating items: profitable, powerful, and vulnerable to form the vulnerability index (Cronbach’s a = .715); virtuous, prosocial, and moral to form the morality index (Cronbach’s a = .839); and single items of “human”, “likable” and “big” using slider scales from 0 (“Not at all”) to 100 (“Completely”). The order of the nine items was randomized.

Before the demographic measures, we also asked what percentage of participants taking this survey and faced with the same decision would choose to leave a Colossal/ Lil/ Mid/ Z Mart without reporting the under-billing error to explore if perceived social norms also varied for different target businesses.

Results and discussion

Consistent with our primary hypothesis of a business-size bias, a Chi-square test revealed that participants were more likely to make the dishonest choice of leaving the store without correcting their bill when the scenario placed them in a big-business context (“Colossal Mart”: 69.0%), rather than a small-business context (“Lil Mart”: 48.1 %), χ^2(1, 588) = 26.51, p < .001, odds ratio (OR) = 2.40, 95% CI = 1.72, 3.37]. The difference corresponds to a treatment effect of almost 21 percentage points, and is highly significant.

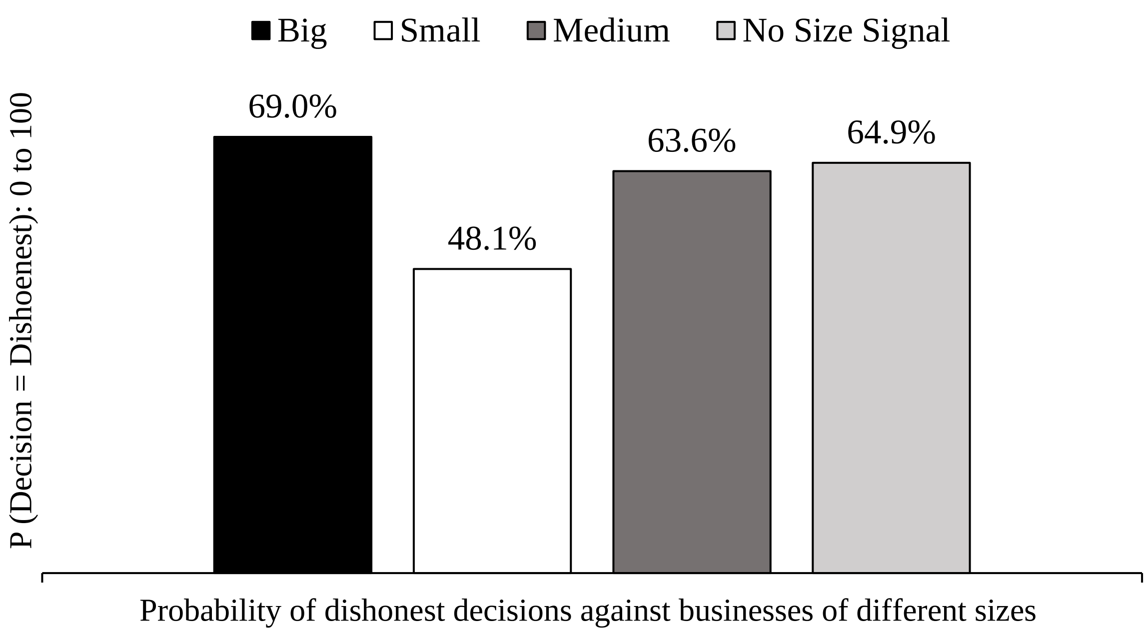


Figure 5. Proportions of participants intending a dishonest choice against businesses varying in size. Individuals were less likely to cheat small businesses than big, medium, and no-size signaling businesses (all p's < .001, odds ratios > 1.89. The likelihood of dishonesty toward either medium and no-size signaling businesses was no different from big, suggesting it is the *small* size of the business that restrained the dishonesty of individuals.

A binomial logistic regression with participant's choices as the outcome and experimental condition as the predictor showed that the business-size bias is attributable to the small-size signal as pairwise comparisons with medium and no size rendered significant results for small (p's <.001). Statistically significant support was not found for pairwise comparisons between big vs. medium ($p = .164$) and big vs. no-size ($p = .282$). Alternatively, compared to a small business (48.1%), individuals were more likely to make a dishonesty choice against big (69.0%: OR = 2.40, 95% CI [1.72, 3.37]), mid (63.6%: OR = 1.89, 95% CI [1.35, 2.62]), and no-size (64.9%: OR = 1.99, 95% CI [1.43, 2.77] signaling businesses. See Figure 5 for illustration.

To examine whether differences in organizational perception can explain the business-size effect on dishonest intentions, separate post hoc t-tests show that Lil Mart was perceived as both more vulnerable, (index $M = 53.8$, $SD = 16.0$ vs. $M = 23.7$, $SD = 15.8$, t $(1174) = 28.59$, $p < .001$, $d = 1.70$, $g = 1.70$, 95 % CI [1.52, 1.88]) and more moral (index $M = 57.4$, $SD = 16.0$ vs. $M = 45.7$, $SD = 22.4$, $t\ (1174) = 7.31$, $p < .001$, $d = .60$, $g = .60$, 95% CI [.44, .77] than Colossal Mart. These results replicate the results from Experiment 3a for size-dependent differences in the perceived vulnerability and morality of big versus small organizations. See Figures S3B.3 and S3B.4 for all pairwise comparisons across the four conditions.

In terms of the psychological process, we first conducted a parallel mediation analysis to test how different organizational perceptions were uniquely associated with differences in dishonest intentions toward big vs. small businesses. The results showed that the indirect mediation effects for both perceived vulnerability (indirect effect = .49, BootSE = .19, 95% CI [.13, .89]) and morality (indirect effect = .39, BootSE = .10, 95% CI [.21, .62]) were statistically significant. Statistically significant support was not found for perceived likability (indirect effect = -.09, BootSE = .08, 95% CI [-.27, .06]) and humanness (indirect effect = .01, BootSE = .10, 95% CI [-.18, .22]) as mediators.

Then, we conducted the same serial mediation analysis as Experiment 3a, predicting that the business-size bias in dishonesty operates through a process of perceived vulnerability and morality (also pre-registered). The results were consistent with this model, as there was a significant indirect effect of that kind, Ind3 (adb) = .195, bootstrapped SE = .061, 95% CI [.015, .277]; Process model 6: Hayes 2018). Specifically, the big (vs. small) organization was perceived as less vulnerable, low vulnerability perception was associated with low morality perception, and low perception of the morality of the organization was associated with greater dishonesty toward that organization. (see Figure 6 for illustration).

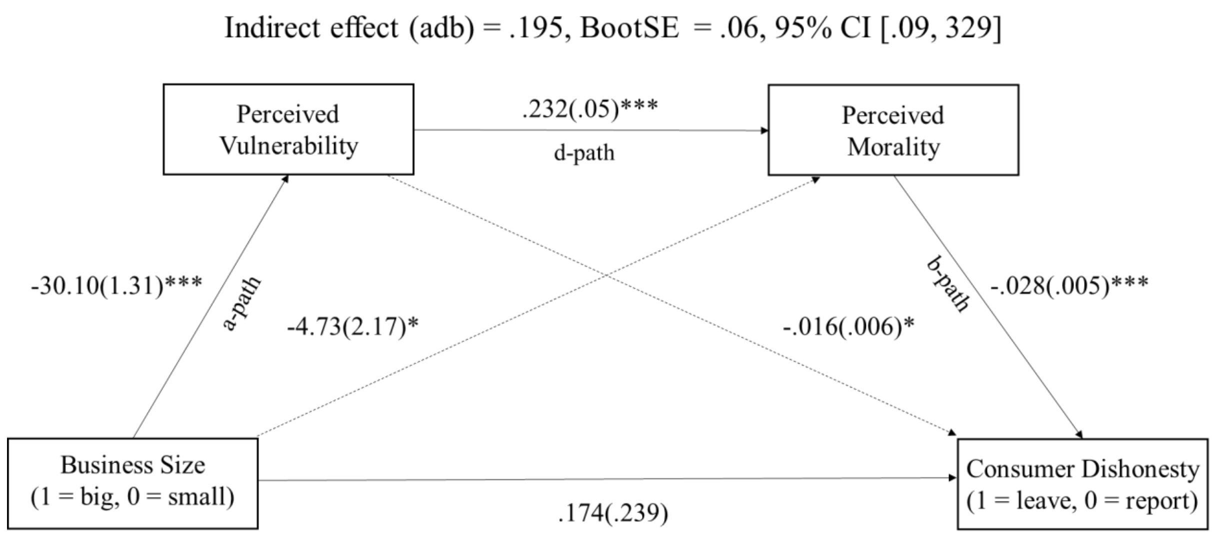


Figure 6. Mediation analysis (Experiment 3b). The big (vs. small) business organization was perceived as less vulnerable and less moral, and these perceptions were associated with greater dishonesty intentions toward big businesses when modeled as a serial mediation process. Path coefficients are unstandardized betas with standard errors in parentheses. *p < .05, ***p < .01, ***p < .001.

Taken together with the mediation results in Experiment 3a, we find more support for our proposal that greater dishonesty against big businesses is driven by a combined process of perceiving these organizations as less vulnerable *and* less moral. Although we find this explanation appealing and plausible at the theoretical level, we acknowledge that this type of mediation analysis does not rule out other, alternative models in accounting for the observed main effect.

Internal meta-analysis: The business-size bias across all 8 experiments

At the end of this research project, we performed an internal meta-analysis to estimate the general effect of the business-size bias across all 8 experiments in the current paper (N =

5,465[13]). Nine main effects were included in the analysis, using effect sizes in Cohen's d as the outcome measure. A random-effects model found that the standardized estimate of the effect was significantly different from zero, coefficient = .306, z = 5.06, p < .001, 95% CI [.187, .424]), with an estimated average treatment effect of d = .31. Thus, when integrating all current experiments in one meta-analytic test, we find robust evidence for an overall effect of greater dishonesty toward big than small businesses, with a moderately strong average effect size of 31% of a standard deviation (see Figure 7 for illustration).

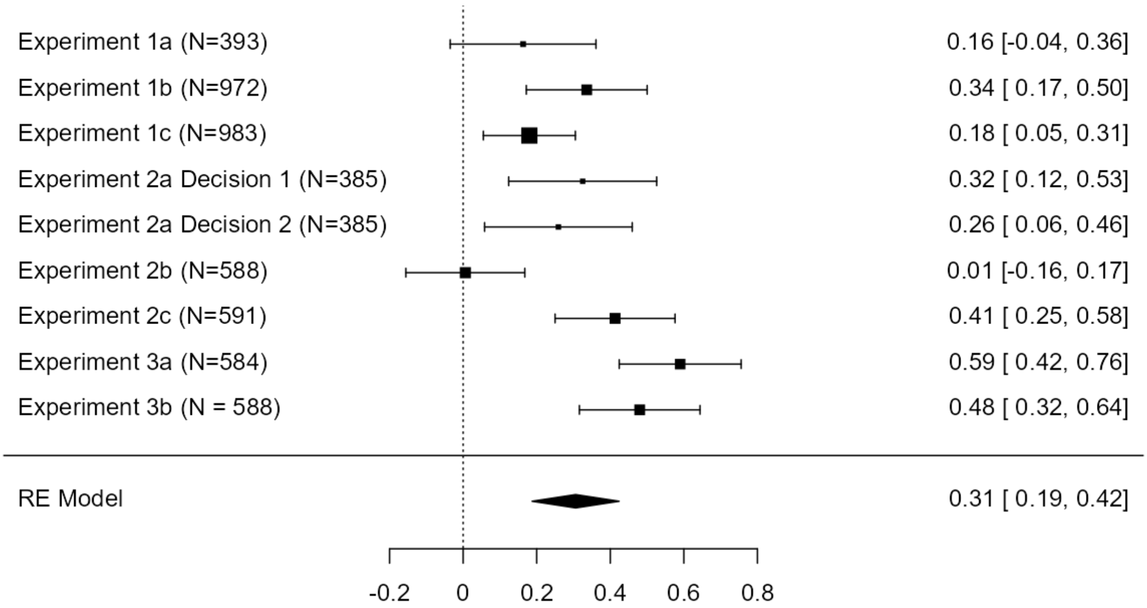


Figure 7. Meta-analysis (k = 9, N = 5,465). The forest plot shows that across all 8 experiments, on average people both intend and actually behave more dishonestly when interacting with big rather than small businesses (standardized estimate = .306, z = 5.06, p < .001, 95% CI [.187, .424].).

[13] This number is determined by simply summing all cells where participants were either assigned to a big vs. small condition. This means that it excludes the mid-size and no-size conditions from Experiment 3b. Also, the N of Experiment 2a is doubled as participants made decisions in two rounds.

General Discussion

Although the principle of equality under the law affords organizations of different sizes the same rights and equal protection against economic crime, the layperson may find it more acceptable to steal from a big than a small business. Informed by moral typecasting theory (Gray & Wegner, 2009), we therefore proposed the hypothesis of a ‘business-size bias’ in moral concern.

Across eight experiments and a total of 5,670 participants, the results provided robust evidence for the predicted business-size bias: On average, people expressed greater dishonesty toward big than small businesses. We found the same effect both for dishonest intentions in hypothetical scenarios and for dishonest *behavior* in incentivized decisions, with an overall meta-analytic effect size corresponding to 31% of a standard deviation (Cohen’s d = 0.31). This suggests that when people make consequential decisions regarding telling the truth or lying, their moral cost of dishonesty is not fixed, but is partly a function of the *size* of the organization they are interacting with. As a possible implication, large and growing organizations might be facing a moral “size penalty” in the marketplace, by attracting consumers that act in somewhat less honest and trustworthy ways than they would have done toward a small business. To that end, our findings add to the nascent literature on how *victim characteristics* can affect dishonesty (Rotman et al., 2018; Soraperra et al., 2019; Yam & Reynolds, 2016). Practically, our findings offer implications both for organizations and policymakers, suggesting the need for strategies to mitigate inherent biases in consumer morality. For small businesses, the current research suggests that they might actually have an advantage in some settings, as people might be more reluctant to steal or lie to them for personal gain.

In terms of psychological mechanisms, perhaps the least surprising explanation would be if the business-size bias in moral concern could be accounted for by mere *likeability*. Since there is evidence that people tend to like small businesses better than big businesses, perhaps this difference can explain the higher willingness to cheat a big business. In contrast to that possibility, the current experiments found no mediating role of likeability.

In line with our theoretical framework of organizational moral typecasting, however, we found that people perceived big businesses as *less vulnerable* and *less moral* than small businesses, and that this perception mediated the greater willingness to cheat big businesses for personal gain. These findings provide a new understanding of how moral psychology interacts with organizational features when people make important decisions about lying or telling the truth. That is, people's moral compass is not only guided by the act itself and the incentives involved, but is also significantly influenced by the perceived characteristics of *whom* it affects.

Limitations and Future Research

While our findings have interesting implications for psychological theory, organizations, and moral decision-making, several limitations must be acknowledged. First, our interpretations are based on empirical findings from a series of online experiments conducted in a single cultural setting (USA), making cross-cultural generalizability a question for future research. Second, although we manipulated size signals using established metrics such as number of employees and branches (Wirtz & McColl-Kennedy, 2010), future research may benefit from other metrics such as revenue, profit margins, and market share. Third, although we found the same main result across different dishonesty measures and contexts, it would also be interesting to see if similar results would be found across other measures of dishonest behavior (e.g., insurance fraud) and different experimental paradigms.

Fourth, we do not claim that organizational typecasting is the *only* explanation of the business-size bias in moral concern, as there might be several other mechanisms than vulnerability and morality perception at play, such as anti-establishment sentiments, victim identifiability, and perceptions of harm dilution. Fifth, although we tested our hypothesis across several industries, future work may also benefit from real-world base rates, for example, by studying whether bigger brands of the same supermarket chain suffer greater retail theft than smaller brands do. Finally, it may also be of interest to examine if our observed bias in moral concern extends to prosocial behaviors: Are individuals less likely to engage in helpful actions toward large than small organizations? Such an extension can reveal if the business-size bias is limited to contexts of dishonesty, or if it reflects an even broader pattern of moral behavior toward organizations of different sizes.

For now, we conclude that the business-size bias in moral concern appears to be a robust empirical finding. Large organizations were perceived as less vulnerable and less moral than small organizations, and when people made hypothetical or real decisions, they were more likely to lie and steal for personal gain when interacting with a large business. Since the business-size bias occurred even in anonymous situations with equal rewards from lying and zero detection risk, the data suggests that people experience a lower *internal moral cost of lying* when the victim of their crime is a big rather than a small organization.

References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for Truth-Telling. *Econometrica*, *87*(4), 1115–1153. https://doi.org/10.3982/ECTA14673

Acemoglu, D., & Robinson, J. A. (2019). Rents and economic development: The perspective of Why Nations Fail. *Public Choice*, *181*(1–2), 13–28. https://doi.org/10.1007/s11127-019-00645-z

Acs, Z. J., & Audretsch, D. B. (1987). Innovation in large and small firms. *Economics Letters*, *23*(1), 109–112. https://doi.org/10.1016/0165-1765(87)90211-4

Ayal, S., Celse, J., & Hochman, G. (2019). Crafting messages to fight dishonesty: A field investigation of the effects of social norms and watching eye cues on fare evasion. *Organizational Behavior and Human Decision Processes*, S0749597818305004. https://doi.org/10.1016/j.obhdp.2019.10.003

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. https://doi.org/10.1111/j.1745-6916.2007.00051.x

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. In N. G. Fielding, A. Clarke, & R. Witt (Eds.), *The Economic Dimensions of Crime* (pp. 13–68). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-62853-7_2

Chandler, A. D. (1959). The Beginnings of "Big Business" in American Industry. *Business History Review*, *33*(1), 1–31. https://doi.org/10.2307/3111932

Cohn, A., Gesche, T., & Maréchal, M. A. (2022). Honesty in the Digital Age. *Management Science*, *68*(2), 827–845. https://doi.org/10.1287/mnsc.2021.3985

Dai, Z., Galeotti, F., & Villeval, M. C. (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Management Science*, *64*(3), 1081–1100. https://doi.org/10.1287/mnsc.2016.2616

Earp, B. D., McLoughlin, K. L., Monrad, J. T., Clark, M. S., & Crockett, M. J. (2021). How social relationships shape moral wrongness judgments. *Nature Communications*, *12*(1), 5776. https://doi.org/10.1038/s41467-021-26067-4

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785–791. https://doi.org/10.1038/nature02043

Fombelle, P. W., Voorhees, C. M., Jenkins, M. R., Sidaoui, K., Benoit, S., Gruber, T., Gustafsson, A., & Abosag, I. (2020). Customer deviance: A framework, prevention strategies, and opportunities for future research. *Journal of Business Research*, *116*, 387–400. https://doi.org/10.1016/j.jbusres.2019.09.012

Freund, A., Flynn, F., & O’Connor, K. (2023). Big Is Bad: Stereotypes About Organizational Size, Profit-Seeking, and Corporate Ethicality. *Personality and Social Psychology Bulletin*.

Fritz, M. S., & MacKinnon, D. P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychological Science*, *18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, *145*(1), 1–44. https://doi.org/10.1037/bul0000174

Gibson, R., Tanner, C., & Wagner, A. F. (2013). Preferences for Truthfulness: Heterogeneity among and within Individuals. *American Economic Review*, *103*(1), 532–548. https://doi.org/10.1257/aer.103.1.532

Giroux, M., Kim, J., Lee, J. C., & Park, J. (2022). Artificial Intelligence and Declined Guilt: Retailing Morality Comparison Between Human and AI. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-022-05056-7

Gray, K., Waytz, A., & Young, L. (2012). The Moral Dyad: A Fundamental Template Unifying Moral Judgment. *Psychological Inquiry*, *23*(2), 206–215. https://doi.org/10.1080/1047840X.2012.686247

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520. https://doi.org/10.1037/a0013748

Greenberg, J. (1993). Stealing in the Name of Justice: Informational and Interpersonal Moderators of Theft Reactions to Underpayment Inequity. *Organizational Behavior and Human Decision Processes*, *54*(1), 81–103. https://doi.org/10.1006/obhd.1993.1004

Greenberg, J. (2002). Who stole the money, and when? Individual and situational determinants of employee theft. *Organizational Behavior and Human Decision Processes*, *89*(1), 985–1003. https://doi.org/10.1016/S0749-5978(02)00039-0

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Second edition). Guilford Press.

Hertwig, R., & Mazar, N. (2022). Toward a taxonomy and review of honesty interventions. *Current Opinion in Psychology*, *47*, 101410. https://doi.org/10.1016/j.copsyc.2022.101410

Hobbes, T. (1651). *Leviathan: Or, The matter, forme, and power of a common-wealth ecclesiasticall and civil*. Wisehouse Classics.

Idson, T. L., & Oi, W. Y. (1999). Workers Are More Productive in Large Firms. *American Economic Review*, *89*(2), 104–108. https://doi.org/10.1257/aer.89.2.104

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, *93*, 328–336. https://doi.org/10.1016/j.jebo.2013.04.003

John, L. K., Loewenstein, G., & Rick, S. I. (2014). Cheating more for less: Upward social comparisons motivate the poorly compensated to cheat. *Organizational Behavior and Human Decision Processes*, *123*(2), 101–109. https://doi.org/10.1016/j.obhdp.2013.08.002

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *The American Economic Review*, *76*(4), 728–741. https://doi.org/10.1017/CBO9780511803475.019

Kajackaite, A., & Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior*, *102*, 433–444. https://doi.org/10.1016/j.geb.2017.01.015

Kocher, M. G., Schudy, S., & Spantig, L. (2018). I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. *Management Science, 64*(9), 3995–4008. https://doi.org/10.1287/mnsc.2017.2800

Lakens, D. (2017). Equivalence Tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

LaMothe, E., & Bobek, D. (2020). Are Individuals More Willing to Lie to a Computer or a Human? Evidence from a Tax Compliance Setting. *Journal of Business Ethics*, *167*(2), 157–180. https://doi.org/10.1007/s10551-019-04408-0

Lee, K., Kim, B.-Y., Park, Y.-Y., & Sanidas, E. (2013). Big businesses and economic growth: Identifying a binding constraint for growth with country panel analysis. *Journal of Comparative Economics*, *41*(2), 561–582. https://doi.org/10.1016/j.jce.2012.07.006

Mars, G. (1985). Cheats at Work: The Anthropology of Workplace Crime. *Anthropology of Work Review*, *6*(4), 44–47. https://doi.org/10.1525/awr.1985.6.4.44

McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What We Owe to Family: The Impact of Special Obligations on Moral Judgment. *Psychological Science*, *31*(3), 227–242. https://doi.org/10.1177/0956797619900321

National Retail Federation. (2022). *NRF Reports Retail Shrink Nearly a $100B Problem*. https://nrf.com/media-center/press-releases/nrf-reports-retail-shrink-nearly-100b-problem

Okereke, M. (2021). Towards vaccine equity: Should big pharma waive intellectual property rights for COVID-19 vaccines? *Public Health in Practice*, *2*, 100165. https://doi.org/10.1016/j.puhip.2021.100165

Pascual-Ezama, D., Prelec, D., Muñoz, A., & Gil-Gómez de Liaño, B. (2020). Cheaters, Liars, or Both? A New Classification of Dishonesty Profiles. *Psychological Science*, *31*(9), 1097–1106. https://doi.org/10.1177/0956797620929634

Pinker, S. (2012). *The better angels of our nature: Why violence has declined*. Penguin Books.

Qualtrics (Version 2022). (2022). [Computer software]. Qualtrics. https://www.qualtrics.com

R Core Team. (2013). *R* [Computer software]. R Foundation for Statistical Computing. www.R-project.org/

Reynolds, T., Howard, C., Sjåstad, H., Zhu, L., Okimoto, T. G., Baumeister, R. F., Aquino, K., & Kim, J. (2020). Man up and take it: Gender bias in moral typecasting.

Organizational Behavior and Human Decision Processes, *161*, 120–141. https://doi.org/10.1016/j.obhdp.2020.05.002

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

Rotman, J. D., Khamitov, M., & Connors, S. (2018). Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically. *Journal of Consumer Psychology*, *28*(2), 353–361. https://doi.org/10.1002/jcpy.1002

Saad, L. (2022). Do Americans Like or Dislike "Big Business"? *GALLUP*. https://news.gallup.com/poll/270296/americans-dislike-big-business.aspx

Schein, C., & Gray, K. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Septianto, F., & Kwon, J. (2021). Too cute to be bad? Cute brand logo reduces consumer punishment following brand transgressions. *International Journal of Research in Marketing*, S016781162100118X. https://doi.org/10.1016/j.ijresmar.2021.12.006

Shepherd, S., Kay, A. C., & Gray, K. (2019). Military veterans are morally typecast as agentic but unfeeling: Implications for veteran employment. *Organizational Behavior and Human Decision Processes*, *153*, 75–88. https://doi.org/10.1016/j.obhdp.2019.06.003

Skowronek, S. E. (2022). DENIAL: A Conceptual Framework to Improve Honesty Nudges. *Current Opinion in Psychology*, 101456. https://doi.org/10.1016/j.copsyc.2022.101456

Smigel, E. O. (1956). Public Attitudes Toward Stealing as Related to the Size of the Victim Organization. *American Sociological Review*, *21*(3), 320. https://doi.org/10.2307/2089287

Soraperra, I., Weisel, O., & Ploner, M. (2019). Is the victim *Max (Planck)* or *Moritz* ? How victim type and social value orientation affect dishonest behavior. *Journal of Behavioral Decision Making*, *32*(2), 168–178. https://doi.org/10.1002/bdm.2104

Strohminger, N., & Jordan, M. R. (2022). Corporate insecthood. *Cognition*, *224*, 105068. https://doi.org/10.1016/j.cognition.2022.105068

Sung, Y. H., Lim, R. E., & Lee, W.-N. (2022). Does company size matter in corporate social responsibility? An examination of the impact of company size and cause proximity fit on consumer response. *International Journal of Advertising*, *41*(2), 284–308. https://doi.org/10.1080/02650487.2020.1850997

The jamovi project. (2022). *Jamovi.* (2.3) [Computer software]. https://www.jamovi.org

Thiel, P. (2014). Competition Is for Losers. *The Wall Street Journal*. https://www.wsj.com/articles/peter-thiel-competition-is-for-losers-1410535536

Thompson, C. J., & Arsel, Z. (2004). The Starbucks Brandscape and Consumers' (Anticorporate) Experiences of Glocalization. *Journal of Consumer Research*, *31*(3), 631–642. https://doi.org/10.1086/425098

Vincent, L. C., Emich, K. J., & Goncalo, J. A. (2013). Stretching the Moral Gray Zone: Positive Affect, Moral Disengagement, and Dishonesty. *Psychological Science*, *24*(4), 595–599. https://doi.org/10.1177/0956797612458806

Wallach, K. A., & Popovich, D. (2023). When Big Is Less than Small: Why dominant brands lack authenticity in their sustainability initiatives. *Journal of Business Research*, *158*, 113694. https://doi.org/10.1016/j.jbusres.2023.113694

Wang, L., & Murnighan, J. K. (2017). How Much Does Honesty Cost? Small Bonuses Can Motivate Ethical Behavior. *Management Science*, *63*(9), 2903–2914. https://doi.org/10.1287/mnsc.2016.2480

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

Wirtz, J., & McColl-Kennedy, J. R. (2010). Opportunistic customer claiming during service recovery. *Journal of the Academy of Marketing Science*, *38*(5), 654–675. https://doi.org/10.1007/s11747-009-0177-6

Woolley, K., Kupor, D., & Liu, P. J. (2022). EXPRESS: Does Company Size Shape Product Quality Inferences? Larger Companies Make Better High-Tech Products, but Smaller Companies Make Better Low-Tech Products. *Journal of Marketing Research*, 002224372211248. https://doi.org/10.1177/00222437221124857

Yam, K. C., & Reynolds, S. J. (2016). The Effects of Victim Anonymity on Unethical Behavior. *Journal of Business Ethics*, *136*(1), 13–22. https://doi.org/10.1007/s10551-014-2367-5

Yang, L. W., & Aggarwal, P. (2019). No Small Matter: How Company Size Affects Consumer Expectations and Evaluations. *Journal of Consumer Research*, *45*(6), 1369–1384. https://doi.org/10.1093/jcr/ucy042

Zhang, Z., Mai, Y., Xu, Z., & McNamara, C. (2023). *Basic and Advanced Statistical Power Analysis* (0.9.4) [Computer software]. https://webpower.psychstat.org

Chapter 4. Article 3: Intergroup Bias in Dishonesty: Selfish versus Coalitional Lying

PREPRINT

Working paper part of a doctoral dissertation (March 2024). All conclusions are preliminary.

Please do not copy, share, or cite without the authors' permission.

Jareef Martuza[1], Hallgeir Sjåstad[1], Helge Thorbjørnsen[1,2]

[1]Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, Bergen 5045, Norway

[2] SNF, Center for Applied Research at NHH, Helleveien 30, Bergen 5045, Norway

Author notes

The authors declare to have no known conflicts of interest. Correspondence about this research should be addressed to either Jareef Martuza (jareef.martuza@nhh.no; PhD student at the Norwegian School of Economics) or Hallgeir Sjåstad (hallgeir.sjastad@nhh.no; Professor of Psychology and Leadership at the Norwegian School of Economics.

Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. As Norwegian laws and regulations do not require review by an institutional review board for anonymous, non-medical, low-risk research with human participants, this project was not submitted to such review. Informed consent was collected from all participants. We received partial funding for our studies from the Center of Ethics and Economics at NHH to conduct our experiments.

Abstract

As individual decisions naturally occur in a social context, the acceptability of dishonest behavior may depend on the *group identity* of whom it affects – posing a serious challenge for intergroup cooperation and societal polarization. To test this possibility, we used an adaptation of the "mind game" to provide anonymous decision-makers with an economic incentive to lie with zero detection risk, in which participants could double their earnings (or the earnings of someone else) by self-reporting a correct guess of a die-roll. In Experiment 1 (N=1,177), Democrat and Republican US voters could lie to increase their personal earnings at the cost of either their political ingroup (same voting preference) or their outgroup (opposite voting preference). Surprisingly, both to us and the participants themselves (prediction data), participants lied at the same rate regardless of whether their self-benefitting lie would come at the cost of an outgroup member or an ingroup member. In Experiment 2 (N=1,710), Democrat and Republican US voters could lie to benefit their political ingroup, their outgroup, or their ingroup *at the cost* of the outgroup – without any selfish reward. This time, participants lied at a significantly higher rate to *benefit* an ingroup member than to benefit an outgroup member (9 p.p.). In addition, we observed no significant lying to *harm* the earnings of an outgroup member, although underreporting correct guesses was also an option. These findings suggest that group identities may not influence self-interested dishonesty in the narrow sense (Exp. 1), but that it robustly promotes other-benefitting behavior in the form of *coalitional dishonesty* (Exp. 2), in which ingroup love, but not outgroup hate, reduces the moral cost of lying.

Keywords: intergroup bias; behavioral dishonesty; selfish lying; coalitional lying.

Intergroup Bias in Dishonesty: Selfish versus Coalitional Lying

In social interaction and intergroup relations, the field of social psychology is deeply concerned with understanding the human tendency to categorize the world into 'us' and 'them' (Brewer, 1999; Cikara et al., 2017; Perdue et al., 1990; Tajfel, 1979; Van Bavel & Packer, 2021). People who see themselves as members of a group behave differently than those who see themselves as isolated individuals (Charness et al., 2007), and being part of a broader 'we' can lead to superior group performance (Ellemers et al., 2004) and long-term resilience in challenging times (Van Bavel et al., 2020). Yet, reflecting the same basic needs for belonging and a positive social identity (Cárdenas & de la Sablonnière, 2020; Oldmeadow & Fiske, 2010), seeing oneself and another person as members of *different* groups can influence people to behave in less cooperative and sometimes even destructive ways (Castano et al., 2002). For instance, compared to members of the ingroup, outgroup members are allocated fewer resources (Tajfel et al., 1971), evoke less moral concerns (Pratto & Glasford, 2008; Struch & Schwartz, 1989), are more likely to be targets of spiteful behavior (Mill & Morgan, 2022), and are even at a higher risk of being sacrificed against their will (Watkins & Laham, 2019). Intergroup bias can also have detrimental consequences at the organizational and societal level, such as racial discrimination in hiring processes (Quillian et al., 2017), reciprocal dehumanization (Kteily et al., 2016), and segregated neighborhoods (Zou & Cheryan, 2022).

Although there is substantial research on the social drivers of cooperation and helping behavior (please see Balliet et al., 2014 for a meta-analysis), less is known about the group dimension of *honesty*: The basic willingness to tell the truth, and to do so even in the face of economic incentives to lie or misreport factual information. Both in academic philosophy and the legal system, the goal is usually to identify universal rights and wrongs, independent from the identities of perpetrators, victims, and beneficiaries. The psychology of right and wrong in everyday life, however, might work differently. As individual decisions naturally occur in a social context of relationships and mutual dependence (Earp et al., 2021), we propose that the moral acceptability of dishonest behavior may partly be a function of the *group identity* of whom it affects. Examining the possibility of an *intergroup bias* in dishonesty is important because the findings can provide a new understanding of how social context, specifically group identities, may shape individual moral decisions. Should the effect be robustly found, it would suggest that people operate with at least two sets of moral codes in social interactions- one for their ingroups and another for their outgroups. Should there *not* be an effect, it would suggest an important boundary condition of intergroup bias in morality.

In the current research, we conducted two incentivized pre-registered experiments (Ns = 1,176 and 1,710) using natural group identities from the political domain, recruiting Democrat and Republican-voting U.S. American participants. Given the current need for robust and credible research in the domain of honesty, and the general systemic problem of publication bias (Scheel et al., 2021; Sterling, 1959), we designed the experiments to be maximally informative in supporting or contradicting our hypotheses. We explore two primary research questions: (1) Are people more likely to cheat outgroups than ingroups for personal gain? (2) Under conditions of no personal gain, are people more likely to cheat when their ingroup than their outgroups receive the gain? To exclude the overt influence of ingroup cooperation and other self-interested motives that could favor the ingroup in repeated interaction with identifiable agents, such as reciprocity and reputation (Axelrod, 1986; Efferson et al., 2024; Romano et al., 2017; Trivers, 1971), we tested our hypotheses using one-shot decisions in *anonymous* experiments with no detection risk, creating a more direct *trade-off* between economic gains and the moral motivation to tell the truth. As a result, differences in behavioral dishonesty would suggest that the "pure" moral cost of lying may be different when it costs and/or benefits ingroups versus outgroups.

1. Intergroup Bias in Dishonesty for Personal Gain

It is well-established in psychology and behavioral economics that human decision-makers, contrary to predictions from the rational model (Becker, 1968), typically try to strike a balance between pursuing economic self-interests and maintaining internal moral standards (Abeler et al., 2019; Gibson et al., 2013). Even in the absence of both reputational concerns and risk of punishment, maintaining one's moral standards can restrict dishonest pursuits of economic self-interest (Gneezy et al., 2018; Thielmann & Hilbig, 2019). While previous research on dishonesty has examined the role of economic incentives (Wang & Murnighan, 2017; Wiltermuth, 2011), the choice environment (Ayal et al., 2019; Cohn et al., 2022; Kocher et al., 2018), and dispositions of the decision-maker (Pascual-Ezama et al., 2015; Vincent et al., 2013), the role of *who* the potential victim is, is underexamined. This is critical because the characteristics of the victim may play a crucial role in regulating dishonest behavior for personal gain (Köbis et al., 2019). Indeed, a small stream of existing research shows that people are more likely to cheat and an identifiable than an anonymous victim (Yam & Reynolds, 2016), a company that commits a moral transgression than one that does not (Rotman et al., 2018), and an organization than a person (Soraperra et al., 2019) – highlighting the importance of considering victim characteristics in studying dishonesty.

Although recent research at the intersection of dishonesty and intergroup relations has examined collaborative dishonesty (van Lent et al., 2023) and contagion (Vives et al., 2022), to our knowledge, no research has examined whether people's willingness to lie for personal gain, also referred to as cheating (Jiang, 2013), is *different* across ingroup and outgroup members as victims. Our question of whether there is an intergroup bias in cheating is also different from seminal work on how people allocate resources between ingroups and outgroups as third-party decision-makers (Tajfel et al., 1971) because allocating resources between ingroups and outgroups relates to biased fairness preferences. Our question is also different from contemporary research on allocating resources between the self and an ingroup or outgroup member (Dimant, 2023) and helping/harming ingroups versus outgroups at personal cost (Halevy et al., 2008)- because these reflect prosocial preferences, without having components of truth-telling or lying. Specifically, the decision-maker is not faced with a trade-off between material gains and the moral motivation to "not cheat" when allocating resources between two others (Tajfel et al., 1971), between the self and another person (Dimant, 2023), and helping/harming another person at personal cost (Halevy et al., 2008). As such, the current research builds and extends on a wide range of previous work on intergroup relations and decision-making, but makes a different contribution by focusing on the group psychology of *behavioral dishonesty*.

Our first hypothesis predicting an intergroup bias in dishonesty for personal gain is supported by a broad range of theoretical perspectives and related empirical evidence. In particular, social identity theory (Tajfel et al., 1971) suggests a joint process of self-categorization and a motive for positive distinction (Turner & Reynolds, 2012) as primary drivers of intergroup bias, which may produce different moral standards for acceptable behavior toward ingroups and outgroups. For instance, social identities can create a special bond of loyalty within the ingroup (Graham et al., 2013), in which people become hypersensitive to any self-serving action that can be perceived as a violation of mutual trust and fairness. In betting behavior, research on "disloyalty aversion" (Morewedge et al., 2018) has found that many people refuse to profit from the demise of their own sports team, even when being offered a free bet that has no impact on the actual outcome of the game. Specifically, two experiments found that more than 45% of basketball fans chose *not* to take the “free" gamble of winning $5 if their team lost and getting $0 if their team won (Morewedge et al., 2018), although there was no downside to taking the bet in purely economic terms. In our view, this aversion to collecting selfish gain from a losing intergroup competition suggests that the valuation of one’s social identity may also outweigh the preference for selfish monetary gains in the domain of honesty.

Recent research also shows that people are more likely to help than harm (Dimant, 2023), readily perceive the pain (Mende-Siedlecki et al., 2019), and avoid destroying the wealth of ingroups than outgroups (Mill & Morgan, 2022). Of course, stronger reciprocity and reputational motivation are likely to be important explanations for such examples of intergroup bias in social behavior. What we suggest in addition, is that even when removing the direct importance of self-serving social motives, people may *still* impose higher moral standards on themselves when interacting with ingroup members than they would with outgroup members, which might create a systematic difference in the willingness to lie for personal gain. If true, that would suggest that even when there are no strategic social motives involved, people operate with at least two different sets of moral codes: one moral code for the ingroup, and a different one for the outgroup.

Combining theoretical predictions and suggestive evidence from related domains, we formally hypothesized the following:

H1: Individuals will be more likely to engage in dishonest behavior toward members of an outgroup compared to members of an ingroup, reflecting an intergroup bias in dishonesty for personal gain.

2. Intergroup Bias in Coalitional Dishonesty

Whereas dishonest behavior for personal gain involves a trade-off between economic incentives and moral standards, people often face a dilemma when their dishonesty benefits *another* person, without direct personal benefits. In contrast with lying for personal gain, lying to benefit someone else in a group context can involve making a trade-off between *conflicting moral standards*: Telling the truth, regardless of who will financially gain or lose from it, or, telling a lie if it can benefit one's ingroup more than the outgroup. Because we are interested in the group psychology of lying, our research moves away from the more general altruistic lying (Brocas & Carrillo, 2021), to examining how group identities specifically influence other-benefitting lies- what we conceptualize as the notion of an intergroup bias in *coalitional dishonesty*. Further, because we eliminate selfish incentives from the equation, this investigation is also different from "altruistic white lies"- where lying benefits another person at a cost to the self (Erat & Gneezy, 2012). Under conditions of no personal gain, we ask if people are more willing to lie to benefit their ingroup than their outgroup. If so, parallel to how the moral cost of cheating an outgroup (vs. ingroup) member for personal gain might be lower, the moral cost of lying to *benefit* an ingroup (vs. outgroup) member might also be lower.

Adjacent research shows that people are more willing to break rules (Brass et al., 1998) and behave dishonestly (Wiltermuth, 2011) when they have an existing relationship with the beneficiary. People are also more likely to cheat under team than individual incentives (Conrads et al., 2013), and even when the benefits are split with another anonymous stranger than when benefits are personal only (Wiltermuth, 2011; for a meta-analysis on collaborative dishonesty, please see Leib et al., 2021). Turning to the current research, we propose that this tendency to lie to benefit others may also be influenced by whether the *beneficiary* is an ingroup or an outgroup member. An anecdote supporting this notion comes from the world of sports: Comparing judges' scores from the 2002 Olympics and other major international competitions showed that figure skating and ski jumping judges scored their national compatriots about 0.13 standard deviations higher than judges with different nationalities (Zitzewitz, 2006).

Our second hypothesis of an intergroup bias in coalitional dishonesty is also informed by multiple theoretical perspectives and related empirical evidence. Moral foundations theory proposes "ingroup loyalty" as one of the five moral foundations explaining human moral thinking at the innate-intuitive level (Graham et al., 2013). It has also been suggested that morality itself has evolved to promote cooperation and pro-sociality within the tribe and other narrowly defined ingroups (Curry, 2016; Curry et al., 2019; Greene & Haidt, 2002). Indeed, this "functionalist approach" to morality can enable people to act *against* internal moral standards when it serves group interests more than individual self-interest (Shalvi & De Dreu, 2014). Whereas cheating an ingroup member for personal gain would be seen as a severe violation of norms, lying to benefit an ingroup member can be seen as an expression of altruism or loyalty toward the ingroup, as suggested by evolutionary perspectives on coalitional psychology (Cikara, 2021; Cosmides & Tooby, 2013; Tooby & Cosmides, 2007). Indeed, altruistic concerns, such as increasing proceeds going to a favorable charity, have been shown to increase levels of dishonest behavior among experimental participants (Hochman et al., 2021).

Related findings from experimental economics also offer suggestive support for our hypothesis of an intergroup bias in coalitional dishonesty. In a lab-in-the-field experiment conducted at a traditional Italian music festival, where self-reported false information of participants could increase payoffs received by an ingroup or an outgroup (at the cost of the experimenters), neither Northern nor Southern Italians were willing to lie to benefit their outgroups; but interestingly, Southerners (not Northerners) lied significantly more to benefit an ingroup (Michailidou & Rotondi, 2019). Another two-period lab-in-the-field experiment with coffee farmers in Guatemala found that participants were likely to cheat to increase payoffs of another fellow villager (vs. stranger from another village) under conditions of abundance (during the harvest season). However, this difference was non-significant under conditions of scarcity (before the harvesting season) (Aksoy & Palma, 2019). In a controlled lab experiment with Chinese university students (Cadsby et al., 2016), participants cheated not only to benefit themselves, but also to benefit an ingroup (fellow student from the same university) at the cost of an outgroup (student from a different university).

Although the previously mentioned studies offer important empirical support to inform our hypothesis, small sample sizes (N = 248 in Michailidou & Rotondi, 2019; N = 109 in Aksoy & Palma, 2019; N = 360 in Cadsby et al., 2016) compared to current standards and significant effects only observed at specific levels of a covariate (effect present only among Southern but not Northern Italians in Michailidou & Rotondi, 2019; during harvesting season but not before in Aksoy & Palma, 2019) necessitates rigorous testing of the general effect. Further, even though Cadsby and colleagues (2016) found a significant and overall intergroup effect in lying to benefit others, their study primarily examined lying to benefit the ingroup at the cost of the outgroup. So, it remains unclear if individuals may also be more likely to lie to benefit their ingroup than the outgroup, even when there is no direct competition and the decision of supporting one’s ingroup or not does not affect any outgroup.

Further, with respect to the findings of Cadsby and colleagues (2016), any inference about the motivation to benefit the ingroup versus harm the outgroup would be conflated. Given that both “ingroup love” and “outgroup hate” have been shown to distinctly influence consequential judgments and decisions (Abbink & Harris, 2019; Dimant, 2023; Halevy et al., 2008; Yudkin et al., 2016), it would be both theoretically and empirically informative to examine how intergroup bias in other-benefitting dishonesty may be a function of either ingroup love, outgroup hate, or both. In the current research, we designed the current experiments accordingly to test these possibilities.

Although our main interest lies in examining whether people are more likely to lie to benefit their ingroups than they would outgroups, it is possible that people may also lie to actively *harm* their outgroups. Intergroup bias has been shown to elicit counter-empathic responses such as feeling pleasure in response to the pain (Schadenfreude) and pain in response to the pleasure of outgroups, suggesting the presence of “harm for harm’s sake” (Cikara et al., 2011) and even antipathy (Cikara et al., 2014) toward outgroups. These findings may suggest that the moral costs of lying may not only be lower when lying to benefit an ingroup but might also be lower when lying to harm an outgroup.

So, going beyond much of existing research, we compare lying to simply benefit an ingroup vs. an outgroup in addition to lying to benefit an ingroup at the cost of an outgroup, in the same experiment. Our design also enables us to test if individuals significantly lie on average to *reduce* the earnings of an outgroup member- that is, whether people might also lie to express outgroup hate. Together, our design and methods can shed light on how ingroup love versus outgroup might distinctly underlie intergroup bias in coalitional dishonesty.

Based on different theoretical predictions and existing empirical evidence, we formally hypothesized the following:

H2: Individuals will be more likely to engage in dishonest behavior to benefit members of an ingroup compared to members of an outgroup, reflecting an intergroup bias in dishonesty for others.

3. The Current Research

We sought to examine two main propositions: (1) intergroup bias in dishonesty for personal gain, and (2) intergroup bias in dishonesty to benefit others *without* any personal gain. To test our hypotheses, the current research moves beyond the design of most previous studies on intergroup dynamics in two significant ways. First, the current research goes beyond self-reports (Baumeister et al., 2007), and offers *behavioral* evidence from decisions with real consequences. In both experiments, we used a modified version of a well-established incentive-aligned guessing task from behavioral economics, known as the *mind-game paradigm* (Fischbacher & Föllmi-Heusi, 2013; Jiang, 2013) to tempt participants to lie under conditions of zero detection, punishment, or reputational concerns, while holding incentivizes constant. We recruited adult U.S. Americans who identify as either Democrats or Republicans in both experiments.

In Experiment 1, participants guessed in advance whether the score of a future die roll would be odd or even, and self-reported after observing the score if their guess was correct or not. If their guess was reported as correct, the participant received $0.75 as a bonus. If their guess was reported as incorrect, their paired participant- a political ingroup or an outgroup- received the $0.75 bonus. In Experiment 2, participants completed the same die roll task with one key difference: Their decisions would not affect their personal gains. Here, if they reported their guess as correct, a political ingroup or an outgroup received a $0.75 bonus. If they reported their guess as incorrect, either there was no bonus, or their political outgroup received the $0.75 bonus.

Beyond testing the main effect, we also pre-registered exploration of how intergroup bias in dishonesty may be moderated by levels of ingroup identification (Leach et al., 2008), sectarianism (Landry et al., 2024), and affective polarization. Recent research suggests that intergroup effects are more likely to exist among the more "groupy" than "non-groupy" individuals (Kranton et al., 2020). So, we chose to include two established scales: One that captures the tendency for ingroup love (identification with the ingroup by Leach and colleagues) and another that captures outgroup hate (sectarianism- a combination of othering, aversion, and negative moralization- by Landry and colleagues). We expect that the intergroup bias in dishonesty would be stronger for individuals reporting higher levels on each scale.

Because our design allowed us to infer the actual rates of dishonesty among Democrats and Republicans across conditions, we also explored *beliefs* about the dishonesty of ingroups and outgroups. Specifically, we measured participants' *estimated* rate of dishonesty in two ways: (1) by asking participants to estimate the rate of dishonesty among their political ingroups and outgroups, and (2) by asking participants to estimate the rate of dishonesty among their political ingroups and outgroups toward outgroups a more *specific* belief measure of intergroup dishonesty. We also pre-registered our overall prediction: Individuals will overestimate the likelihood of dishonest behavior among political outgroup members more than they would overestimate among political ingroup members. That is, although there may be a general tendency to overestimate the dishonesty of others, the difference between predicted and actual rates of dishonesty will be larger when considering outgroup behavior than ingroup behavior.

Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. Given that **country redacted for anonymous peer review** laws and regulations do not require a review by an institutional review board for anonymous, non-medical, low-risk research involving human participants, we did not submit the project for such a review.

The experiments were implemented in Qualtrics (Qualtrics, 2022), and the analyses were conducted in the latest version of R (R Core Team, 2013) and jamovi (The jamovi project, 2022). Visualizations were made using the ggplot2 (Wickham, 2016), and regression outputs using sjplot (Lüdecke, 2023).

The main hypotheses, planned analyses, sample sizes, and exclusion criteria were all pre-registered before data collection. We report all measures, manipulations, and conditions in the methods sections. Any deviations from pre-registrations are marked explicitly. Non-preregistered analyses are flagged as secondary analyses.

4. Experiment 1

In Experiment 1 (preregistered: open PDF here), we examined how intergroup bias affects dishonest behavior for personal gain using *natural groups* in the political domain. Using Prolific, an online participant recruitment platform known to provide high-quality data (Peer et al., 2022), we recruited U.S. Americans who identified as either Democrats or Republicans. To manipulate group identities, we used a between-participants design and paired each participant with another participant from the same party (ingroup condition) *or* a different party (outgroup condition). As the outcome measure, participants had the opportunity to dishonestly increase their earnings at the cost of their paired participant with an ingroup or outgroup signal, without any surveillance, reputation, or punishment concerns. To measure behavioral dishonesty, we used an adaptation of the mind-game task, as it creates a clear trade-off between financial gain and moral standards, and tests how the latter can vary based on the group identity (ingroup or outgroup) of the potential victim.

4.1. Method

4.1.1. Participants and power

We received complete responses from 595 Democrat and 601 Republican-identifying U.S. American participants. We recruited participants using filters of U.S.A nationality, reported political party affiliation (Democrat or Republican), being fluent in English, having an approval rate of 95% and higher in 100 or more submissions, and a 1:1 sex balance. Participants received a base payment of \$0.75 with the possibility of another \$0.75 as a bonus depending on the outcome of the task.

To achieve balanced pairings and cell sizes, we implemented two recruitment arms in Prolific based on participants' stated party membership, and randomly assigned them to an ingroup or an outgroup condition. In one arm, we recruited participants who identified themselves as Democrats. In another arm, we recruited participants who identified as Republicans. By recruiting participants through two simultaneous arms and randomly assigning them, we aim to achieve an even distribution of participants across the ingroup and outgroup conditions in a 1:1 ratio.

As per our pre-registration, we excluded responses from participants ($N = 19$) who failed the simple attention check at the start of the study. In the attention check question, participants were presented with general instructions about the importance of reading instructions carefully and not rushing through the survey, before being asked to select “A lot” to the question “How much do you like painting?” when presented with five options as potential answers. This left us with a final dataset consisting of 1,177 valid responses ($M_{age} = 43.84$, $SD = 14.12$; 565 male, 606 female, 14 other). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample could detect an effect size of $w = .09$ or larger in a “goodness-of-fit” contingency tables Chi-square test ($df = 1$, $p < .05$) with 90% power.

4.1.2. Procedure

The experiment consisted of four main parts: (1) participant pairing, (2) honesty task, (3) individual measures, and (4) belief measures.

Participant pairing. Right after the attention check question, participants responded to demographic measures such as age, gender, education level, household income, political party identification, voting history in the 2020 U.S. Presidential election, and voting intention in the 2024 election.

Then, participants were informed that they had been paired with another participant described as a Democratic or Republican Party voter, had voted for Joe Biden/ Donald Trump in the 2020 U.S. Presidential election, and plans to vote for the same candidate in the 2024 election. To make identities more salient, participants were presented with the symbol (donkey or elephant) of the corresponding party. Participants in both conditions then confirmed that they understood their paired participant's political leaning, and that their decisions in the next task can impact both their own and their pair's earnings. Please see section *1.2. Experimental Manipulations* in Supplemental Materials for a participant perspective preview across conditions.

Honesty task. For our outcome measure of behavioral dishonesty, participants in both conditions completed the mind-game paradigm. Participants were presented with instructions for a die-roll game where they were to guess in their minds in advance whether the outcome of a future die-roll would be odd or even. They were asked to remember their guess, and upon observing the result, report honestly whether their guess turned out to be correct or not. In the bonus mechanism part, they were informed that if they guessed correctly, they would receive a $0.75 bonus, thereby doubling their payment to $1.50, while their pair would not receive any bonus. If they guessed incorrectly, their paired participant would receive the same

$0.75 bonus, thereby doubling the paired participant's payment to $1.50, while the decision-making participant would not receive any bonus.

After confirming they understood the instructions, participants were directed to the game pages. First, after guessing in their minds, participants confirmed that they had complied with the instruction and actively chose "I have now made up my mind and guessed whether the score of the die roll will be ODD or EVEN" before proceeding. On the next page, they were prompted to click "ROLL" to reveal the number on the virtual die, and subsequently shown the score of the roll on the following page.

On the page where the score was revealed, participants were asked to confirm the number displayed on the displayed die. Then, participants were asked to self-report whether they guessed in advance correctly concerning whether the number would be odd (1, 3, 5) or even (2, 4, 6). As responses, participants were presented with two options: (1) they guessed correctly and would receive a $0.75 bonus while their pair would not receive any bonus (2) they did not guess correctly and their paired participant (a Democrat or Republican Party voter) would receive a $0.75 bonus while the decision-making participant would not receive any bonus.

Individual measures. After completing the honesty task, participants completed several individual-level measures. First, they completed a 14-item identification with the ingroup scale (Leach et al., 2008; a = .973) and a 9-item sectarianism scale (Landry et al., 2024; a = .944) on separate pages. Participants were presented with corresponding items where they rated their agreement from 1 (strongly disagree) to 7 (strongly agree) to each. The items for both scales were presented in a way that those related directly to Democrats and Republicans, depending on the party affiliation of the participant (Example items measuring ingroup identification among Republican participants: "I feel a bond with Republicans"/ "It is pleasant to be a Republican", example item measuring sectarianism among Democrat participants: "I feel distant from Republicans"/ "I hate Democrats"). After responding to these scales, participants reported their affective polarization, rating their feelings toward the Republican Party and the Democratic Party from 0 (Very cold) to 100 (Very warm). The measure of feelings toward the opposing party was reverse-coded as we are interested in the magnitude of polarization.

Belief measures. After the individual-level measures, participants were first asked what percentage of Democratic/ Republican Party voters (same as the *paired* participant's affiliation), completing the same study as them, would falsely state they guessed correctly in the die-roll game. To incentivize belief elicitation, participants were told that five participants with the closest guess would receive a $1 additional bonus. In a similar format, participants were also asked what percentage of Democratic/Republican Party voters (same as the paired participant's affiliation), *paired* with a Republican/Democratic Party voter (*opposite* as the paired participant's affiliation), would falsely state that they guessed correctly. This question probed what percentage of their ingroup or outgroup participants believed would engage in dishonest behavior specifically against members of the *opposing* political party. The idea here was to explore beliefs directly about intergroup dishonesty. Response to this question was also incentivized with a potential $1 bonus for the five closest guesses.

At the very end of the study, participants were asked to indicate their extent of agreement to three single-item measures exploring (lack of anticipated guilt ("If I were to cheat a Democratic/ Republican Party voter, I would <u>not</u> feel guilty about it"), cheating expectation ("My paired Democratic/ Republican party voter would expect me to cheat them"), and expected reciprocity ("My paired Democratic/ Republican Party voter would have cheated me if I were on the receiving end") on 7-point scales anchored at 1 (Strongly disagree) to 7 (Strongly agree). Please see Table S1 in Supplemental Materials for a full list of measures in Experiment 1.

4.2 Results

4.2.1. Preregistered analyses

Behavioral dishonesty. We pre-registered the hypothesis that "Individuals will be more likely to engage in dishonest behavior toward members of an outgroup compared to members of an ingroup, reflecting an intergroup bias in dishonesty."

In line with our preregistration, we first estimated the number of dishonestly behaving participants in each condition. Please see Table 1 for exact calculations derived for comparison.

Table 1. The estimated number of participants behaving dishonestly vs. honestly after observing a die-roll with an outcome opposite to what they had guessed in advance.

Condition/ Count	N	Expected correct, *E* (N/2)	Honesty Decisions, *D*	Reported correct guesses, *R*	Estimated Dishonest, ED (*R* - *E*)	Estimated Honest (*D* – *ED*)	Dishonest Participants (*ED* / *D*)
Ingroup	622	311	311	458	147	164	47.3%
Outgroup	555	278	277	400	122	155	44.0%

Descriptively, the estimated percentage of dishonest participants in the *ingroup* condition is higher, thereby going against our predicted direction. A chi-squared test with Yate's continuity correction rendered a non-significant difference between percentage dishonesty in the outgroup (44.0%) versus ingroup (47.3%) conditions, χ^2 (1, 588) = 0.55, p = .460. Thus, there is no statistical support for our preregistered hypothesis H1. The presence of any intergroup bias in dishonest behavior was <u>not</u> found to be statistically significant in either the predicted direction or the opposite direction.

Dishonesty beliefs. We pre-registered the hypothesis that "Individuals will overestimate the likelihood of dishonest behavior among outgroup members more than among ingroup members. Specifically, the difference between individuals' predicted rates of dishonesty and the actual rates will be greater for outgroup members compared to ingroup members."

In line with our preregistration, we first calculated the difference between predicted rates of dishonesty and actual rates of dishonesty for each cell. An independent samples t-test to compare the differences showed that the difference in the outgroup condition (M = 10.39, SD = 25.40) was significantly higher than that in the ingroup condition (M = 1.11, SD = 23.49), t (1176) = 6.65, p < .001, d = .388, 95% CI [.273, .504]. Thus, there is statistically significant support for our preregistered hypothesis H2. That is, participants overestimated the percentage of their outgroups behaving dishonesty significantly more than that of ingroups.

4.2.2. Secondary analyses

Behavioral dishonesty. Given the null effect of experimental condition on behavioral dishonesty, we conducted an equivalence test (Lakens et al., 2018) of proportions to examine if the proportions of dishonest participants in the outgroup (44.4%) and ingroup (47.1%) conditions are statistically "equivalent". This was not preregistered and should be seen as exploratory.

Contrary to how traditional hypothesis testing looks for significant differences, an equivalence test examines if the difference between two conditions is smaller than a pre-specified smallest effect size of interest (SESOI). Here, the null hypothesis posits that the true difference between the conditions is greater in magnitude (that is either above the upper bound or below the lower bound). The alternative hypothesis posits that the true differences lie within the bound, hence suggesting equivalence between the groups, and so statistical equivalence.

For our equivalence test of the two proportions of dishonest participants, we set the equivalence bounds at .05, assuming that a 5% point difference in the proportions of dishonest participants may be the smallest effect size of interest (SESOI). The results showed that the estimated 90% equivalence bounds [-.035, .10] were non-significantly different from the TOST lower ($Z = -0.432$, $p = .332$) but significantly different from the TOST upper ($Z = 2.00$, $p = .023$). This suggests that although the SESOI (5% point difference) cannot be ruled out in the direction *opposite* to our hypothesis, there is statistical equivalence in the direction of our hypotheses (proportion outgroup > proportion ingroup). Therefore, this provides evidence in favor of the null hypothesis concerning our predicted *direction* of intergroup bias in dishonesty. That is, contradicting our Hypothesis 1, individuals are no more likely to cheat outgroups than ingroups for personal gain.

One question we wondered was whether our treatment effect, which was null overall, may exist for some individuals but not others. So, we conducted a series of binary logistic regressions with reported guesses as the outcome (0 = incorrect, 1 = correct) and predictors of interactions of moderators (preregistered under secondary analysis) with the experimental condition (0 = ingroup, 1 = outgroup). Please see section 1.3 in Supplemental Materials for outputs from all models.

Notably among the moderators, the coefficient for the condition X ingroup identification scale interaction was significant (OR = 1.27, 95% CI [1.05, 1.53], p = .013). Follow-up floodlight analyses (Spiller et al., 2013) of the Johnson-Neyman test showed that the ingroup vs. outgroup effect was significant when ingroup identification was *outside* the interval [3.38, 7.59]. Interestingly, simple slopes analysis showed that along the ingroup identification scale, the slope of the experimental condition (0 = ingroup, 1 = outgroup) at -1 SD (3.25) was significant and *negative* (b = -.38, SE = .18, z = -2.05, p = .04), at the mean (4.64) non-significant (b = -.05, SE = .13, z = -.35, p = .72), and at + 1 SD (6.03) non-significant (b = .28, SE = .19, z = 1.48, p = .14). Further analysis showed that the number of participants in the region of significant was 237 (20.1% of the sample).

Together, this presents weak evidence regarding the presence of intergroup bias in dishonesty for personal gain in the direction *opposite* to our prediction at *low* levels of ingroup identification. That is, individuals with low ingroup identification may be more likely to cheat their *ingroups* than their outgroups for personal gain.

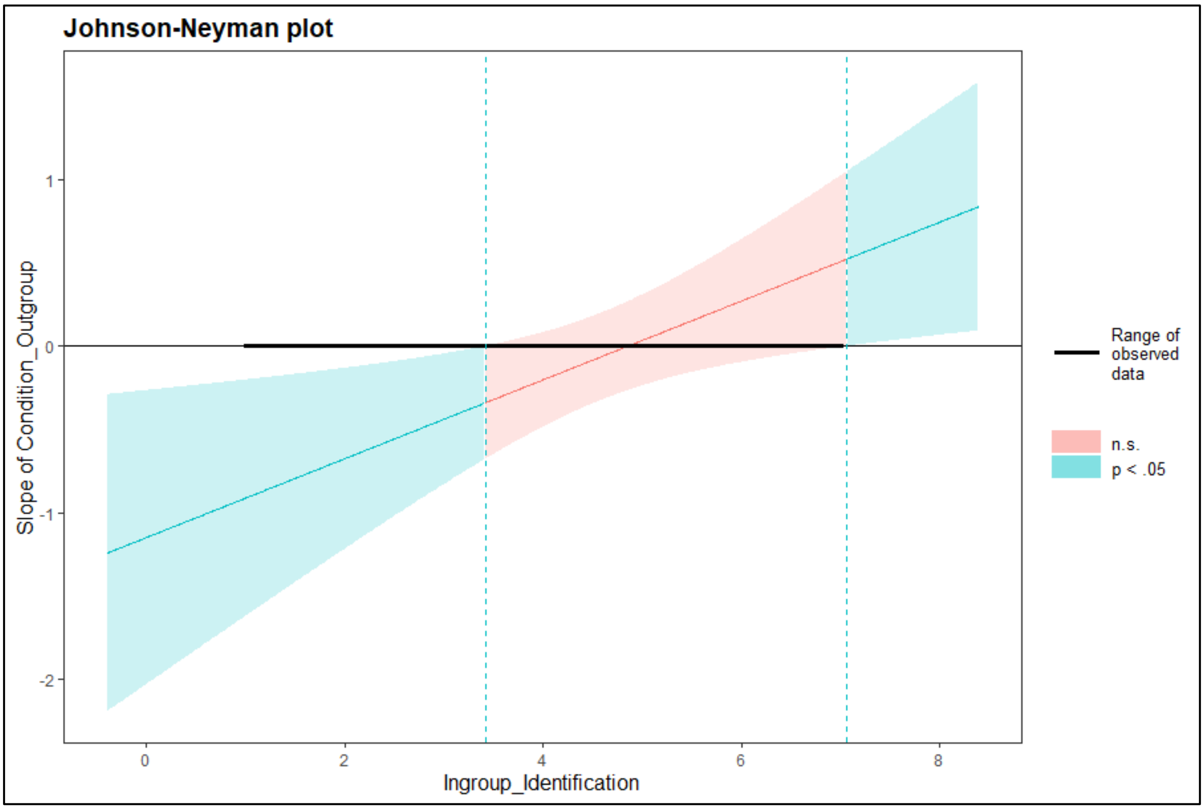


Figure 1. Flood analysis of intergroup bias in dishonesty for personal gain moderated by identification with the ingroup. Experiment 1 (N = 1,177). The figure illustrates the interaction effect (OR = 1.27, 95% CI [1.05, 1.53], p = .013) between experimental condition (cheating outgroup vs. ingroup) and intergroup identification as a continuous scale (1-7), using reported guess (0 = incorrect, 1 = correct) as the outcome. The estimated "region of significance" of the treatment effect is *outside* the interval [3.38, 7.59] on the ingroup identification scale. This suggests speculative evidence of the presence of an effect opposite to our predictions for individuals lower in ingroup identification (237 participants, 20.1% of the sample).

Dishonesty beliefs. We also conducted a series of multiple linear regressions with the difference between predicted and actual rates of dishonesty behavior as the outcome and predictors of interactions of moderators (preregistered as secondary) with the experimental condition (0 = ingroup, 1 = outgroup). Notably, we found that the intergroup bias in beliefs about others' dishonesty was higher among Democrats, and no significant differences among Republicans.

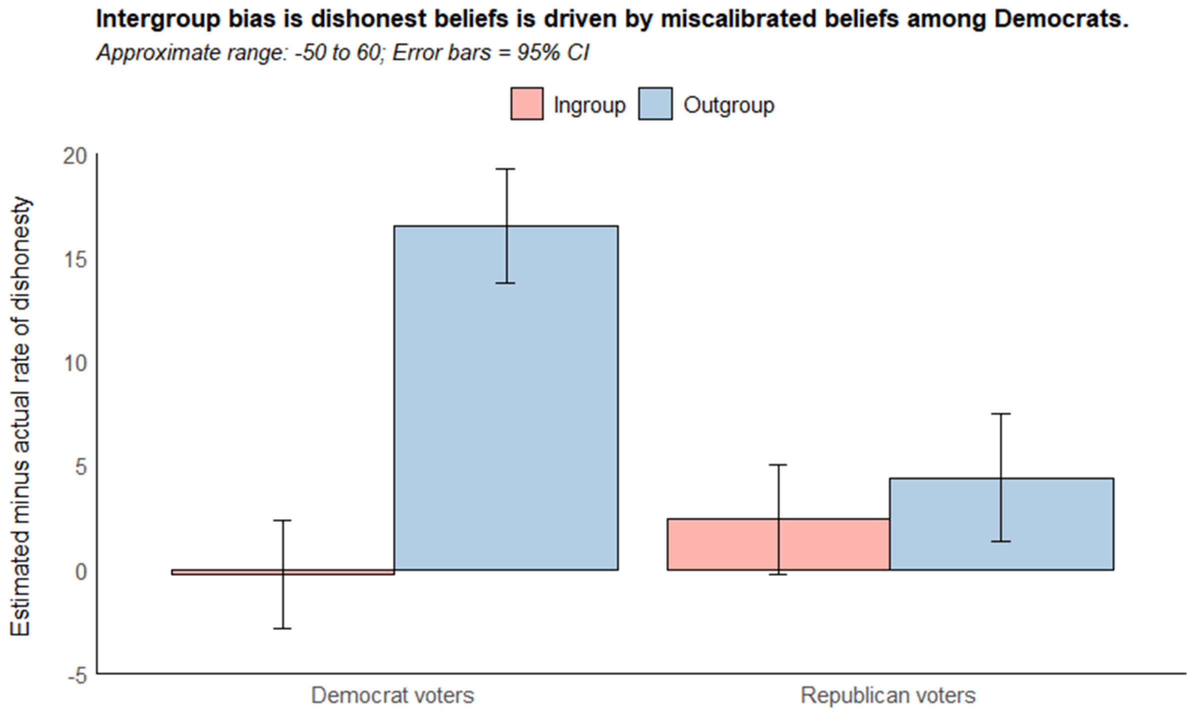


Figure 2. Estimated minus the actual rate of dishonest behavior across experimental conditions and political affiliations. Experiment 1 (N = 1,177). Democrats exhibited greater overestimation in the outgroup (M = 16.38, SD = 23.14) than ingroup (M = -0.51, SD = 23.19) condition, t (1173) = 8.48, p < .001, d = .702, whereas that in Republicans between outgroup (M = 4.11, SD = 26.12) and ingroup (M = 2.32, SD = 23.75) conditions was *not* statistically significant, t (1173) = -0.91, p = .802, d = .075.

A 2 (affiliation: Democratic vs. Republican) X 2 (condition: ingroup vs. outgroup) ANOVA showed that the condition X affiliation interaction term was significant, (F (1173) = 27.38.17, p < .001, $\eta^2{}_p$ = .023. Post hoc tests showed that Democrats exhibited greater overestimation in the outgroup (M = 16.48, SD = 23.14) than ingroup (M = -0.21, SD = 23.19) condition, t (1173) = 8.38, p_tukey < .001, d = .694, 95% CI [.529, .859]. However, the overestimation among Republicans between outgroup (M = 4.41, SD = 26.12) and ingroup (M = 2.42, SD = 23.75) conditions was *not* statistically significant, t (1173) = 1.01, p_tukey = .746, d = .783, 95% CI [-.079, .244]. This suggests that intergroup bias in dishonest beliefs may be asymmetrical across political ideologies, with Democrats especially believing Republicans to be more dishonest than reality, than Republicans not exhibiting an intergroup bias in dishonest beliefs.

Intergroup dishonesty beliefs. Our additional belief measure asked participants to estimate what percentage of their ingroup or outgroup members would engage in dishonest behavior when paired with another participant of the opposing party, that is, *specifically* estimate the rate of intergroup dishonesty. We first calculated the difference between estimated rates and actual rates in each cell. A 2 (affiliation: Democratic vs. Republican) X 2 (condition: ingroup vs. outgroup) showed that only the condition X affiliation interaction term was significant, F (1173) = 18.17, p < .001, $\eta^2{}_p$ = .015.

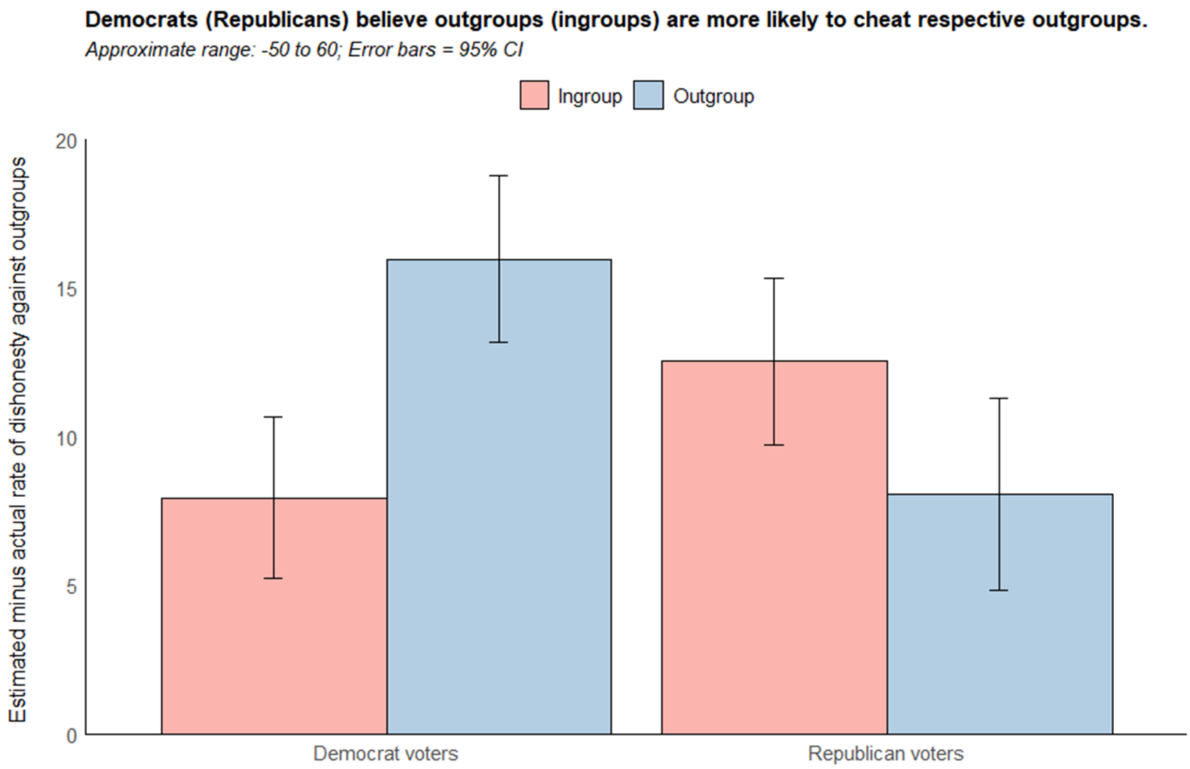


Figure 3. Estimated minus the actual rate of dishonest behavior in outgroup contexts across

experimental conditions and political affiliations. Experiment 1 (N = 1,177). Democrats significantly overestimated Republicans cheating Democrats (M = 15.97, SD = 23.72) more than they overestimated Democrats cheating Republicans (M = 7.95, SD = 24.24), t (1173) = 3.85, p_tukey = .001, d = .319. However, Republicans non-significantly overestimated Republicans cheating Democrats (M = 12.54, SD = 25.09) more than they overestimated Democrats cheating Republicans (M = 8.06, SD = 27.32), t (1173) = 2.17, p_tukey = .132, d = .179.

Post hoc comparisons showed that Democrats significantly overestimated Republicans cheating Democrats (M = 15.97, SD = 23.72) more than they overestimated Democrats cheating Republicans (M = 7.95, SD = 24.24), t (1173) = 3.85, p_tukey = .001, d = .319, 95% CI [.156, .482]. However, Republicans non-significantly overestimated Republicans cheating Democrats (M = 12.54, SD = 25.09) more than they overestimated Democrats cheating Republicans (M = 8.06, SD = 27.32), t (1173) = 2.17, p_tukey = .132, d = .179, 95% CI [.017, .340]. This suggests that both Democrats and Republicans believe Republicans are more likely to exhibit bias against outgroups than Democrats.

5. Experiment 2

In Experiment 2 (preregistered: open PDF here), we examined how intergroup bias affects dishonest behavior *without* personal gain among Democrats and Republicans. Like Experiment 1, we recruited U.S. Americans who identified as either Democrats or Republicans, used the mind-game task as the measure of behavioral dishonesty, and measured the same individual-level measures. Participants who completed Experiment 1 were excluded from the recruitment call.

Using a between-participants design, we randomly assigned participants to one of three conditions. The only difference across conditions was who would be the potential recipient of a \$0.75 bonus based on a self-reported "correct guess" of the outcome of a die-roll (an ingroup or outgroup member), and what would happen if self-reporting an incorrect guess (no bonus or bonus to the outgroup). The three conditions are as follows: (1) an outgroup member receives a bonus (correct guess) vs. no bonus (incorrect guess), (2) an ingroup member receives a bonus (correct guess) vs. no bonus (incorrect guess), and (3) an ingroup member receives a bonus (correct guess) vs. an outgroup receives a bonus (incorrect guess).

5.1. Method

5.1.1. Participants and power

We received complete responses from 900 Democrat and 897 Republican-identifying U.S. American participants. We recruited participants using filters of U.S.A nationality, reported political party affiliation (Democrat or Republican), being fluent in English, having an approval rate of 95% and higher in 50 or more submissions, and a 1:1 sex balance. Participants received a base payment of \$0.75. Similar to Experiment 1, we implemented two recruitment arms in Prolific based on participants' stated party membership, and randomly assigned them to one of three conditions.

As per our pre-registration, we excluded responses from participants flagged as duplicate or fraudulent by Qualtrics (N = 54) and who failed the simple attention check at the start of the study (N = 33). In the attention check question, participants were presented with general instructions about the importance of reading instructions carefully and not rushing through the survey, before being asked to select “A lot” to the question “How much do you like painting?” when presented with five options as potential answers. This left us with a final dataset consisting of 1,710 valid responses (M_age = 43.35, SD = 14.25; 796 male, 896 female, 18 other; 1322 White, 149 Black, 104 Asian, 86 Mixed, 44 Other, 5 unknown).

A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample could detect an effect size of w = .09 or larger in pairwise comparisons using "goodness-of-fit" contingency tables Chi-square test (df = 1, $p < .05$) with 90% power.

5.1.2. Procedure

Experiment 2 consisted of four main parts: (1) demographic information, (2) honesty task, (3) belief measures, and (4) individual measures.

Demographic information. Like Experiment 1, right after the attention check question, participants responded to demographic measures such as age, gender, education level, household income, political party identification, voting history in the 2020 U.S. Presidential election, and voting intention in the 2024 election.

Honesty task. For our outcome measure of behavioral dishonesty, participants completed the mind-game paradigm, just like in Experiment 1. Participants were presented with instructions for a die-roll game where they were to guess in their minds in advance whether the outcome of a future die-roll would be odd or even. They were asked to remember their guess, and upon observing the result, report honestly whether their guess turned out to be correct or not.

Unlike Experiment 1, our experimental manipulation was implemented in the very bonus mechanism part. In Condition 1, participants were informed that if they guessed correctly, a *same*-party voter would receive a $0.75 bonus; and if they guessed incorrectly, there would be no bonus. In Condition 2, participants were informed that if they guessed correctly, an *opposing* party voter would receive a $0.75 bonus; and if they guessed incorrectly, there would be no bonus. In Condition 3, participants were informed that if they guessed correctly, a same party voter would receive a $0.75 bonus; and if they guessed incorrectly, an opposing party voter would receive a $0.75. In all three conditions, participants were told that the outcome of the die roll does not affect their personal payoffs and that their identity remains anonymous.

After confirming they understood the instructions, participants were directed to the game pages, the same as in Experiment 1. Participants guessed odd or even in their minds, rolled the die, observed the score, and confirmed the number. But unlike Experiment 1, here participants were presented with options to self-report based on their randomly assigned *conditions*.

When asked to self-report whether they guessed in advance correctly concerning whether the number would be odd (1, 3, 5) or even (2, 4, 6), participants were reminded of the outcomes of reporting their guess as correct or incorrect. Depending on the condition, the choice "YES, I guessed correctly" was accompanied by an additional statement that the same voting participant would receive $0.75 (Condition 1 and 3), or the opposing voting participant would receive a $0.75 bonus (Condition 2). The choice "NO, I did not guess correctly" was accompanied by an additional statement that there would be no bonus (Condition 1 and 2) or an opposing voting participant receiving a $0.75 bonus (Condition 3). Please see section *2.2. Experimental Manipulations* for a participant perspective preview across conditions.

Belief measures. After the individual-level measures, participants were asked, in separate questions, what percentage of Democratic and Republican Party voters completing the same study as them they thought would falsely state they guessed correctly in the die-roll game. The order of the questions was randomized with each question being presented on a separate page. Like Experiment 1, participants were incentivized with a $1 additional bonus for the closest five guesses.

Individual measures. Like Experiment 1, participants completed several individual-level measures, including a 14-item identification with the ingroup scale (Leach et al., 2008; a = .973) and a 9-item sectarianism scale (Landry et al., 2024; a = .944) on scales anchored at 1 (strongly disagree) and 7 (strongly agree) to each. Participants also reported their affective polarization, rating their feelings toward the Republican Party and the Democratic Party from 0 (Very cold) to 100 (Very warm), with feelings toward the opposing party reverse-coded.

At the very end of the study, participants were asked to indicate their extent of agreement to three single-item measures exploring (lack of anticipated guilt ("If I were to cheat in this task, I would <u>not</u> feel guilty about it"), cheating expectation ("In this task, the Democratic/ Republican Party voter would have expected me to cheat."), and expected reciprocity ("In this task, the Democratic/ Republican Party voter would have cheated if I was on the receiving end.") on a 7-point scale anchored at 1 (Strongly disagree) to 7 (Strongly agree). Please see Table S2 in Supplemental Materials for a full list of measures used in Experiment 2.

5.2 Results

5.2.1. Preregistered analyses

Behavioral dishonesty. We preregistered the hypothesis that "The proportion of participants reporting a correct guess will be higher when a correct guess generates a $0.75 bonus to an ingroup member vs. no bonus if incorrect, than when a correct guess generates a $0.75 bonus to an outgroup member vs. no bonus if incorrect, between-subjects." A 2 X 2 Chi-square test comparing Condition 1 (beneficiary = ingroup) vs. Condition 2 (beneficiary = outgroup) showed that the proportion of correct guesses in the ingroup condition (61.9%) was significantly higher than that in the outgroup condition (52.8%), continuity corrected χ^2 (1, 1137) = 9.37, p = .002, OR = 1.45, 95% CI [1.15, 1.84]. This translates to a "small" effect size of d = .20 (95% CI [.08, .34]), providing statistically significant support to our hypothesis that people are more likely to lie to benefit an ingroup than an outgroup member.

We also preregistered that, “The proportion of participants reporting a correct guess will be higher when a correct guess generates a $0.75 bonus to an ingroup member vs. a $0.75 bonus to an outgroup member if incorrect, than when a correct guess generates a $0.75 bonus to an outgroup member vs. no bonus if incorrect, between-subjects.” A 2 X 2 Chi-square test comparing Condition 1 vs. 3 showed that the proportion of correct guesses in the ingroup condition (60.0%) was statistically higher than in the outgroup condition (52.8%), continuity corrected χ^2 (1, 1126) = 5.75, p = .016, OR = 1.34, 95% CI [1.06, 1.70]. This translates to a “small” effect size of d = .17 (95% CI [.03, .29]), providing statistically significant support to our hypothesis that people are more likely to lie to benefit an ingroup at the cost of an outgroup member than to simply benefit an outgroup member.

Finally, we preregistered the hypothesis that “The proportion of participants reporting a correct guess will be higher than the statistical expectation of 50% when a correct guess generates a $0.75 bonus to an ingroup member vs. a $0.75 bonus to an outgroup member if incorrect, within-subjects. A binomial proportion test showed that the proportion of correct guesses within Condition 3, where an ingroup could benefit from a lie at the cost of an outgroup, (.60) was significantly higher than the statistical expectation of .50 if guesses about the outcome of the die roll were reported truthfully, p < .001, 95% CI [.56, .64], Bayes factor = 4698, 95% credibility intervals [.56, .64]. Thus, we find statistically highly significant support for our hypothesis of a significant rate of lying to benefit an ingroup at the cost of an ingroup.

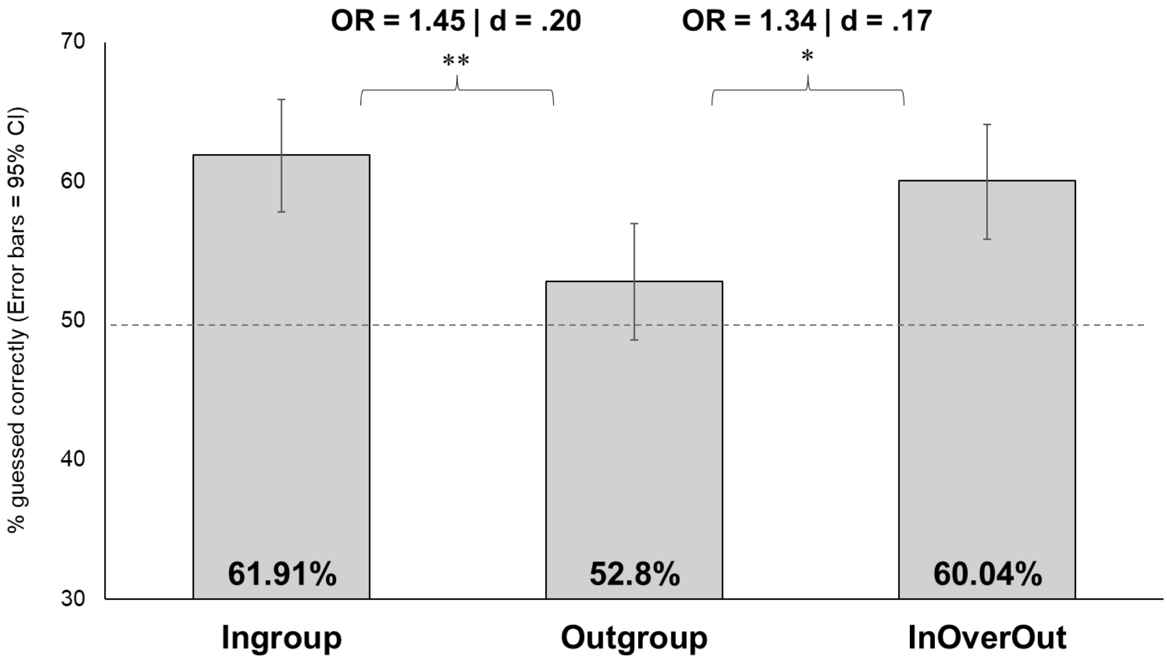


Figure 4. Percentage of reported correct guesses across conditions. Experiment 3 (N = 1,710). Compared to the percentage reporting a correct guess of a future die roll being odd or even to benefit an outgroup (52.8%), participants were more likely to report a correct guess both when it simply benefitted an ingroup member (61.91%, or an ingroup member at the cost of an outgroup member (60.04%).

Dishonesty beliefs. We preregistered the hypothesis that "Individuals will overestimate the likelihood of dishonest behavior among political outgroup members, when comparing individuals' predicted rates of dishonesty in the mind game with the actual rates of behavioral dishonesty." A one-sample t-test showed that the difference between the estimated and actual rate of dishonest behavior ($M = 35.12$, $SD = 25.70$) among outgroups was significantly higher than 0, $t(1, 1710) = 56.5$, $p < .001$, $d = 1.37$, 95% CI [1.30, 1.43]. This provides statistically highly significant support for our hypothesis, and suggests the presence of a "large" effect concerning overestimating the dishonest behavior of outgroup members.

We also preregistered the hypothesis that "Individuals will overestimate the likelihood of dishonest behavior among political outgroup members more than among political ingroup members, meaning that the difference between predicted and actual dishonesty will be larger when considering outgroup behavior than ingroup behavior." A paired-sample t-test showed that the *difference* between self-reported and actual dishonesty rates for outgroups ($M = 35.12$, $SD = 25.70$) was greater than that for ingroups ($M = 25.84$, $SD = 23.89$), $t(1709) = 17.19$, $p < .001$, $d = .42$, 95% CI [.37, .46]. This provides statistically highly significant support for our hypothesis, and suggests the presence of a "small-to-medium" effect concerning overestimating the dishonest behavior of outgroup members <u>more</u> than that of ingroup members.

5.2.2. Secondary analyses

Behavioral dishonesty. How do ingroup love and outgroup hate underlie intergroup bias in other-benefitting dishonesty? We conducted two non-preregistered analyses to explore this. A binomial proportion test showed that the proportion of correct guesses within Condition 1 (.62), where simply an ingroup could benefit from a lie, was significantly higher than the statistical expectation of .50 if guesses about the outcome of the die roll were reported truthfully, $p < .001$, 95% CI [.58, .66], Bayes factor = 727247, 95% credibility intervals [.58, .66]. This suggests the presence of ingroup love in other benefitting dishonesty, as people lie at a significant rate to simply benefit an ingroup member.

Another binomial proportion test showed that the proportion of correct guesses within Condition 2 (.63), where simply an outgroup could benefit from a lie, was not significantly different from the statistical expectation of .50 if guesses about the outcome of the die roll were reported truthfully, $p = .195$, 95% CI [.49, .57], Bayes factor = .128, 95% credibility intervals [.49, .57]. This suggests the *absence* of outgroup hate in other benefitting dishonesty, as there seemed to be no significant rate of lying on average, neither to benefit nor harm the earnings of outgroup members.

Given that we found speculative evidence of moderation in intergroup bias in dishonesty Experiment 1, we conducted a series of binary logistic regressions with reported guesses as the outcome (0 = incorrect, 1 = correct) and predictors of interactions of moderators (preregistered under secondary analysis) and dummy variables of experimental conditions.

First, we explore moderation of the simply lying for ingroup vs. outgroup effect. The coefficients for the condition (0 = outgroup, 1 = ingroup) X ingroup identification scale interaction (OR = 1.25, 95% CI [1.04, 1.50], p = .015) and condition X sectarianism interaction were significant (OR = 1.29, 95% CI [1.10, 1.51], p = .001), together suggesting individual-level differences in the strength of the effects.

Exploring the moderating role of identification with the ingroup, follow-up floodlight analyses (Spiller et al., 2013) of the Johnson-Neyman test showed that the outgroup vs. ingroup effect was significant when ingroup identification was *outside* the interval [-4.89, 3.88]. Simple slopes analysis showed that along the ingroup identification scale, the slope of the experimental condition (0 = outgroup, 1 = ingroup) at -1 SD (3.11) was non-significant (b = .80, SE = .17, z = 0.50, p = .62), at the mean (4.43) significant and positive (b = .38, SE = .12, z = 3.12, p < .01), and at + 1 SD (5.75) significant and positive (b = .67, SE = .17, z = 3.86, p < .01). This suggests that individuals reporting higher levels of identification with their ingroup are more likely to lie to simply benefit an ingroup (vs. outgroup) member.

Exploring the moderating role of sectarianism, another Johnson-Neyman test showed that the outgroup vs. ingroup effect was significant when sectarianism was *outside* the interval [-0.13, 3.75]. Simple slopes analysis showed that along the sectarianism scale, the slope of the experimental condition (0 = outgroup, 1 = ingroup) at -1 SD (2.74) was non-significant ($b = -0.00$, $SE = .17$, $z = -0.01$, $p = .99$), at the mean (4.29) significant and positive ($b = .39$, $SE = .12$, $z = 3.19$, $p < .001$), and at + 1 SD (5.82) significant and positive ($b = .78$, $SE = .17$, $z = 4.55$, $p < .01$). This suggests that individuals reporting higher levels of sectarianism are more likely to lie to simply benefit an ingroup (vs. outgroup) member.

Second, we explore moderation of the lying to benefit the ingroup at the cost of an outgroup vs. simply benefitting an outgroup effect. Here, although the coefficient for condition (0 = outgroup, 1 = ingroup at the cost of outgroup) X sectarianism interaction was significant (OR = 1.45, 95% CI [1.24, 1.70], $p < .001$), the coefficient for the condition X ingroup identification scale interaction was not significant (OR = 1.11, 95% CI [0.92, 1.33], $p = .264$). Follow-up analyses of the Johnson-Neyman test showed that the outgroup vs. ingroup effect was significant when sectarianism was *outside* the interval [2.47, 4.12]. Simple slopes analysis showed that along the sectarianism scale, the slope of the experimental condition (0 = outgroup, 1 = ingroup at the cost of outgroup) at -1 SD (2.73) was non-significant ($b = -0.27$, $SE = .17$, $z = -1.57$, $p = .12$), at the mean (4.29) significant and positive ($b = .30$, $SE = .12$, $z = 2.48$, $p = .01$), and at + 1 SD (5.83) significant and positive ($b = .88$, $SE = .17$, $z = 5.11$, $p < .001$). This suggests that individuals reporting higher levels of sectarianism are more likely to lie to benefit an ingroup member at the cost of an outgroup.

6. General Discussion

The current research examined two primary questions using incentivized pre-registered experiments: (1) Are people more likely to cheat outgroups than ingroups for personal gain? (2) Under conditions of no personal gain, are people more likely to cheat when their ingroup than their outgroups receive the gain? In Experiment 1, we found that individuals lied at the same rate for personal gain, irrespective of costing their outgroup or ingroup members. This suggests an absence of intergroup bias in dishonesty for *personal* gain. In Experiment 2, where the decision-maker's own outcomes were removed from the equation, individuals lied at a significantly higher rate to benefit their ingroups. This suggests the presence of intergroup bias in other-benefitting behavior in the form of coalitional dishonesty.

In both our experiments, people could lie anonymously under conditions of zero detection, no punishment, and without any reputational consequences. Personal material incentives were either held constant at $0.75 in Experiment 1 or completely eliminated in Experiment 2 where lying would benefit another participant by the same $0.75 without affecting own outcomes. Further, in both experiments, participants made one-shot decisions while knowing that their decisions would not affect decisions by others toward them- which enables us to rule out conflation with reciprocal expectations, a key factor underlying ingroup favoritism (Yamagishi & Kiyonari, 2000).

Findings from Experiment 1 add to the growing literature (such as Rotman et al., 2018; Soraperra et al., 2019; Yam & Reynolds, 2016) on how characteristics of the victim can affect dishonest behavior against them. Our specific result that political group identities did not affect cheating another person for personal gain is surprising. Because if there was an intergroup bias in dishonesty for personal gain, we would have been likely to find it a setting with intergroup animosity as high as between Democrats and Republicans (Kranton et al., 2020). This perhaps suggests a positive aspect in the sense that people may not adopt different moral standards for dishonest behavior against ingroup versus outgroup members. When it comes to selfish lying to increase personal payoffs, people are similarly likely to cheat ingroups and outgroups. As a result, we speculate that self-interest may trump group identities in the moral domain. To that end, it may also be worth seeing this in light of how several "honesty nudges" have been found to be ineffective in consistently reducing incentivized dishonest behavior (e.g., Kristal et al., 2020).

In contrast to the recent stream of research that finds political group identities are especially likely to be a factor in important counterproductive judgments and behaviors such as animosity on social media (Rathje et al., 2021), processing misinformation (Langdon et al., 2024; Rathje et al., 2023), and even wealth destruction (Mill & Morgan, 2022), we find no effect of intergroup bias in interpersonal dishonesty. Nonetheless, the null finding in Experiment 1 aligns with some of the null findings in related domains. For example, in a minimal group (N = 196) study conducted in a competitive setting, participants were similarly likely to cheat to increase their payoffs at the cost of their competitors regardless of sharing the same group identity or not (Benistant & Villeval, 2019).

Findings from Experiment 2 add to the previously suggestive evidence from prior research in experimental economics (Aksoy & Palma, 2019; Cadsby et al., 2016; Michailidou & Rotondi, 2019) that people may be more likely to cheat to benefit their ingroup than their outgroup. Our findings add empirical robustness in the sense that we found people are more likely to lie to benefit their ingroups, regardless of the benefit coming at the cost of the experimenter *or* their outgroup. This suggests that when personal benefits are eliminated, the group identity of the beneficiary seems to influence the "pure" moral costs of cheating. That is, people seem to impose *lower* moral standards for their behavior when it can benefit an ingroup than they would if it would benefit an outgroup. Further, we found that ingroup love, not outgroup hate, is an important explanation for this biased dishonesty. This is based on our observation of no significant negative difference from the statistical expectation of 50% correct guesses, suggesting people did not lie on average to harm the earnings of an outgroup member. To that end, we also add a group psychology dimension to the literature on altruistic lying (Brocas & Carrillo, 2021; Erat & Gneezy, 2012) and what drives it, by documenting the presence of *ingroup love* underlying intergroup bias in other-benefitting dishonesty.

The fact that we found evidence of moderation, especially in Experiment 2, suggests that coalitional lying for the ingroup compared to the outgroup is more likely for individuals who report higher levels of identification with the ingroup or sectarianism. This makes intuitive sense that intergroup bias in other-benefitting dishonesty should be higher for more polarized individuals, as supported by similar evidence of moderation in the literature.

Practically, findings from the current research paint a mixed picture. In the current age of challenges for intergroup cooperation and societal polarization, it seems that the group dimension may not affect egotistic lying. People may not be any more likely to selfishly cheat their outgroups than their ingroups. On the positive side, it may imply that individuals in negotiations, transactions, and other economic interactions may not cheat someone more just because the potential victim is an outgroup. On the negative side, it may also be that seeing a potential victim as part of the ingroup may not curb the dishonest pursuit of selfish gains at their cost. Keeping this in mind may be beneficial to policymakers, peer-to-peer marketplace managers, and also social workers interested in intergroup relations.

Further, our robust finding that people may be more likely to cheat to benefit their ingroups than their outgroups, regardless of the costs being borne by a third party or an outgroup, may be relevant to organizational leaders and decision-makers at large. Broadly, employees may actively lie, or misreport factual information merely based on the premise that it can benefit another employee merely by virtue of sharing a group identity. More specific examples may include lying in different capacities for another employee from the same team, department, or firm to exploit customers, cover up wrongdoings, and even secure promotions. This can imply that auditors and regulators should consider group identities in examining statements made about others, even when the statement-maker does not have any clear incentive to lie. To that end, ethical training content within organizations may benefit from incorporating the potential effects of group identities in other-benefitting dishonesty. Even societally, recognizing the influence of group identities in moral decision-making may help reduce the negative effects of intergroup biases.

7. Limitations and Future Research

Although our results have important implications for intergroup relations and dishonesty literature, several limitations must be acknowledged. First, our interpretations are based on online experiments in the political group identity context with US American participants. So, we do not claim intergroup dishonesty for personal gain may be absent in all groups across cultures; nor do we claim that intergroup dishonesty to benefit others may be present in all groups across cultures. So, future research may find it fruitful to test if our findings generalize beyond Democrat and Republican-voting participants as decision-makers.

Another limitation of our experiments is that our design allowed comparison of *rates* of dishonesty across conditions and not levels of dishonesty. That is, across both experiments, participants could only make the decision to cheat or not, and not *how much* to cheat. Perhaps the null findings from Experiment 1 speak only to intergroup dishonesty in a binary decision context. Future research may benefit from also comparing levels of dishonesty for personal gain using a *continuous* measure of behavioral dishonesty.

In Experiment 2, we found no evidence of antisocial cheating- that is, falsely reporting that the guess was wrong to *not* benefit the outgroup. This was also surprising as a finding “outgroup hate” (e.g., Dimant, 2023) would have predicted that people may have also cheated to deprive an outgroup of receiving the $0.75. So, future research may benefit from using a continuous measure to test levels of antisocial cheating rather than in a binary context like ours.

Finally, we tested our hypotheses on intergroup bias in dishonesty in controlled settings which although appropriate for basic research in judgment and decision-making, would be interesting to test how our findings may compare with dishonesty in more applied settings. For example, are people also likely to lie and benefit their ingroups more in organizational/workplace settings? Are consumers differently or similarly likely to commit returns and/or insurance fraud against businesses with similar or different political identity signals? It may also be interesting to compare real-world base rates. For example, do citizens cheat more on their taxes at different rates when their affiliated versus opposing political party is elected?

Despite the limitations, our research found mixed effects of intergroup bias in dishonesty opening avenues for further research at the intersection of intergroup relations and morality. For now, we conclude that group identities may not influence self-interested dishonesty in the narrow sense, but that it robustly promotes other-benefitting behavior in the form of *coalitional dishonesty*, in which ingroup love, but not outgroup hate, reduces the moral cost of lying.

References

Abbink, K., & Harris, D. (2019). In-group favouritism and out-group discrimination in naturally occurring groups. *PLOS ONE*, *14*(9), 1–13. https://doi.org/10.1371/journal.pone.0221616

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for Truth-Telling. *Econometrica*, *87*(4), 1115–1153. https://doi.org/10.3982/ECTA14673

Aksoy, B., & Palma, M. A. (2019). The effects of scarcity on cheating and in-group favoritism. *Journal of Economic Behavior & Organization*, *165*, 100–117. https://doi.org/10.1016/j.jebo.2019.06.024

Axelrod, R. (1986). An Evolutionary Approach to Norms. *American Political Science Review*, *80*(4), 1095–1111. https://doi.org/10.2307/1960858

Ayal, S., Celse, J., & Hochman, G. (2019). Crafting messages to fight dishonesty: A field investigation of the effects of social norms and watching eye cues on fare evasion. *Organizational Behavior and Human Decision Processes*, S0749597818305004. https://doi.org/10.1016/j.obhdp.2019.10.003

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

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. https://doi.org/10.1111/j.1745-6916.2007.00051.x

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. In N. G. Fielding, A. Clarke, & R. Witt (Eds.), *The Economic Dimensions of Crime* (pp. 13–68). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-62853-7_2

Benistant, J., & Villeval, M. C. (2019). Unethical behavior and group identity in contests. *Journal of Economic Psychology*, *72*, 128–155. https://doi.org/10.1016/j.joep.2019.03.001

Brass, D. J., Butterfield, K. D., & Skaggs, B. C. (1998). Relationships and Unethical Behavior: A Social Network Perspective. *The Academy of Management Review*, *23*(1), 14. https://doi.org/10.2307/259097

Brewer, M. B. (1999). The Psychology of Prejudice: Ingroup Love and Outgroup Hate? *Journal of Social Issues*, *55*(3), 429–444. https://doi.org/10.1111/0022-4537.00126

Brocas, I., & Carrillo, J. D. (2021). Self-serving, altruistic and spiteful lying in the schoolyard. *Journal of Economic Behavior & Organization*, *187*, 159–175. https://doi.org/10.1016/j.jebo.2021.04.024

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

Cárdenas, D., & de la Sablonnière, R. (2020). Participating in a new group and the identification processes: The quest for a positive social identity. *British Journal of Social Psychology*, *59*(1), 189–208. https://doi.org/10.1111/bjso.12340

Castano, E., Yzerbyt, V., Bourguignon, D., & Seron, E. (2002). Who May Enter? The Impact of In-Group Identification on In-Group/Out-Group Categorization. *Journal of Experimental Social Psychology*, *38*(3), 315–322. https://doi.org/10.1006/jesp.2001.1512

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual Behavior and Group Membership. *American Economic Review*, *97*(4), 1340–1352. https://doi.org/10.1257/aer.97.4.1340

Cikara, M. (2021). Causes and consequences of coalitional cognition. In *Advances in Experimental Social Psychology* (Vol. 64, pp. 65–128). Elsevier. https://doi.org/10.1016/bs.aesp.2021.04.002

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Cikara, M., Bruneau, E., Van Bavel, J. J., & Saxe, R. (2014). Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of Experimental Social Psychology*, *55*, 110–125. https://doi.org/10.1016/j.jesp.2014.06.007

Cikara, M., Van Bavel, J. J., Ingbretsen, Z. A., & Lau, T. (2017). Decoding "us" and "them": Neural representations of generalized group concepts. *Journal of Experimental Psychology: General*, *146*(5), 621–631. https://doi.org/10.1037/xge0000287

Cohn, A., Gesche, T., & Maréchal, M. A. (2022). Honesty in the Digital Age. *Management Science*, *68*(2), 827–845. https://doi.org/10.1287/mnsc.2021.3985

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

Cosmides, L., & Tooby, J. (2013). Evolutionary Psychology: New Perspectives on Cognition and Motivation. *Annual Review of Psychology*, *64*(1), 201–229. https://doi.org/10.1146/annurev.psych.121208.131628

Curry, O. S. (2016). Morality as Cooperation: A Problem-Centred Approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The Evolution of Morality* (pp. 27–51). Springer International Publishing. https://doi.org/10.1007/978-3-319-19671-8_2

Curry, O. S., Jones Chesters, M., & Van Lissa, C. J. (2019). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire.

Journal of Research in Personality, *78*, 106–124. https://doi.org/10.1016/j.jrp.2018.10.008

Dimant, E. (2023). Hate Trumps Love: The Impact of Political Polarization on Social Preferences. *Management Science*, mnsc.2023.4701. https://doi.org/10.1287/mnsc.2023.4701

Earp, B. D., McLoughlin, K. L., Monrad, J. T., Clark, M. S., & Crockett, M. J. (2021). How social relationships shape moral wrongness judgments. *Nature Communications*, *12*(1), 5776. https://doi.org/10.1038/s41467-021-26067-4

Efferson, C., Bernhard, H., Fischbacher, U., & Fehr, E. (2024). Super-additive cooperation. *Nature*, *626*(8001), 1034–1041. https://doi.org/10.1038/s41586-024-07077-w

Ellemers, N., De Gilder, D., & Haslam, S. A. (2004). Motivating Individuals and Groups at Work: A Social Identity Perspective on Leadership and Group Performance. *The Academy of Management Review*, *29*(3), 459. https://doi.org/10.2307/20159054

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

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

Gibson, R., Tanner, C., & Wagner, A. F. (2013). Preferences for Truthfulness: Heterogeneity among and within Individuals. *American Economic Review*, *103*(1), 532–548. https://doi.org/10.1257/aer.103.1.532

Gneezy, U., Kajackaite, A., & Sobel, J. (2018). Lying Aversion and the Size of the Lie. *American Economic Review*, *108*(2), 419–453. https://doi.org/10.1257/aer.20161553

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral Foundations Theory. In *Advances in Experimental Social Psychology* (Vol. 47, pp. 55–130). Elsevier. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523. https://doi.org/10.1016/S1364-6613(02)02011-9

Halevy, N., Bornstein, G., & Sagiv, L. (2008). "In-Group Love" and "Out-Group Hate" as Motives for Individual Participation in Intergroup Conflict: A New Game Paradigm. *Psychological Science*, *19*(4), 405–411. https://doi.org/10.1111/j.1467-9280.2008.02100.x

Hochman, G., Peleg, D., Ariely, D., & Ayal, S. (2021). Robin Hood meets Pinocchio: Justifications increase cheating behavior but decrease physiological tension. *Journal of Behavioral and Experimental Economics*, *92*, 101699. https://doi.org/10.1016/j.socec.2021.101699

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, *93*, 328–336. https://doi.org/10.1016/j.jebo.2013.04.003

Köbis, N. C., Verschuere, B., Bereby-Meyer, Y., Rand, D., & Shalvi, S. (2019). Intuitive Honesty Versus Dishonesty: Meta-Analytic Evidence. *Perspectives on Psychological Science*, *14*(5), 778–796. https://doi.org/10.1177/1745691619851778

Kocher, M. G., Schudy, S., & Spantig, L. (2018). I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. *Management Science*, *64*(9), 3995–4008. https://doi.org/10.1287/mnsc.2017.2800

Kranton, R., Pease, M., Sanders, S., & Huettel, S. (2020). Deconstructing bias in social preferences reveals groupy and not-groupy behavior. *Proceedings of the National Academy of Sciences*, *117*(35), 21185–21193. https://doi.org/10.1073/pnas.1918952117

Kteily, N., Hodson, G., & Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, *110*(3), 343–370. https://doi.org/10.1037/pspa0000044

Kristal, A. S., Whillans, A. V., Bazerman, M. H., Gino, F., Shu, L. L., Mazar, N., & Ariely, D. (2020). Signing at the beginning versus at the end does not decrease dishonesty. *Proceedings of the National Academy of Sciences*, *117*(13), 7103–7107. https://doi.org/10.1073/pnas.1911695117

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Landry, A., Finkel, E., Hoyle, R. H., Druckman, J., & Van Bavel, J. J. (2024). *Partisan Antipathy and the Erosion of Democratic Norms* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/ahgy6

Langdon, J. A., Helgason, B. A., Qiu, J., & Effron, D. A. (2024). “It’s Not Literally True, But You Get the Gist:” How nuanced understandings of truth encourage people to condone and spread misinformation. *Current Opinion in Psychology*, *57*, 101788. https://doi.org/10.1016/j.copsyc.2024.101788

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., Ouwerkerk, J. W., & Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, *95*(1), 144–165. https://doi.org/10.1037/0022-3514.95.1.144

Leib, M., Köbis, N., Soraperra, I., Weisel, O., & Shalvi, S. (2021). Collaborative dishonesty: A meta-analytic review. *Psychological Bulletin*, *147*(12), 1241–1268. https://doi.org/10.1037/bul0000349

Lüdecke, D. (2023). *sjPlot: Data visualization for statistics in social science* (R package version 2.8.15) [Computer software]. https://vps.fmvz.usp.br/CRAN/web/packages/sjPlot/sjPlot.pdf

Mende-Siedlecki, P., Qu-Lee, J., Backer, R., & Van Bavel, J. J. (2019). Perceptual contributions to racial bias in pain recognition. *Journal of Experimental Psychology: General*, *148*(5), 863–889. https://doi.org/10.1037/xge0000600

Michailidou, G., & Rotondi, V. (2019). I'd lie for you. *European Economic Review*, *118*, 181–192. https://doi.org/10.1016/j.euroecorev.2019.05.014

Mill, W., & Morgan, J. (2022). The cost of a divided America: An experimental study into destructive behavior. *Experimental Economics*, *25*(3), 974–1001. https://doi.org/10.1007/s10683-021-09737-4

Morewedge, C. K., Tang, S., & Larrick, R. P. (2018). Betting Your Favorite to Win: Costly Reluctance to Hedge Desired Outcomes. *Management Science*, *64*(3), 997–1014. https://doi.org/10.1287/mnsc.2016.2656

Oldmeadow, J. A., & Fiske, S. T. (2010). Social status and the pursuit of positive social identity: Systematic domains of intergroup differentiation and discrimination for high- and low-status groups. *Group Processes & Intergroup Relations*, *13*(4), 425–444. https://doi.org/10.1177/1368430209355650

Pascual-Ezama, D., Fosgaard, T. R., Cardenas, J. C., Kujal, P., Veszteg, R., Gil-Gómez de Liaño, B., Gunia, B., Weichselbaumer, D., Hilken, K., Antinyan, A., Delnoij, J., Proestakis, A., Tira, M. D., Pratomo, Y., Jaber-López, T., & Brañas-Garza, P. (2015). Context-dependent cheating: Experimental evidence from 16 countries. *Journal of*

Economic Behavior & Organization, *116*, 379–386. https://doi.org/10.1016/j.jebo.2015.04.020

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology*, *59*(3), 475–486. https://doi.org/10.1037/0022-3514.59.3.475

Pratto, F., & Glasford, D. E. (2008). Ethnocentrism and the value of a human life. *Journal of Personality and Social Psychology*, *95*(6), 1411–1428. https://doi.org/10.1037/a0012636

Qualtrics (Version 2022). (2022). [Computer software]. Qualtrics. https://www.qualtrics.com

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, *114*(41), 10870–10875. https://doi.org/10.1073/pnas.1706255114

R Core Team. (2013). *R* [Computer software]. R Foundation for Statistical Computing. www.R-project.org/

Rathje, S., Roozenbeek, J., Van Bavel, J. J., & Van Der Linden, S. (2023). Accuracy and social motivations shape judgements of (mis)information. *Nature Human Behaviour*, *7*(6), 892–903. https://doi.org/10.1038/s41562-023-01540-w

Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences*, *118*(26), e2024292118. https://doi.org/10.1073/pnas.2024292118

Romano, A., Balliet, D., & Wu, J. (2017). Unbounded indirect reciprocity: Is reputation-based cooperation bounded by group membership? *Journal of Experimental Social Psychology*, *71*, 59–67. https://doi.org/10.1016/j.jesp.2017.02.008

Rotman, J. D., Khamitov, M., & Connors, S. (2018). Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically. *Journal of Consumer Psychology*, *28*(2), 353–361. https://doi.org/10.1002/jcpy.1002

Scheel, A. M., Schijen, M. R. M. J., & Lakens, D. (2021). An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. *Advances in Methods and Practices in Psychological Science*, *4*(2), 251524592110074. https://doi.org/10.1177/25152459211007467

Shalvi, S., & De Dreu, C. K. W. (2014). Oxytocin promotes group-serving dishonesty. *Proceedings of the National Academy of Sciences*, *111*(15), 5503–5507. https://doi.org/10.1073/pnas.1400724111

Soraperra, I., Weisel, O., & Ploner, M. (2019). Is the victim *Max (Planck)* or *Moritz* ? How victim type and social value orientation affect dishonest behavior. *Journal of Behavioral Decision Making*, *32*(2), 168–178. https://doi.org/10.1002/bdm.2104

Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., & McClelland, G. H. (2013). Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression. *Journal of Marketing Research*, *50*(2), 277–288. https://doi.org/10.1509/jmr.12.0420

Sterling, T. D. (1959). Publication Decisions and their Possible Effects on Inferences Drawn from Tests of Significance—Or Vice Versa. *Journal of the American Statistical Association*, *54*(285), 30–34. https://doi.org/10.1080/01621459.1959.10501497

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*(3), 364–373. https://doi.org/10.1037/0022-3514.56.3.364

Tajfel, H. (1979). Individuals and groups in social psychology*. *British Journal of Social and Clinical Psychology*, *18*(2), 183–190. https://doi.org/10.1111/j.2044-8260.1979.tb00324.x

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

The jamovi project. (2022). *Jamovi.* (2.3) [Computer software]. https://www.jamovi.org

Thielmann, I., & Hilbig, B. E. (2019). No gain without pain: The psychological costs of dishonesty. *Journal of Economic Psychology*, *71*, 126–137. https://doi.org/10.1016/j.joep.2018.06.001

Tooby, J., & Cosmides, L. (2007). Evolutionary psychology, ecological rationality, and the unification of the behavioral sciences. *Behavioral and Brain Sciences*, *30*(1), 42–43. https://doi.org/10.1017/S0140525X07000854

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755

Turner, J. C., & Reynolds, K. J. (2012). Self-Categorization Theory. In P. Van Lange, A. Kruglanski, & E. Higgins, *Handbook of Theories of Social Psychology* (pp. 399–417). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222.n46

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A., Jetten, J., … Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, *4*(5), 460–471. https://doi.org/10.1038/s41562-020-0884-z

Van Bavel, J. J., & Packer, D. J. (2021). *The power of us: Harnessing our shared identities to improve performance, increase cooperation, and promote social harmony* (First edition). Little, Brown Spark.

van Lent, T., Verwijmeren, T., & Bijlstra, G. (2023). Dishonest collaboration in an intergroup context. *British Journal of Social Psychology*, bjso.12675. https://doi.org/10.1111/bjso.12675

Vincent, L. C., Emich, K. J., & Goncalo, J. A. (2013). Stretching the Moral Gray Zone: Positive Affect, Moral Disengagement, and Dishonesty. *Psychological Science*, *24*(4), 595–599. https://doi.org/10.1177/0956797612458806

Vives, M.-L., Cikara, M., & FeldmanHall, O. (2022). Following Your Group or Your Morals? The In-Group Promotes Immoral Behavior While the Out-Group Buffers Against It. *Social Psychological and Personality Science*, *13*(1), 139–149. https://doi.org/10.1177/19485506211001217

Wang, L., & Murnighan, J. K. (2017). How Much Does Honesty Cost? Small Bonuses Can Motivate Ethical Behavior. *Management Science*, *63*(9), 2903–2914. https://doi.org/10.1287/mnsc.2016.2480

Watkins, H. M., & Laham, S. (2019). The influence of war on moral judgments about harm. *European Journal of Social Psychology*, *49*(3), 447–460. https://doi.org/10.1002/ejsp.2393

Wickham, H. (2016). *Ggplot2*. Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

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

Yam, K. C., & Reynolds, S. J. (2016). The Effects of Victim Anonymity on Unethical Behavior. *Journal of Business Ethics*, *136*(1), 13–22. https://doi.org/10.1007/s10551-014-2367-5

Yamagishi, T., & Kiyonari, T. (2000). The Group as the Container of Generalized Reciprocity. *Social Psychology Quarterly*, *63*(2), 116. https://doi.org/10.2307/2695887

Yudkin, D. A., Rothmund, T., Twardawski, M., Thalla, N., & Van Bavel, J. J. (2016). Reflexive intergroup bias in third-party punishment. *Journal of Experimental Psychology: General*, *145*(11), 1448–1459. https://doi.org/10.1037/xge0000190

Zitzewitz, E. (2006). Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making. *Journal of Economics & Management Strategy*, *15*(1), 67–99. https://doi.org/10.1111/j.1530-9134.2006.00092.x

Zou, L. X., & Cheryan, S. (2022). Diversifying neighborhoods and schools engender perceptions of foreign cultural threat among White Americans. *Journal of Experimental Psychology: General*, *151*(5), 1115–1131. https://doi.org/10.1037/xge0001115

Chapter 5. Article 4: Beliefs versus Reality: People Overestimate the Actual Dishonesty of Others

PREPRINT

Working paper part of a doctoral dissertation (March 2024). All conclusions are preliminary. Please do not copy, share, or cite without the authors' permission.

Jareef Martuza[1], Hallgeir Sjåstad[1], Helge Thorbjørnsen[1,2]

[1]Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, Bergen 5045, Norway

[2] SNF, Center for Applied Research at NHH, Helleveien 30, Bergen 5045, Norway

Author notes

The authors declare to have no known conflicts of interest. Correspondence about this research should be addressed to Jareef Martuza (jareef.martuza@nhh.no) PhD student at the Norwegian School of Economics.

Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. As Norwegian laws and regulations do not require review by an institutional review board for anonymous, non-medical, low-risk research with human participants, this project was not submitted to such review. Informed consent was collected from all participants. We did not receive any funding for the current research.

Abstract

Beliefs about the dishonesty of other people can shape our behavior in powerful ways, both in the marketplace and in society. How accurate are these beliefs? And do we believe that other people are similarly, more, or less dishonest than they truly are? In a research program on moral decision-making spanning three years (2022-24) and a total of 31 different effects (N = 8,127), we placed participants in different situations where they could lie for personal gain, without any repercussions or detection risk. Crucially, we also asked all participants to estimate what percentage of other people would lie in a similar situation. Our meta-analysis of these experiments, including both incentivized choice experiments and hypothetical marketplace scenarios that were initially designed to test a broad collection of different hypotheses, revealed a significant *overestimation* of others' dishonesty by an average of 14 percentage points (meta-analytic effect: Hedge's g = 0.58). That is, people are less dishonest than we tend to think. In addition, four controlled experiments demonstrated that this overestimation is robust, irrespective of the order of belief elicitation and decision-making, and regardless of variations in the instructions. The findings reveal a pervasive tendency to overestimate the actual dishonesty of other people, suggesting a widespread belief that the world is less moral than it actually is. In the long run, such miscalibrated beliefs may undermine mutual trust in society, encourage misguided policies, and trigger over-surveillance of the public at the expense of individual freedom (e.g., employees, consumers, voters).

Keywords: dishonesty; beliefs; meta-analysis.

Beliefs versus Reality:

People Overestimate the Actual Dishonesty of Others

In the current research, we study one simple question: Do people believe that others are similarly, more, or less dishonest than they truly are? Moral decision-making in general, and truth-telling in particular, is currently one of the most active research areas in behavioral social science. Ranging from psychology (Coltheart et al., 2011), economics (Schotter & Trevino, 2014), sociology (Kluegel & Smith, 1981), organizational behavior (Greenberg, 2002) and consumer research (Rotman et al., 2018) behavior, there is now an extensive literature studying the frequency of dishonest behavior, and what specific characteristics of the social context and individual person (Gerlach et al., 2019) that are most important in shaping the basic willingness to cheat others for personal gain. However, far less is known about the *beliefs* people have about the dishonesty of others.

In our view, beliefs about dishonesty are equally important to study as actual dishonesty, because most of our actions are filtered through our mental representations of the world, rather than the external world directly. That is, it does not help much if most people are honest, if the average person does not *believe* that other people will tell the truth. Conversely, if people should turn out to be less honest than we think, that would also be important to know, as it could create its own set of problems – including a social environment of exploitation, free-riding, and other violations of mutual trust. The current paper will therefore analyze a large set of recent experiments on moral decision-making, providing high-quality data that enables us to compare *actual dishonesty* with *beliefs about dishonesty* in the exact same test situation.

1. Actual Dishonesty

Honesty is like the air around us, we do not usually notice its presence, but its absence can have detrimental consequences. Charles Ponzi's pyramid scheme, the Enron accounting scandal, and Sam Bankman-Fried defrauding crypto investors are just a few examples of severe dishonest behavior in modern times. While honesty is positively associated with trust, cooperation, and economic prosperity (Gächter & Schulz, 2016), dishonest behavior inflicts great costs on society. For example, merchandise returns fraud cost US retailers $23.2 billion in 2021 (NRF, Appriss Retail, 2022). Further, 13% of tax revenue ($428 billion) in the United States goes uncollected (Internal Revenue Service, 2022).

Beyond fraudsters and career criminals, "mostly honest" ordinary people also commit tax evasion to some extent (Bott et al., 2020), consumer fraud (National Retail Federation, 2022), and corruption bribery (Sulitzeanu-Kenan et al., 2022). The dishonesty of everyday people has therefore increasingly captured the attention of behavioral scientists, both for theoretical and societal reasons. While Becker's (1968) theory of rational crime predicts that people will cheat as long as the benefits exceed costs, it is now well-established in psychology and behavioral economics that people rarely cheat to the maximum extent (Abeler et al., 2019), basically "leaving money on the table", even under anonymity and zero punishments (Gerlach et al., 2019; Rosenbaum et al., 2014). Why? Research on moral decision-making has examined the role of incentives (Wang & Murnighan, 2017; Wiltermuth, 2011), the effects of context (Ayal et al., 2019; Cohn et al., 2022; Kocher et al., 2018), traits of the perpetrator (Pascual-Ezama et al., 2020; Vincent et al., 2013), and characteristics of the victim (Rotman et al., 2018; Soraperra et al., 2019). Another important stream of research has examined how to promote honesty using different interventions (Bicchieri et al., 2023; Bott et al., 2020; Kristal et al., 2020), albeit with mixed findings and a lack of insights into underlying processes when something worked or did not work (Hertwig & Mazar, 2022; Skowronek, 2022).

Taken together, the study of behavioral dishonesty is quite prevalent and draws broad interests. From newsrooms to research labs to economic processes in society, it is a topic of conversation at dinner tables, board rooms, and academic discourse. Of particular interest for the current research, is to gain a better understanding of whether people's *beliefs* about dishonesty are generally accurate, or whether they might be systematically biased toward overestimation or underestimation of actual dishonesty.

2. Beliefs about Dishonesty

People hold fundamental beliefs about the social world, including how similar others are to the self (Robbins & Krueger, 2005), how favorable they are (Tarrant et al., 2012), and how to attribute the behaviors of others (Hewstone, 1990). In general, beliefs can shape our expectations and behavior in important ways (Jervis, 2006), such as the decision to rent an Airbnb apartment from a stranger (Nødtvedt et al., 2021), fare systems in public transport (Galeotti et al., 2021), or whether to cooperate in a public goods dilemma (Weber et al., 2023). Importantly, beliefs can also affect behavior when they are wrong or poorly calibrated, meaning that they deviate from the best available evidence about external reality. As one timely example, during the beginning of the election year of 2024, American citizens have been *feeling* dissatisfied about the state of the economy, expressing strong beliefs about a spiraling negative trend, leading them to widespread disapproval of political leaders and harboring a pessimistic outlook of the future. This is a striking pattern; because the best available *empirical data* suggests that the US economy has performed quite well in terms of GDP growth, stock market returns, and low unemployment during the same period (Pew Research Center, 2023). So, although beliefs and external reality are presumably positively correlated over time, in some cases they may differ drastically, in which beliefs about the economy can be a stronger predictor of personal and democratic behavior than the actual economy (Dias et al., 2023).

Similarly, people also hold beliefs about how honest or dishonest other people tend to be (Bicchieri et al., 2023; Dimant & Gesche, 2023). For instance, beliefs about others' moral inclinations can shape pre-emptive actions in curbing employee misconduct (List & Momeni, 2021), facilitating transactions among strangers (Köbis et al., 2021), and even asking strangers for help (Zhao & Epley, 2022). On the one hand, believing others to be less dishonest than they truly are and developing unwarranted trust can lead to naïve behaviors and policies vulnerable to exploitation by bad actors (Köbis et al., 2021). On the other hand, if people believe others are more dishonest than they truly are, the public may develop unwarranted distrust, which may hinder economic growth (Algan & Cahuc, 2013), state functions (Herreros, 2023), and even social relations (Schilke et al., 2021). Thus, it is important to not only study behavioral dishonesty, but to also assess the realism of people's *beliefs* about dishonest behavior.

Seminal research on the "moral forecasting error" (Teper et al., 2011), compared moral beliefs with actual behavior using an elegant method. In a between-participants design, individuals were randomly assigned to either make a prediction about how they would act in a moral dilemma, or given the opportunity to make a choice in that situation. The results showed that participants in the behavior condition cheated significantly *less* for personal gain than participants in the prediction condition thought they would have done in the same situation (Teper et al., 2011). Although the authors compared estimates of predicted moral behavior to actual behavior, the current work differs from Teper and colleagues (2011) as theirs focused exclusively on potential forecasting errors in *self-prediction*. That is, how predictions about one's own moral behavior may differ from actual behavior, and not how moral beliefs about other people compare to their actual behavior. We, therefore, consider previous research on the moral forecasting error an important first step, but where the current research will expand the scope by studying the accuracy and potential bias in social beliefs.

Given the many differences in judgments about oneself versus others (Cusimano & Goodwin, 2020; Molouki & Pronin, 2015; Vazire, 2010), we consider it to be an open question whether we will find overestimation or underestimation when aggregating all our available data on predictions about the behavioral dishonesty of other people.

Regarding moral beliefs about the behavior of others, recent research shows that people wrongly believe that society as a whole has become less moral over time (Mastroianni & Gilbert, 2023), suggesting the presence of a negative bias in beliefs about others in the moral domain. However, this particular study assessed *changes* in beliefs about others' morality, compared to actual changes over time, and was not able to make a direct comparison of moral beliefs about a specific behavior compared to actual moral behavior in the same test situation. As such, this research provides relevant background for how people might mistakenly misperceive historical trends, although it cannot provide an empirical assessment of whether people have accurate beliefs in the here and now.

3. Beliefs vs. Reality: Competing Hypotheses

Although related research on the moral forecasting error (Teper et al., 2011) and the illusion of moral decline (Mastroianni & Gilbert, 2023) is highly relevant for the current research, it does *not* answer whether people believe that others are similarly, more, or less dishonest than they truly are. The answer to that question may be less straightforward, in which there are empirical and theoretical bases for competing hypotheses. For instance, a global study on civic honesty found that when asked to predict the rate of lost wallets being returned, the general population overestimated the rate of returned lost wallets (i.e., civic honesty) when the wallet contained no money and some money, but underestimated when the wallet contained more money; interestingly, professional economists overestimated the rate of returned wallets both when it contained no money and some money (Cohn et al., 2019).

This suggests that both lay people and social scientists may have inaccurate beliefs about the moral behavior of other people, and so a more focused approach is needed to assess the accuracy or potential direction of a bias in moral beliefs.

Unlike the current research, the lost wallet study was conducted in the field, and the participants making predictions did not see or experience the same test situation themselves. Thus, there might be other unobserved differences between the prediction scenario and the actual behavior scenario, representing potential confounds or competing explanations. The current meta-analysis will therefore only focus on controlled incentivized experiments, in which all participants who are making a prediction about the moral behavior of other people have just been part of the exact same test situation themselves. In our view, this method provides a unique comparison of moral beliefs to actual behavior, combining 31 different effects across more than eight thousand study participants, keeping all else constant.

In the current meta-analysis, the null hypothesis is that there is no significant difference between beliefs and actual behavior in the domain of dishonesty. If true, that would be "good news" both for ordinary citizens and policymakers, as it would suggest that the beliefs we have about each other are accurately calibrated with empirical reality – providing a solid basis to make well-informed decisions about trust, trustworthiness, and cooperation. In contrast with this view, a set of competing hypotheses suggests that beliefs and behavior will systematically deviate. According to what we call the underestimation hypothesis (H1), people will predict that other people are less dishonest than they truly are. This finding would be consistent with research from other life domains on optimism (Sharot et al., 2007) and wishful thinking (Babad, 1987; Babad & Katz, 1991), in which people tend to form rosy beliefs about other people and the future, including "defaulting" to the truth when judging information from others (Levine, 2014, 2022), that systematically exceeds empirical reality. According to what we call the overestimation hypothesis (H2), the exact opposite will happen: On average, people will predict that other people are more dishonest than they truly are. If true, this finding would be consistent with the moral forecasting error documented in self-prediction (Teper et al., 2011), with the crucial extension that people might be making a similar mistake when predicting the aggregate behavior of *other people*.

4. The current investigation

A study providing a careful examination of the accuracy or potential bias in moral beliefs, should measure both the beliefs about dishonesty and actual dishonesty rates as similarly as possible. Ideally, in a controlled test environment where miscalibration does not stem from differences in the prediction context from the behavior context, or other confounds that can make people wrongly estimate how concerned others would be about external factors like detection, punishment, and reputation, when these factors are in fact irrelevant in the specific behavior under study. This is important because many forms of dishonest behavior, such as tax evasion, insurance fraud, or retail theft, go undetected (Amankwah & Van Schoubroeck, 2022; Guyton et al., 2021; Harris, 2008), which makes it important to study beliefs about others' internal motivation to be truthful when it would pay to lie. Second, research comparing beliefs about dishonesty with actual dishonesty should test the robustness of this comparison across different study contexts. Again, this is important because of the wide range of dishonest behaviors one can commit, making it an additional contribution to examine the generalizability of (any) miscalibrated beliefs about others' dishonesty, rather than relying on one specific context like returning wallets (Cohn et al., 2019) or overreporting incentivized performance on a math task (Teper et al., 2011).

In trying to overcome these two challenges, we have conducted an exhaustive internal meta-analysis (k = 31, N = 8,127) of our effect of interest, comparing beliefs about others' dishonest behavior with actual rates of dishonest behavior in the exact same situation. These effects come from *all* experiments conducted as part of a research program on moral decision-making spanning three years (2022-24), which were initially designed to test a broad collection of different hypotheses about the drivers and mechanisms of dishonesty. This approach provides complete access to all relevant data where average beliefs about others' dishonesty and rates of actual dishonesty could be compared. Importantly, beyond enabling us to include effects from a range of contexts of dishonest behavior (e.g., mind game, returns fraud, underbilling), the current meta-analysis is also free from the validity threat arising from the systematic omission of null results in the published literature (Sterling, 1959; Sterling et al., 1995) and unobservable "researcher degrees of freedom" in conducting analyses per study (Simmons et al., 2011). Further, the internal validity of our meta-analysis is bolstered by the fact that we included all relevant studies *and* conducted exactly one type of analysis (Pennycook & Rand, 2022; Vosgerau et al., 2019), comparing beliefs about others' dishonest behavior to rates of dishonest behavior. To allay potential method confounds in our meta-analytic estimate of accuracy vs. potential bias in dishonesty beliefs (such as task-belief order and/or participant's previous behavior), we also conducted four controlled experiments as follow-ups. In those experiments, we used between-subject designs to assess these concerns empirically, by including conditions where participants reported their beliefs about others *before* completing the task themselves, without them being informed that they would have a chance to participate in the same task afterward. In the sections that follow, we describe our meta-analysis and four follow-up experiments, which provided consistent results.

3. Meta-Analysis

3.1. Method

We first present internal meta-analytic evidence comparing beliefs about the rate of others' dishonest behavior to the corresponding rate of in-sample dishonest behavior. There were four criteria for an effect to be included: (1) The study conducted by the authors for a project part in a research program on moral decision-making (e.g., dishonesty against big vs small businesses), (2) The study contained a general measure about beliefs about the rate of dishonesty of other participants completing the survey (e.g., "What percentage of participants do you think, completing the same study as you, would falsely state (intentionally) that they guessed correctly?"), (3) The data from the study could estimate the rate of dishonest behavior (e.g., percentage false reporting for bonus), and (4) the smallest cell size in the study was equal to or greater than 100 participants. This led to the inclusion of all the conditions of 14 separate studies between 2022 and 2024, comprising 31 unique effects. Please see Table 1 for a brief description of the effects. All studies from which effects were extracted were conducted by recruiting U.S. American participants from Prolific (www.prolific.com), a reputable online participant recruitment platform that generates high-quality data (Peer et al., 2021).

Now, we turn to outlining how the data for the meta-analysis was prepared. For each effect, we computed the mean and standard deviations of participants' beliefs on what *percentage* of others, from 0 to 100[14], they thought would behave dishonestly. Then, we conducted 31 one-sample t-tests for each separate effect by testing the mean reported beliefs about others against the target value of the in-sample proportion of dishonest participants the beliefs were reported about. From there, we calculated Cohen's d and subsequently Hedge's g with corresponding standard errors for each effect.

[14] Beliefs regarding effect sizes 28 and 29 were reported from 0 to 83 as we explicitly mentioned that 17% of participants will be statistically able to attain additional bonuses by reporting a correct guess of a future die roll *without* having to lie.

Of the 14 separate studies, 9 measured actual dishonesty behavior using adaptations of the mind game paradigm (Fischbacher & Föllmi-Heusi, 2013; Gneezy & Kajackaite, 2020; Jiang, 2013) which enabled participants to cheat anonymously, and the color game (Albertazzi, 2022) which allowed for individual-level dishonest behavior detection. Each study had between-participant designs which enabled us to estimate the percentage of dishonestly behaving participants against a particular target, and beliefs about that percentage. Two studies had repeated measures, in the sense that participants were asked about their beliefs about others on two rounds of an honest task. Nonetheless, we considered each belief measure across rounds and targets *separately*.

In two of the studies, we asked participants about whether they intended to behave honestly or dishonesty in marketplace scenarios (merchandise returns fraud and not reporting being underbilled). Then, we asked what percentage of others would intend to behave dishonestly in these scenarios. This enabled us to compare how accurate people's beliefs are about others' moral intentions in the marketplace. Here too, we considered each belief measure regarding the rate of dishonest behavior against each target separately, and computed the beliefs-reality difference measure by subtracting the percentage of dishonest choices from the participants' beliefs about the percentage of dishonest choices.

Table 1. Description of the experiments included in the meta-analysis.

Effect ID	Date	Cell size	What the study was originally designed to examine	Dishonesty measure	Dishonesty belief measure prompt	Sample Mean age (SD)	Belief minus reality Mean (SD)
1	May 2022	195	Differences in dishonest behavior against a big	Mind Game	What percentage of participants do you think would state	27.4 (4.7)	10.7 (23.6)
2		190				27.5 (4.9)	17.3 (25.8)

3		195	vs. small-signaling business.		that they guessed correctly?	27.4 (4.7)	14.0 (27.3)
4		190				27.5 (4.9)	20.1 (25.0)
5	July 2022	199	Differences in dishonest behavior against a participant with a male vs. female avatar.	Color Game	As a potential bonus, what percentage of participants do you think will get A. £5 + £2 B. £5 + £5	35.1 (13.1)	14.0 (24.3)
6		199				35.1 (13.1)	14.2 (25.9)
7		202				35.2 (13.4)	9.3 (25.3)
8		202				35.2 (13.4)	25.3 (24.3)
9	August 2022	300	Differences in dishonest behavior against a big vs. small-signaling business.	Mind Game	What percentage of participants do you think would state that they guessed correctly?	27.5 (4.8)	12.9 (22.3)
10		288				27.7 (4.7)	10.1 (22.3)
11	August 2022	308		Mind Game	What percentage of participants do you think would state that they guessed correctly?	38.5 (13.9)	8.6 (21.7)
12		283				38.3 (13.8)	6.1 (20.6)
13	November 2022	409	Differences in intention to commit “returns fraud” when paired with different types of chatbot avatars	Returns Fraud	Out of 100 participants, how many would you guess will choose a free return option with X? *10 participants with the closest guesses will get a £2 bonus.*	41.6 (14.6)	16.5 (21.6)
14		418				39.1 (13.8)	11.6 (21.6)
15		414				39.7 (14.0)	20.3 (23.3)
16		364				39.6 (13.6)	12.9 (21.4)
17	July 2023	291	Differences in intention to commit self-checkout fraud	Self-checkout	Imagine that 100 participants are taking this survey and are faced with	38.8 (14.7)	9.9 (21.1)
18		297				38.0 (14.7)	-3.0 (21.1)

19		294	against big vs. small vs. medium vs. no-size signaling business		the same decision. How many of them do you think would choose to leave X Mart without reporting the billing error?	36.1 (13.1)	-0.5 (21.6)
20		296				38.1 (14.4)	-1.0 (21.6)
21	March 2023	299	Differences in beliefs about others' dishonest behavior when asked before vs. after completing a dishonesty task.	Mind Game	Out of 100 participants completing this task, how many do you think will report they guessed the same score as the die roll? Five participants with the closest guesses will receive an additional £5 within a week.	38.0 (13.0)	21.4 (23.3)
22		296				36.3 (12.2)	17.6 (25.1)
23	April 2023	205	Whether informed about erroneous beliefs about others' dishonest behavior, in various ways, influences subsequent dishonest behavior	Mind Game	We asked several American participants on Prolific to participate in a task last week…… What percentage of participants do you think reported they had guessed the same score as the die roll? Please enter a number only from 0 to 100, no text.	34.4 (11.5)	21.6 (23.7)
24		203				36.7 (12.9)	20.4 (26.2)
25		207				35.2 (11.8)	14.2 (25.6)

					Five participants with the closest estimates will receive an additional £5 within a week.		
26	April 2023	301	Whether informed about erroneous beliefs about others' dishonest behavior, in various ways, influences subsequent dishonest behavior	Mind Game	We asked several American participants on Prolific to participate in a task earlier this month…. What percentage of participants do you think reported they had guessed the same score as the die roll? Please enter a number only from 0 to 100, no text. Five participants with the closest estimates will receive an additional £5 within a week.	35.1 (12.1)	18.7 (25.8)
27		298				34.7 (12.2)	17.3 (26.5)
28	September 2023	292	Whether discrepancy between beliefs about and reality of dishonesty persist when explicitly	Mind Game	What percentage of participants do you think will falsely report they guessed the same score as the die roll? Please insert number only.	37.7 (12.)	26.2 (24.4)
29		291				38.1 (13.3)	24.2 (24.9)

			mentioning others' dishonest behavior in the instructions		Five participants with the closest estimates will receive an additional £5 within a week.		
30	Febraury 2024	100	Whether people are more likely to cheat their outgroup than their ingroup in artificial groups induced using a minimal group creation.	Mind game	What percentage of participants do you think, completing the same study as you, would **falsely state (intentionally) that they guessed correctly**? That is, among all participants who were not able to correctly guess whether it would be an odd or even number, what percentage of these participants do you think would falsely claim that they guessed correctly? Two participants with the closest guess will receive **$1** as additional bonuses within a week of completion of the survey.	38.7 (12.9)	11.3 (22.2)
31		101				38.6 (12.6)	15.2 (25.3)

3.2. Results

We first examine the overall effect sizes of how different the reported beliefs about others' dishonesty were from actual proportions. Using each effect size in Hedge's g as the outcome and the ID of the effect as the study label, we conducted a meta-analysis using the Meta-Analysis Module (MAJOR) in the open-source statistical software Jamovi (The jamovi project, 2022). A random effects model ($k = 31$) showed that the meta-analytic estimate was statistically significant ($b = .58$, $SE = .05$, $Z = 10.9$, $p < .001$, 95% CI [.48, .68]. This suggests that participants overestimated what percentage others were dishonest by 58% of a standard deviation, and more specifically, an overestimation by 13.9 (weighted average) percentage points.

Finding significant heterogeneity in effect sizes across studies, Cochran's Q (30) = 626.9, I^2 = 94.8%, $p < .001$, we conducted a series of moderator analyses.

We intuited that people's beliefs about marketplace dishonesty may be closer to reality than those about dishonesty in economic games. This is because people may be more likely to have information about others committing returns or self-checkout frauds- naturally occurring in the real world- than information about abstract decision-making paradigms from experimental economics. Accordingly, with the addition of a moderator regarding what context beliefs about others were reported (0 = economic game, 1 = marketplace behavior), a random-effect meta-analysis showed that the meta-analytic effect's estimated intercept ($b = .70$, $SE = .05$, $Z = 15.1$, $p < .001$, 95% CI [.61, .79]) was qualified by a statistically significant and negative moderator ($b = -.45$, $SE = .09$, $Z = -5.1$, $p < .001$, 95% CI [-.63, -.28]. This suggests that although people overestimate the percentage of dishonest others overall, for contexts where participants have arguably less experience, they systematically resort to overestimating the percentage of dishonest others.

We also conducted two separate moderated meta-analyses to check for robustness. First, with the addition of a moderator regarding whether dishonest behavior was detectable at the individual level (0 = no, 1 = yes), a random-effect meta-analysis showed that the meta-analytic effect's estimated intercept ($b = .71$, $SE = .06$, $Z = 12.1$, $p < .001$, 95% CI [.59, .82]) was qualified by a statistically significant negative moderator ($b = -.32$, $SE = .09$, $Z = -3.5$, $p < .001$, 95% CI [-.51, -.14]. This suggests the face validity of our results as the overestimation of the percentage of dishonest others decreased for contexts where dishonest actions could be detected.

Second, with the addition of a moderator regarding whether the study was conducted to *specifically* examine differences between beliefs and reality, that is whether the study was conducted specifically for this paper, (0 = no, 1 = yes), A random-effect meta-analysis showed that the meta-analytic effect's estimated intercept ($b = .49$, $SE = .06$, $Z = 8.7$, $p < .001$, 95% CI [.38, .59]) was qualified by a statistically significant and positive moderator ($b = .32$, $SE = .10$, $Z = 3.1$, $p = .002$, 95% CI [.12, .53]. This suggests that the overestimation of the percentage of dishonest others *increased* for contexts when the victim was not specified. This matches intuition because, for most acts of dishonesty, the perpetrator usually has some idea of who the victim is, and so because of greater commonality with real-world dishonest behavior, people's beliefs about dishonesty are likely to be less miscalibrated.

4. Experiment 1. Decision-then-beliefs vs. Beliefs-then-decision

In many of the studies included in the meta-analysis (Effect IDs 1-20, 30-31, participants' prior decisions could have influenced their subsequently reported beliefs about others. For instance, those who have *already* made a dishonest (honest) decision may have reported exaggerated (understated) beliefs. Because belief measures always come after the behavioral measure, we cannot pare out the task effects from participants' beliefs about others.

So, in Experiment 1 (preregistration here), designed specifically to compare estimated and actual rates of dishonesty, participants were randomly assigned to one of two conditions. In one condition, participants completed an honesty task, and then reported their beliefs about others, similar to other studies in the meta-analysis. In the other condition, importantly, participants were presented with the same information about the task and asked to express their beliefs about others who would be completing the task, before they completed the same task themselves. In this experiment, the first condition serves as the ground truth for actual dishonest behavior in a population. The second condition serves as a clean measure of beliefs because participants did not behave honestly or dishonestly beforehand.

We chose the mind-game paradigm (Fischbacher & Föllmi-Heusi, 2013; Gneezy & Kajackaite, 2020; Jiang, 2013) as the measure of behavioral dishonesty for two reasons. First, this task creates a clear tradeoff between economic incentives and moral costs on the decision-maker's side. This is because participants guess in their minds, and then self-report where they are incentivized to lie to increase their potential payoffs. Therefore, our dishonesty measure enables examination of people's beliefs about others' dishonesty in the *absence* of punishment, detection, and reputation concerns, and how that compares with actual dishonest behavior in the same risk-free conditions. So, the results would really represent how people think of the prevalence of dishonest behavior in society should there be no material carrots or sticks to behave honestly.

Second, because dishonesty directly affects the experimenter and not the participants who report their beliefs, self-preservation or wishful thinking motives should not affect the beliefs they report. Because participants are being asked about the dishonesty of others *not* against them but a third-party entity, they would have no incentive to be over/under-cautious.

4.1. Method

603 U.S. American participants were recruited from Prolific to participate in a study for a participation fee. After participants provided informed consent, they were presented with a simple attention check where they were asked to indicate their agreement with the statement “I swim across the Atlantic Ocean to get to work every day.”. 8 participants selected “Agree”, thereby failing the attention check, and so their response was excluded from analyses, leaving us with a final dataset of 595 participants (M_age = 37.1, SD = 12.6; 293 Male, 288 Female, 14 Other).

Then, participants were randomly assigned to either the “decision first, beliefs second” or the “beliefs first, decision second” condition. The “decision first, beliefs second” condition closely followed the studies reported in the meta-analysis. Participants were presented with information about a “Roll a Die Task” where they can increase their payoff by up to £6 by correctly guessing the result of a future die-roll. Participants were first asked to guess the score (an integer from 1 to 6) in their minds, observe and confirm the result, and then report if the score they guessed was the same or different from the result. Then, participants were asked, “Out of 100 participants completing this task, how many do you think will report they guessed the same score as the die roll?”. This was incentivized by telling participants that those with the closest guesses would receive an additional £5 within a week.

In the “beliefs first, decision second” condition, participants were told that several other participants on Prolific would participate in a Roll a Die Task. Participants were told to carefully read the exact same information that was shown to the other participants. Same as the other condition, participants were asked, “Out of 100 participants completing this task, how many do you think will report they guessed the same score as the die roll?”. The crucial difference here is that the participants in this condition did not make any decision before reporting their beliefs. To incentivize accuracy, we explicitly mentioned that “Five participants with the closest guesses will receive an additional £5 within a week”.

Having reported their beliefs, participants were then asked to participate in the same Roll a Die Task with the same incentives. At the end of the study, all participants responded to several measures, such as an 11-item measure of cynicism, and demographic questions including age, gender, political orientation, education, and income.

4.2. Results

First, we calculated the proportions of reported correct guesses in each condition: 53.0% in the “decision first, beliefs second” condition, and 45.2% in the “beliefs first, decision second” condition. Both were clearly above the statistical expectation of 16.7% (1 in 6 chance) for correctly guessing the score of a future die roll (p’s $< .001$).

Our main outcome variable of interest is the difference between participants’ beliefs about what percentage of others would report a correct guess, and the actual percentage of reported correct guesses. This served as a proxy for comparing belief versus reality of others’ dishonesty. To test our hypothesis, we used a one-sample t-test with reported beliefs in the beliefs first roll second condition ($M = 62.7$, $SD = 25.4$) as the outcome, against the test value of 45.2 (the proportion of reported correct guesses in the roll first beliefs second condition). Indeed, the 17.6 percentage point difference between average dishonesty beliefs was statistically significant, t (295) = 11.9, $p < .001$, and corresponded to a medium-to-large effect size, Hedge’s $g = .69$, 95% CI [.57, .82].

5. Experiments 2a-b: Beliefs about others’ dishonesty from last week and last month

Recruiting U.S. American participants from Prolific, Experiments 2a (N = 610; M_age = 35.4, SD = 12.1, 302 Male, 297 Female, 17 Other; registration here) and 2b (N = 599; M_age = 34.9, SD = 12.1; 299 Male, 287 Female, 13 Other; preregistration here) were conceptual replications, with participants recruited using Prolific as well. We wanted to explore if overestimation in beliefs was robust when using slightly different instructions.[15] Taking the proportion of reported correct guesses in Experiment 1 as the point of comparison, we asked participants to predict the percentage of dishonest others from last week (Experiment 2a) and last month (Experiment 2b).

5.1. Method

At the beginning of Experiment 2a, participants were told "We asked several American participants on Prolific to participate in a task last week. On the next page, you will see the exact instructions that those participants saw." and then shown the exact same instructions seen by participants in the beliefs first, decision second condition of Experiment 2.

Similarly, at the beginning of Experiment 2b, participants were told "We asked several American participants on Prolific to participate in a task earlier this month. On the next page, you will see the exact instructions that those participants saw." shown instructions in the same way as Experiment 2a.

In both Experiments 2a and 2b, participants were also asked to respond to a comprehension check regarding the bonus mechanism. Then most importantly, participants

[15] Experiments 2a-b also had a secondary purpose exploring if communicating differences in beliefs vs. reality can reduce subsequent dishonest behavior in the same mind-game paradigm. Although not relevant to the current investigation, none of the formulations of communicating the belief vs. reality difference significantly reduced dishonesty.

were asked to enter a number to the question, “What percentage of participants do you think reported they had guessed the same score as the die roll?”. The belief elicitations were incentivized in the same way as Experiment 2, disclosing that five participants with the closest estimates would receive an additional £5.

5.2 Results

A one-sample t-test showed that mean beliefs about what percentage of others would report a correct guess in Experiment 2a were (M = 63.9, SD = 25.3) far exceeded the actual proportion of 45.2 (from Experiment 2), t (614) = 18.4, p < .001, g = .74, 95% CI [.65, .83] when asked about participants from last week. Similarly, beliefs about others also exceeded reality in Experiment 2b (M = 63.2, SD 26.1), t (598) = 16.9, p < .001, g = .69, 95% CI [.60, .78] when asked about participants from last month. This shows that even when asking participants about beliefs about others’ dishonesty across time periods, the overestimation is robust.

6. Experiment 3. Beliefs about explicitly described dishonest behavior

Although we were able to show that exaggerated beliefs about others’ dishonesty persisted regardless of the order in which the die roll task and belief report were completed (Experiment 1) and making predictions across time frames (Experiments 2a-b), it can still be argued that we did not *explicitly* ask the participants about the dishonest behavior of others[16]. Therefore, we may have elicited participants’ beliefs about others’ dishonest behavior in the clearest possible way. So, in Experiment 3, we not only explicitly told participants about how anyone can behave dishonestly *without* detection in the mind-game task, but also asked them about their beliefs about what *exact* percentage they think will behave dishonestly.

---

[16] Effect IDs 30 and 31 come from studies where participants were explicitly asked to estimate the percentage of dishonest participants in the task.

6.1 Method

583 participants based in the U.S.A. were recruited using Prolific to participate in the study for a participation fee (preregistration here). After removing duplicate IDs, the final dataset had 571 responses (M_age = 37.9, SD = 12.9; 285 Male, 287 Female, 11 Other). Similar to Experiment 1, participants were randomly assigned to either the "decision first, beliefs second" or the "beliefs first, decision second" condition. However, there were two crucial additions.

First, before asking participants about their beliefs about others, we asked participants to read carefully a section that described in detail how statistically 1 in 6 participants (17%) would correctly guess the score of a six-sided die-roll; but because everyone guesses in their minds and then self-reports, some may falsely report that they have guessed the same number. We further illustrated that we will never know who specifically falsely reported because only the participant knows what is on their mind.

Further, we showed participants in a table what different percentages of reported correct guesses can suggest in terms of percentages of participants falsely reporting (e.g., 17% = 0% false reported, 37% = 20% falsely reported, …., 100 % = 83 % falsely reported). Then, we presented participants with a comprehension check asking, "If X % of participants completing this task report that they guessed the same, what percentage of participants do you think falsely reported having guessed the same?" with response options[17] of "X % minus 17%" and "X %".

Second, participants were explicitly asked, "What percentage of participants do you think will falsely report they guessed the same score as the die roll? Please insert number

[17] The results did not meaningfully change when excluding the participants who failed the check.

only." in an open response form. Note that we specifically mention the false reporting, rather than indirectly asking what percentage will state they guessed correctly. Similar to Experiments 1, 2a, and 2b, we incentivized responses by disclosing that five participants with the closest guesses would receive an additional £5 bonus within a week. Additionally, on the next page, participants were also asked about how confident they were about their estimate on a slider scale from 0 (Not at all confident) to 100 (Completely confident). The two differences were presented in the exact same way in both conditions of "decision first, beliefs second" and "belief first, decision second".

The main mind-game/die-rolling task was the same as in Experiments 1, 2a, and 2b. Participants guessed in their minds the outcome of a future die roll, observed the score, and then self-reported if their guess was the same as the roll. Reporting their guess as the same as the score allows them to be considered for a random lottery where they can receive an additional payoff corresponding to the score.

At the end, participants responded to demographic questions regarding age, gender, political beliefs, and lastly 5-item exploratory measures for projection, negativity, and availability biases.

6.2. Results

Similar to Experiment 1, we first calculated the proportions of reported correct guesses in each condition: 41.8% in the "roll first beliefs second" condition, and 43.6% in the "beliefs first and roll second" condition. Both were clearly above the statistical expectation of 16.7% (1 in 6 chance) for correctly guessing the score of a future die roll (p's $< .001$), suggesting the proportions of dishonest participants were 25.1% and 26.9% respectively.

Our main statistic of interest is how far reported beliefs about others' dishonesty are from the actual rate of dishonest behavior. So, we used a one-sample t-test with beliefs about the percentage of others falsely reporting in the "beliefs first roll second" condition (M = 49.3%, SD = 24.9) as the outcome, against the test value of 25.1[18] (the inferred proportion of dishonesty). Indeed, the 24.2 percentage point difference between average dishonesty beliefs and in-sample dishonesty was statistically significant, t (290) = 16.6, $p < .001$, $g = .97$, 95% CI [.83, 1.11].

7. General Discussion

The current research examines a fundamental question: Do people believe others are similarly, more, or less dishonest than they actually are? The results from our meta-analysis reveal a significant *overestimation* of the actual dishonesty of other people, by an average of 14 percentage points (meta-analytic effect: Hedge's g = .0.58), when comparing aggregate beliefs about dishonesty with actual dishonesty in the same test situation. This finding provides robust support to the overestimation hypothesis, and goes against the underestimation hypothesis. Across our specific experiments, the results showed that the overestimation effect persisted across belief-decision order, time of behavior, and variations in instructions.

Given how dishonesty in the lab also predicts real-world behavior outside the lab (Dai et al., 2018), the current meta-analysis and follow-up experiments suggest that people may wrongly believe that the world is more dishonest than it truly is. This systematic overestimation has important implications for social cognition in the moral domain and practical consequences for organizational practices, market dynamics, and interpersonal relations.

[18] We subtracted 17% to calculate the proportion of dishonest participants in this study as per our simplified instructions.

Challenging the applicability of the 'wisdom of crowds' approach (Galton, 1907; Surowiecki, 2005; Van Dolder & Van Den Assem, 2017) to moral judgments, our study reveals that collective estimates regarding dishonest behavior are systematically biased toward pessimism. To that end, our findings contribute to the negativity bias literature (Baumeister et al., 2001; Rozin & Royzman, 2001) by demonstrating its significant role in shaping social perceptions of dishonesty. That is, individuals might prioritize negative information when forming beliefs about dishonest behaviors of others *even* when it does not affect them.

Although we found that people who were more likely to have been dishonest were also more likely to overestimate the dishonest behaviors of others, people who were more likely to have been honest also overestimated, albeit to a lower extent. This suggests that social projection (Krueger, 2007) is partially a factor in driving how people think about others behaving dishonestly, but the negative bias in beliefs is pervasive throughout the population.

Another theoretical contribution our results can make is speaking to the mixed effects of social norm nudges in reducing tax evasion (John & Blume, 2018) and insurance fraud (Martuza et al., 2022). It may be so that people are quite entrenched in their beliefs in overestimating others' dishonest behavior, which can lead to people simply not believing descriptive social norm messaging such as "90% of your peers report correct information when filling forms".

Further, our finding, that people overestimate in the context of economic games than marketplace scenarios, implies that under conditions of lower information, people are *systematically* more likely to err on the pessimistic side regarding how dishonest others would behave. This goes to show that the notion of loss aversion (Kahneman & Tversky,

1979) not only applies to judgments under uncertainty where it affects the self, but also to general social perceptions.

Practically, the systematic overestimation of others' dishonesty signals the need for trust-building interventions, such as information campaigns or educational programs, aimed at correcting these misperceptions to foster a more trustful society. For example, more and more supermarkets are putting (even) low-value products such as deodorant, toothpaste, and soap behind lock and key (Meyersohn, 2022). In the workplace, a surge in monitoring employee activities may be fueling worker distrust (Christian, 2022).

Based on our finding that people overestimate the dishonesty of others pervasively, organizations may benefit from reconsidering policies enacted from mistrust and excessive surveillance, for example for remote/hybrid work and expense accounts. By adopting a more balanced approach that emphasizes trust, decision-makers could enhance employee morale and productivity without compromising integrity.

Broadly, firms may benefit from scaling back their security measures, especially in today's increasingly common self-service environments, where surveillance can reduce the gains from self-service customer journeys. In peer-to-peer platforms, communicating that other users are more honest and trustworthy than customers may think can help in designing more trust-oriented and richer experiences in the sharing economy. Lastly, for public awareness, addressing the misperception of societal moral decline could foster greater social cohesion and encourage more trusting behavior.

Several limitations must be acknowledged in our interpretations, which also creates a potential for interesting subsequent research. First, given the U.S. American participant pool in all our studies, the generalizability of our findings across cultures remains an open question. Second, experimental honesty tasks, while illustrative of the economic gain versus moral cost tradeoffs in their basic form, may not fully capture the complexity of real-world dishonesty, where factors such as stakes, punishment, and reputational concerns play significant roles. So future research may benefit from examining how estimates and actual rates of dishonesty compare in settings that may mirror real-world decisions to a greater extent.

Third, our reliance on mostly one-shot reports of beliefs precludes the analysis of within-individual variations over time, which could offer further insights into the dynamics of belief formation. So, future work may benefit from Fourth, although we robustly document the presence of an important phenomenon- overestimation of others’ dishonesty- we did not investigate what drives this overestimation. As a next step, future research may benefit from exploring the mechanisms underlying the effect.

Nonetheless, our investigation into the accuracy of beliefs about others' dishonesty reveals a consistent overestimation, with important theoretical and practical implications. By deepening our understanding of the psychological underpinnings of these beliefs, we can begin to address the societal consequences of such misperceptions, fostering environments where trust and integrity are valued and nurtured.

References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for Truth-Telling. *Econometrica*, *87*(4), 1115–1153. https://doi.org/10.3982/ECTA14673

Albertazzi, A. (2022). Individual cheating in the lab: A new measure and external validity. *Theory and Decision*, *93*(1), 37–67. https://doi.org/10.1007/s11238-021-09841-0

Algan, Y., & Cahuc, P. (2013). Trust and Growth. *Annual Review of Economics*, *5*(1), 521–549. https://doi.org/10.1146/annurev-economics-081412-102108

Amankwah, J., & Van Schoubroeck, C. (2022). Fraud detection in motor insurance: Privacy and data protection concerns under EU Law. *International Data Privacy Law*, *12*(3), 220–238. https://doi.org/10.1093/idpl/ipac009

Ayal, S., Celse, J., & Hochman, G. (2019). Crafting messages to fight dishonesty: A field investigation of the effects of social norms and watching eye cues on fare evasion. *Organizational Behavior and Human Decision Processes*, S0749597818305004. https://doi.org/10.1016/j.obhdp.2019.10.003

Babad, E. (1987). Wishful thinking and objectivity among sports fans. *Social Behaviour*, *2*(4), 231–240.

Babad, E., & Katz, Y. (1991). Wishful Thinking—Against All Odds. *Journal of Applied Social Psychology*, *21*(23), 1921–1938. https://doi.org/10.1111/j.1559-1816.1991.tb00514.x

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is Stronger than Good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. In N. G. Fielding, A. Clarke, & R. Witt (Eds.), *The Economic Dimensions of Crime* (pp. 13–68). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-62853-7_2

Bicchieri, C., Dimant, E., & Sonderegger, S. (2023). It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion. *Games and Economic Behavior*, *138*, 321–354. https://doi.org/10.1016/j.geb.2023.01.005

Bott, K. M., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2020). You've Got Mail: A Randomized Field Experiment on Tax Evasion. *Management Science*, *66*(7), 2801–2819. https://doi.org/10.1287/mnsc.2019.3390

Christian, A. (2022). The employee surveillance that fuels worker distrust. *BBC Worklife*. https://www.bbc.com/worklife/article/20220621-the-employee-surveillance-that-fuels-worker-distrust

Cohn, A., Gesche, T., & Maréchal, M. A. (2022). Honesty in the Digital Age. *Management Science*, *68*(2), 827–845. https://doi.org/10.1287/mnsc.2021.3985

Cohn, A., Maréchal, M. A., Tannenbaum, D., & Zünd, C. L. (2019). Civic honesty around the globe. *Science*, *365*(6448), 70–73. https://doi.org/10.1126/science.aau8712

Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional Belief. *Annual Review of Psychology*, *62*(1), 271–298. https://doi.org/10.1146/annurev.psych.121208.131622

Cusimano, C., & Goodwin, G. P. (2020). People judge others to have more voluntary control over beliefs than they themselves do. *Journal of Personality and Social Psychology*, *119*(5), 999–1029. https://doi.org/10.1037/pspa0000198

Dai, Z., Galeotti, F., & Villeval, M. C. (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Management Science*, *64*(3), 1081–1100. https://doi.org/10.1287/mnsc.2016.2616

Dias, R., Sharma, E., & Fitzsimons, G. J. (2023). *Consumer Wealth and Price Expectations*. https://marketing.wharton.upenn.edu/wp-content/uploads/2023/09/10.19.2023-Dias-Rodrigo-PAPER-PriceExpectations-UnderReviewJCR.pdf

Dimant, E., & Gesche, T. (2023). Nudging enforcers: How norm perceptions and motives for lying shape sanctions. *PNAS Nexus*, *2*(7), pgad224. https://doi.org/10.1093/pnasnexus/pgad224

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

Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, *531*(7595), 496–499. https://doi.org/10.1038/nature17160

Galeotti, F., Maggian, V., & Villeval, M. C. (2021). Fraud Deterrence Institutions Reduce Intrinsic Honesty. *The Economic Journal*, *131*(638), 2508–2528. https://doi.org/10.1093/ej/ueab018

Galton, F. (1907). Vox Populi. *Nature*, *75*(1949), 450–451. https://doi.org/10.1038/075450a0

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, *145*(1), 1–44. https://doi.org/10.1037/bul0000174

Gneezy, U., & Kajackaite, A. (2020). Externalities, stakes, and lying. *Journal of Economic Behavior & Organization*, *178*, 629–643. https://doi.org/10.1016/j.jebo.2020.08.020

Greenberg, J. (2002). Who stole the money, and when? Individual and situational determinants of employee theft. *Organizational Behavior and Human Decision Processes*, *89*(1), 985–1003. https://doi.org/10.1016/S0749-5978(02)00039-0

Guyton, J., Langetieg, P., Reck, D., Risch, M., & Zucman, G. (2021). *Tax Evasion at the Top of the Income Distribution: Theory and Evidence* (w28542; p. w28542). National Bureau of Economic Research. https://doi.org/10.3386/w28542

Harris, L. (2008). Fraudulent Return Proclivity: An Empirical Analysis. *Journal of Retailing*, *84*(4), 461–476. https://doi.org/10.1016/j.jretai.2008.09.003

Herreros, F. (2023). The State and Trust. *Annual Review of Political Science*, *26*(1), 117–134. https://doi.org/10.1146/annurev-polisci-051921-102842

Hertwig, R., & Mazar, N. (2022). Toward a taxonomy and review of honesty interventions. *Current Opinion in Psychology*, *47*, 101410. https://doi.org/10.1016/j.copsyc.2022.101410

Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology*, *20*(4), 311–335. https://doi.org/10.1002/ejsp.2420200404

Internal Revenue Service. (2022). *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2014-2016*. Research, Applied Analytics & Statistics. https://www.irs.gov/pub/irs-pdf/p1415.pdf

Jervis, R. (2006). Understanding Beliefs. *Political Psychology*, *27*(5), 641–663. https://doi.org/10.1111/j.1467-9221.2006.00527.x

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, *93*, 328–336. https://doi.org/10.1016/j.jebo.2013.04.003

John, P., & Blume, T. (2018). How best to nudge taxpayers? The impact of message simplification and descriptive social norms on payment rates in a central London local authority. *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.10

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Kluegel, J. R., & Smith, E. R. (1981). Beliefs about Stratification. *Annual Review of Sociology*, *7*(1), 29–56. https://doi.org/10.1146/annurev.so.07.080181.000333

Köbis, N. C., Soraperra, I., & Shalvi, S. (2021). The Consequences of Participating in the Sharing Economy: A Transparency-Based Sharing Framework. *Journal of Management*, *47*(1), 317–343. https://doi.org/10.1177/0149206320967740

Kocher, M. G., Schudy, S., & Spantig, L. (2018). I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. *Management Science*, *64*(9), 3995–4008. https://doi.org/10.1287/mnsc.2017.2800

Kristal, A. S., Whillans, A. V., Bazerman, M. H., Gino, F., Shu, L. L., Mazar, N., & Ariely, D. (2020). Signing at the beginning versus at the end does not decrease dishonesty. *Proceedings of the National Academy of Sciences*, *117*(13), 7103–7107. https://doi.org/10.1073/pnas.1911695117

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology*, *18*(1), 1–35. https://doi.org/10.1080/10463280701284645

Levine, T. R. (2014). Truth-Default Theory (TDT): A Theory of Human Deception and Deception Detection. *Journal of Language and Social Psychology*, *33*(4), 378–392. https://doi.org/10.1177/0261927X14535916

Levine, T. R. (2022). Truth-default theory and the psychology of lying and deception detection. *Current Opinion in Psychology*, *47*, 101380. https://doi.org/10.1016/j.copsyc.2022.101380

List, J. A., & Momeni, F. (2021). When Corporate Social Responsibility Backfires: Evidence from a Natural Field Experiment. *Management Science*, *67*(1), 8–21. https://doi.org/10.1287/mnsc.2019.3540

Martuza, J. B., Skard, S. R., Løvlie, L., & Thorbjørnsen, H. (2022). Do honesty-nudges really work? A large-scale field experiment in an insurance context. *Journal of Consumer Behaviour*, *21*(4), 927–951. https://doi.org/10.1002/cb.2049

Mastroianni, A. M., & Gilbert, D. T. (2023). The illusion of moral decline. *Nature*, *618*(7966), 782–789. https://doi.org/10.1038/s41586-023-06137-x

Meyersohn, N. (2022). Why Old Spice, Colgate and Dawn are locked up at drug stores. *CNN Business*. https://edition.cnn.com/2022/07/30/business/drug-stores-locked-products/index.html

Molouki, S., & Pronin, E. (2015). Self and other. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology, Volume 1: Attitudes and social cognition.* (pp. 387–414). American Psychological Association. https://doi.org/10.1037/14341-013

National Retail Federation. (2022). *NRF Reports Retail Shrink Nearly a $100B Problem*. https://nrf.com/media-center/press-releases/nrf-reports-retail-shrink-nearly-100b-problem

Nødtvedt, K. B., Sjåstad, H., Skard, S. R., Thorbjørnsen, H., & Van Bavel, J. J. (2021). Racial bias in the sharing economy and the role of trust and self-congruence. *Journal of Experimental Psychology: Applied*. https://doi.org/10.1037/xap0000355

Pascual-Ezama, D., Prelec, D., Muñoz, A., & Gil-Gómez de Liaño, B. (2020). Cheaters, Liars, or Both? A New Classification of Dishonesty Profiles. *Psychological Science*, *31*(9), 1097–1106. https://doi.org/10.1177/0956797620929634

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature Communications*, *13*(1), 2333. https://doi.org/10.1038/s41467-022-30073-5

Pew Research Center. (2023). *In Divided Washington, Americans Have Highly Negative Views of Both Parties' Leaders*. https://www.pewresearch.org/politics/2023/04/07/evaluations-of-the-economy-and-the-state-of-the-nation/#:~:text=While%20Americans'%20views%20of%20current,think%20the%20economy%20will%20improve.

Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3

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

Rotman, J. D., Khamitov, M., & Connors, S. (2018). Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically. *Journal of Consumer Psychology*, *28*(2), 353–361. https://doi.org/10.1002/jcpy.1002

Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2

Schilke, O., Reimann, M., & Cook, K. S. (2021). Trust in Social Relations. *Annual Review of Sociology*, *47*(1), 239–259. https://doi.org/10.1146/annurev-soc-082120-082850

Schotter, A., & Trevino, I. (2014). Belief Elicitation in the Laboratory. *Annual Review of Economics*, *6*(1), 103–128. https://doi.org/10.1146/annurev-economics-080213-040927

Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, *450*(7166), 102–105. https://doi.org/10.1038/nature06280

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Skowronek, S. E. (2022). DENIAL: A Conceptual Framework to Improve Honesty Nudges. *Current Opinion in Psychology*, 101456. https://doi.org/10.1016/j.copsyc.2022.101456

Soraperra, I., Weisel, O., & Ploner, M. (2019). Is the victim *Max (Planck)* or *Moritz* ? How victim type and social value orientation affect dishonest behavior. *Journal of Behavioral Decision Making*, *32*(2), 168–178. https://doi.org/10.1002/bdm.2104

Sterling, T. D. (1959). Publication Decisions and their Possible Effects on Inferences Drawn from Tests of Significance—Or Vice Versa. *Journal of the American Statistical Association*, *54*(285), 30–34. https://doi.org/10.1080/01621459.1959.10501497

Sterling, T. D., Rosenbaum, W. L., & Weinkam, J. J. (1995). Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa. *The American Statistician*, *49*(1), 108–112. https://doi.org/10.1080/00031305.1995.10476125

Sulitzeanu-Kenan, R., Tepe, M., & Yair, O. (2022). Public-Sector Honesty and Corruption: Field Evidence from 40 Countries. *Journal of Public Administration Research and Theory*, *32*(2), 310–325. https://doi.org/10.1093/jopart/muab033

Surowiecki, J. (2005). *The wisdom of crowds* (Nachdr.). Anchor Books.

Tarrant, M., Branscombe, N. R., Warner, R. H., & Weston, D. (2012). Social identity and perceptions of torture: It's moral when we do it. *Journal of Experimental Social Psychology*, *48*(2), 513–518. https://doi.org/10.1016/j.jesp.2011.10.017

Teper, R., Inzlicht, M., & Page-Gould, E. (2011). Are We More Moral Than We Think?: Exploring the Role of Affect in Moral Behavior and Moral Forecasting. *Psychological Science*, *22*(4), 553–558. https://doi.org/10.1177/0956797611402513

The jamovi project. (2022). *Jamovi.* (2.3) [Computer software]. https://www.jamovi.org

Van Dolder, D., & Van Den Assem, M. J. (2017). The wisdom of the inner crowd in three large natural experiments. *Nature Human Behaviour*, *2*(1), 21–26. https://doi.org/10.1038/s41562-017-0247-6

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300. https://doi.org/10.1037/a0017908

Vincent, L. C., Emich, K. J., & Goncalo, J. A. (2013). Stretching the Moral Gray Zone: Positive Affect, Moral Disengagement, and Dishonesty. *Psychological Science*, *24*(4), 595–599. https://doi.org/10.1177/0956797612458806

Vosgerau, J., Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2019). 99% impossible: A valid, or falsifiable, internal meta-analysis. *Journal of Experimental Psychology: General*, *148*(9), 1628–1639. https://doi.org/10.1037/xge0000663

Wang, L., & Murnighan, J. K. (2017). How Much Does Honesty Cost? Small Bonuses Can Motivate Ethical Behavior. *Management Science*, *63*(9), 2903–2914. https://doi.org/10.1287/mnsc.2016.2480

Weber, T. O., Schulz, J. F., Beranek, B., Lambarraa-Lehnhardt, F., & Gächter, S. (2023). The behavioral mechanisms of voluntary cooperation across culturally diverse societies:

Evidence from the US, the UK, Morocco, and Turkey. *Journal of Economic Behavior & Organization*, *215*, 134–152. https://doi.org/10.1016/j.jebo.2023.09.006

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

Chapter 6. Article 5: A Registered Report on Gender Bias in Interpersonal Dishonesty: Are Females and Males Cheated Differently?

PREPRINT

Working paper part of a doctoral dissertation (March 2024). All conclusions are preliminary. Please do not copy, share, or cite without the author's permission.

Jareef Martuza[1]

[1]Department of Strategy and Management, Norwegian School of Economics, Helleveien 30, Bergen 5045, Norway

Author notes

The authors declare to have no known conflicts of interest. Correspondence about this research should be addressed to Jareef Martuza (jareef.martuza@nhh.no) PhD student at the Norwegian School of Economics.

Our studies comply with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. As Norwegian laws and regulations do not require review by an institutional review board for anonymous, non-medical, low-risk research with human participants, this project was not submitted to such review. Informed consent was collected from all participants. I received funding for the experiment from the Distribusjonsøkonomiske forskningsprosjekter (Distribution economics research project) at NHH.

Abstract

This registered report investigated the effect of the target's sex in selfish dishonest behavior. Informed by gender biases and moral typecasting theory, I hypothesized that individuals may have different moral standards when their lying may harm a female target than when a male target. A total of 3,166 male and female participants from nine countries were incentivized to lie and increase personal payoffs at the cost of another same-sex, opposite-sex, or unmentioned-sex participant- with zero risks of detection, punishment, and reputational consequences. Female targets were cheated 22% less than unmentioned-sex targets. There was also a significant target and decision-maker interaction: Female decision-makers cheated female (vs. male) targets 53.6% less, with no such difference among male decision-makers. Several exploratory self-reports suggest people think it is less acceptable to cheat women than men. Together, this research highlights the role of gender in understanding how individual dishonesty may depend on who it affects.

Keywords: Gender bias; dishonesty; moral typecasting.

A Registered Report on Gender Bias in Interpersonal Dishonesty:

Are Females and Males Cheated Differently?

If you were to cheat another person for personal gain- during a one-shot interaction with zero risk of detection, punishment, or reputational consequences- would the target's gender influence your decision? If asked, most individuals would likely assert that they would not cheat *regardless* of the target's gender. Yet, instances of individuals cheating others for personal gain are not uncommon. Ordinary people, not just career criminals, cheat others in negotiations (Kray et al., 2014), marketplace interactions (Ross & Robertson, 2000), and even economic games (Gneezy, 2005). In life, many interpersonal interactions occur in zero-sum contexts where economic incentives and moral standards are at odds (Jacobsen et al., 2018). Individuals may be tempted to lie and gain at the expense of strangers whom the decision-maker may never come across again. Then, does it matter *who* is affected by the decision- the target- for whether they get cheated? In that spirit, I ask two questions: (1) Are individuals more likely to cheat males or females? and (2) Are the effects symmetrical or asymmetrical across male and female decision-makers?

It has been well-documented in psychology and behavioral economics that both economic self-interest and moral standards affect the dishonesty of individuals (Gibson et al., 2013). Indeed, people do not cheat to the maximum when incentivized, basically leaving money on the table even when their decisions are completely anonymous (Gerlach et al., 2019). This supports the reasoning that individuals may have an internal "moral cost" of lying (Kajackaite & Gneezy, 2017), which restrains dishonest pursuits of self-interest, even when there are zero risks of detection, punishment, or reputational consequences (Gneezy et al., 2018; Thielmann & Hilbig, 2019). In other words, this internal force can be viewed as the "moral cost" of dishonesty that promotes honest behavior even when there are no repercussions for dishonest behavior. Recent research hints that this moral cost of dishonesty is not fixed, but may *vary* depending on “at whose cost” the gains of dishonesty are attained. For example, when tempted with identical risk-free gains through dishonest behavior, individuals are more likely to cheat another person than an organization (Soraperra et al., 2019), a small organization than a large organization (Martuza et al., 2023), and a harmless corporation than a harmful corporation (Rotman et al., 2018). These suggest that the mere characteristics of the target, holding all else equal, can play an important role when individuals face trade-offs between economic incentives and moral standards.

In that vein, I argue that the average person may incur different moral costs for potentially cheating *different* individuals as well: being biased in cheating some individuals more than others. Specifically, the current research proposes that the target’s sex (male or female) can affect dishonest behavior against them. Here, dishonest behavior refers to lying to benefit the self at the cost of another entity, also known as “cheating” (Jiang, 2013). On the decision-maker side, a meta-analysis on rates of dishonesty in different experimental paradigms suggests that males are more dishonest than females on average (Gerlach et al., 2019). Yet, it remains unclear how the oft-salient sex of the *target* can influence dishonesty against them. In the social context, whether one’s interaction partner is a male or female is one of the first things an individual notices. Sex is an ever-present category people are socialized to notice from a young age (Eagly, 2013; Perales et al., 2023). So, whether there is a gender bias in interpersonal dishonesty is important to examine because gender prescriptions and proscriptions apply not only to how men and women behave themselves but also to how people behave toward men and women (Kosakowska-Berezecka et al., 2023; Kray et al., 2014; Prentice & Carranza, 2002). Indeed, gender bias is quite pervasive in society (Georgeac et al., 2019), from affecting how people form impressions (Oh et al., 2020) to how they judge music (Boghrati & Berger, 2023). So much so, recent research shows that the target's gender can even shift the altruistic preferences of decision-makers *even* when all other external influences are held constant (FeldmanHall et al., 2016; Graso et al., 2023). So, people may apply different moral standards regarding dishonest behavior toward males and females.

The primary hypothesis proposed a gender bias in interpersonal dishonesty: Decision-makers will be more dishonest against male targets than female targets (Hypothesis 1). That is, in the absence of detection, punishment, and reputational consequences, individuals will be more likely to commit dishonest actions for personal economic gains at the loss of another male than another female person. This is informed by the moral typecasting theory (Gray & Wegner, 2009), which posits that people categorize others as either moral agents or moral patients based on their perceived characteristics. Combined with prevalent gender stereotypes, often conflated with biological sex[19] (Cameron & Stinson, 2019), which portrays men as agentic and women as vulnerable (Bem, 1974), research also suggests evidence for a gender bias in moral typecasting in itself (Reynolds et al., 2020). In this paper, the tendency to perceive men as moral agents who are less deserving of protection from harm, and women as moral patients who require protection (Graso et al., 2023; Reynolds et al., 2020) forms the basis of the prediction that people will cheat males more than females.

The secondary hypothesis proposes that the sex of the decision-maker will moderate the aforementioned effect: Male decision-makers will be more dishonest against male (vs. female) targets than female decision-makers will against other male (vs. female) targets (Hypothesis 2). Specifically, the target's gender bias in cheating would be stronger among

[19] In this research, we examine the role of the target's sex (how males and females are cheated differently) and NOT gender (how men and women are cheated differently). However, it is worth noting that part of previous research on informing our theoretical framework is based on gender research that implied biological male vs. female comparisons. So, when referencing those works, our paper stays consistent in how the original authors framed their comparisons (men vs. women). Moreover, biases toward males and females have been conceptualized as gender bias and not sex bias (e.g., gender bias in moral typecasting (Reynolds et al., 2020)).

male (versus female) decision-makers. Specifically, we propose that male decision-makers will be more likely to cheat other males than females, compared to the extent female decision-makers will cheat other males more than females. This prediction is supported by empirical evidence that suggests males may be particularly more prone to condone the mistreatment of other males (Arnestad et al., 2020; Owuamalam & Matos, 2022). Taken together, the current research extends previous work on gender bias in victimhood (e.g., Graso et al., 2023; Reynolds et al., 2020) to enhance understanding of gender in moral psychology, and can inform policies for a more honest, trustworthy, and less biased society. Formally, we tested the following hypotheses, the theory behind which we detail in the next sections.

H1. Decision-makers will be more dishonest against male targets than female targets.

H2. Male decision-makers will be more dishonest against male (vs. female) targets than female decision-makers will other male (vs. female) targets.

Dishonesty against males and females

The primary hypothesis, that males are more likely to be cheated than females, is informed by combining moral typecasting theory (Gray & Wegner, 2009) and gender stereotypes (Bem, 1974). The influential moral typecasting theory, as proposed by Gray and Wegner (2009), provides a useful framework for understanding how others in social interactions are categorized as moral agents or patients. According to the moral typecasting theory, others classified as agents are considered more as capable performers and less as recipients of actions, whereas those classified as patients are considered more as recipients and less as performers of actions. This categorization of a moral dyad is mutually exclusive, such that in a given interaction, agentic characteristics of an interaction partner can make it easier to see them as a moral agent and less as a moral patient, and vice versa (Gray & Wegner, 2009). In case of targets of moral harm- such as being cheated in an interpersonal interaction- decision-makers may typecast the target as an agentic moral perpetrator or a vulnerable moral victim depending on the characteristics of the target. Because of mutual exclusivity in a moral dyad, agentic perpetrators are perceived to be less capable of feeling pain whereas vulnerable victims are perceived to be more sensitive to feeling pain (Reynolds et al., 2020; Shepherd et al., 2019). This greater vulnerability perception of a target can evoke greater moral concerns (Dijker, 2010; Gray & Wegner, 2009).

Indeed, differences in vulnerability perceptions have been shown to partially explain why big businesses are more likely to be cheated than small businesses (Martuza et al., 2023). Similarly, differences in vulnerability perceptions of males and females make it plausible to expect a gendered bias in interpersonal dishonesty as well. Across history and cultures, gender stereotypes depict men as dominant, independent, and aggressive, whereas women as tender, vulnerable, and pain-sensitive (Bem, 1974). In the moral domain, people are systematically more accepting of harm befalling males than females (Graso et al., 2023), and are less outraged by the same transgression against males than females (Reynolds et al., 2020). Accordingly, many societies have instilled a heuristic that females should be protected from harm (Graso et al., 2023). A meta-analysis shows that females are more likely to receive help than their male counterparts (Eagly & Crowley, 1986). Even in simple allocation choices, the gender of the recipient seems to play a role, with female players being allocated more money than male players in the Dictator game (Saad & Gill, 2001). Contrarily, men are thought to experience less pain (Reynolds et al., 2020), sacrificed more in trolley dilemmas (FeldmanHall et al., 2016), and are blamed more for their academic underperformance (Cappelen et al., 2019). These suggest that males are more likely to be deemed acceptable to bear the costs of negative treatment, which may spill over to dishonesty against them as well.

Because if the average individual is more accepting of (Graso et al., 2023) and less outraged by (Reynolds et al., 2020) the same harm when it befalls males than females, the potential of cheating males may evoke a lesser feeling of wrongness. Much of people's moral decisions are based on what intuitively feels more or less right and wrong (Greene & Haidt, 2002), and cheating females may feel more intuitively wrong than males. In the context of the current research, moral typecasting theory suggests that individuals may be more likely to cheat those they perceive as moral agents (i.e., males) and less likely to cheat those they perceive as moral patients (i.e., females). Specifically, we propose that when faced with the decision to cheat another male or female person, differences in the cognitive ease of seeing males and females as vulnerable victims can lead to systematic differences in cheating preferences against male and female targets.

Counter to the primary hypothesis, there are reasons to expect females to be cheated more than males. Because throughout history, it has been females who have been biased against more- such as being denied the right to vote, own property, and receive higher education. Despite significant progress, bias against females persists across societies and domains, disadvantaging women in hiring processes (Moss-Racusin et al., 2012), job promotions (Régner et al., 2019), and media coverage (Shor et al., 2019). Moreover, research also suggests that negotiators believe women to be easier to mislead than men, leading to them being more likely to use deception toward women in negotiations (Kray et al., 2014). In a related domain, underperforming women are more likely to be given upwardly distorted feedback, and these white lies deprive women of the same quality of feedback as men (Jampol & Zayas, 2021). So, one may counter-argue that these differences in cheating preferences across the target's sex may be due to differences in perceived risks such as repercussions of cheating may differ across male and female targets. However, our proposed experiment allows participants to cheat another male, female, or unmentioned-sex participant without any detection, punishment, or reputational consequences. By making anonymous and risk-free decisions, individuals will face a clear tradeoff between the economic incentives of lying and the motivation to uphold moral standards. This can delineate who would people on average be more likely to cheat- males or females- if the target's sex was the *only* difference, enabling us to examine how the *isolated* moral cost of cheating may be different when potentially cheating males versus females.

The role of the decision-maker's sex

The secondary hypothesis predicts an interaction effect: such that the sex of the decision-maker will moderate the effect of the target's sex on cheating. I hypothesized that male decision-makers will be more dishonest against other males (vs. females) than female decision-makers will other males (vs. females). That is, the gender bias in dishonesty would be stronger in the male decision-maker subgroup than the female decision-maker subgroup of a population. This hypothesized moderation effect is informed by multiple streams of empirical evidence. Research suggests males (vs. females) are more biased in downplaying the harm inflicted on other males (Arnestad et al., 2020; Owuamalam & Matos, 2022), suggesting an asymmetry of the effect across the sex of decision-makers. For example, although female participants judged male-aggressor-female-victim and female-aggressor-male-victim sexual harassment cases as equally bad, whereas male participants judged the male-aggressor-female-victim *more* negatively than female-aggressor-male-victim, suggesting the trivialization of harm inflicted on male victims may be driven by male judgers (Arnestad et al., 2020). Moreover, when cued to the masculinity of homosexual male victims of a hate crime, heterosexual males downplayed their compassion for the victims (Owuamalam & Matos, 2022), suggesting the masculinity signals can typecast the target as less vulnerable, and thereby reduce moral concerns (Dijker, 2010). Furthermore, males (vs.

females) scoring higher than women in benevolent sexism (Glick & Fiske, 1996) suggests that males may also be more likely to typecast females as victims who need to be protected (Reynolds et al., 2020), and thereby males (vs. females) may find it especially less acceptable to cheat females.

Research also suggests that females may be more favorable to other females than males may be to other males (Rudman & Goodwin, 2004), suggesting gendered in-group favoritism among females only. In the context of gendered interpersonal dishonesty, this in-group favoritism view suggests that females may also be especially less likely to cheat other females over males. Given the highly plausible but mechanism-wise unclear moderating effects of the sex of decision-makers, the presence of unmentioned-sex conditions can help us disentangle the dishonesty-facilitating and dishonesty-inhibiting components. If females cheat other females less than they cheat other unmentioned-sex participants, it can be inferred as specific evidence of a “gendered” in-group favoritism among females that reduces the dishonesty of females against other females. Similarly, if males cheat other females less than they cheat other unmentioned-sex participants, it can be inferred as specific evidence of the stronger gendered effect in interpersonal dishonesty among males (vs. females) being driven by their greater readiness to typecast females as victims. The possibilities further underscore the importance of examining the interpersonal dishonesty facilitating vs. inhibiting in their own rights, vis-à-vis strict baselines, not just for the main effects of the target’s sex but also for moderating effects across the sex of decision-makers.

Method

A large-scale, incentivized, and multi-country experiment examined two questions: (1) Are individuals more likely to cheat males or females? and (2) Are the effects symmetrical across the decision-maker's own sex and attitudes? To measure behavioral dishonesty, participants were tempted to cheat another same-sex or opposite-sex participant, with no risk of detection, punishment, or reciprocity, in a modified "mind game" (Agneman & Chevrot-Bianco, 2023; Fischbacher & Föllmi-Heusi, 2013; Jiang, 2013) task.

In a between-subjects design, participants were randomly assigned to one of three conditions depending on the disclosed sex of the paired participant serving as a target: male, female, or unmentioned. Because incentives are held constant and the only experimental difference across conditions was the target participant's sex, this experiment can estimate differences in the isolated *moral cost* of cheating males versus females. Furthermore, the unmentioned-sex condition serves as a baseline to estimate average levels of dishonesty when information about the target's sex is *not* disclosed.

The study complied with all relevant ethical regulations regarding human research participants, including the guidelines from the Helsinki Declaration. As Norwegian laws and regulations do not require review by an institutional review board for anonymous, non-medical, low-risk research with human participants, the project was not submitted to such a review.

Informed consent was collected from all participants. All participants were recruited through Prolific, a crowdsourcing survey platform that provides non-identifiable data and meets the high standards of GDPR. No form of deception was used. Responses remained completely anonymous. Bonus pay-outs were in addition to base payments for participation. The approved protocol after Stage 1 acceptance in principle was registered on the OSF here. All code, data, and study materials can be made available on request.

The experiment was implemented using Qualtrics (Qualtrics, 2022), and the analyses were conducted in the latest version of R (R Core Team, 2013) and jamovi (The jamovi project, 2022). Visualizations were created using the ggplot2 (Wickham, 2016), and regression output tables were generated using sjPlot (Lüdecke, 2023).

Participants and power

A total of 3,380 complete responses were received from nine countries[20]: Australia, Canada, Germany, Italy, Poland, Portugal, South Africa, the UK, and the USA. Participants were recruited using 1:1 sex-balanced recruitment for English-speaking participants with at least a 95% approval rate for at least 50 submissions, with a quota of 375 complete responses per country. As pre-registered, 212 responses were removed from participants failing an attention check where they were asked to select "A lot" to the question "How much do you like painting?" to show that they were paying attention to instructions. Deviating from one pre-registered exclusion criterion, responses deemed "fraudulent" by Qualtrics' metrics could not be removed as the feature was not switched on. However, because Prolific screens participants when joining the platform and issues a unique participant ID to each participant, the chances of duplicate responses are minimal. Finally, 2 responses, from participants who indicated their biological sex being neither male nor female, were removed.

This resulted in a dataset with valid responses from 3,166 participants (M_age = 33.09, SD = 11.45). Power analyses using G*POWER (Faul et al., 2007) showed that the minimal detectable effect size in pairwise comparisons with 90% and two-tailed $p < .05$ ranged from $d = .203$ to $d = .197$.

Procedures

The experiment consisted of three main stages: (1) participant pairing, (2) dishonesty task, and (3) individual-level measures.

[20] Based having at least 1,500 active English-speaking participants on Prolific at the time of originally submitting the Stage 1 proposal.

Participant Pairing. After responding to the attention check, participants indicated their age, income, and education, and were told they would be paired with another participant for the task. Depending on the condition, participants were told, “You have been paired with another male/ female/ Prolific participant. Your decisions in the task can affect both your own and the male/ female/ Prolific participant’s earnings.”

Participants did not engage in any direct communication or interaction with their pair. The structure of the study is designed to implement the *monetary impacts* of participants’ potential dishonesty without actual interpersonal exchange. This approach was inspired by economic games to examine decision-making affecting others using online samples (e.g., Dimant, 2023). This format has been shown to effectively leverage large samples without the need for direct participant interaction.

Once the data collection phase was completed, the same participants who participated as decision-makers were randomly assigned the role of targets (i.e., potential victims of cheating). In the role of “targets”, those participants need not perform *any* additional tasks. They merely *received* the monetary impacts, in terms of additional earnings, of potential cheating by the decision-making participants. For example, if a decision-making participant with Prolific ID-1 claimed £45 as their additional earning, the “target” participant with Prolific ID-2 only needed to receive the £15, to actualize the consequences of the task’s outcome. This eliminated the need to recruit a separate set of target participants because the same decision-making participants can be paid earnings as *targets* via the same bonus awarding feature on Prolific.

Programming in R was used to implement the randomization assignment of decision-makers to targets based on their sex and experimental conditions. A digital "hat" was used to draw participant IDs randomly, ensuring that each participant, in their role as a target, is matched fairly and without bias. This random draw ensured that participants were paired with the sex they were told they would be paired with, no participant was paired with themselves, and the overall balance between different groups was maintained.

Dishonesty task. For the measure of interpersonal dishonest behavior, participants completed a 10-round die-rolling task where they could earn up to £60 (in addition to their base payment)- depending on their self-reported die scores and luck in the lottery. Participants were also instructed that if their additional earnings are less than £60, the remaining amount will be given to the participant (male, female, or no-sex) they have been paired with. To elicit incentive-aligned choices, participants were told that 40 participants in total will be randomly chosen by lottery, and be paid their additional payoffs within a week of completion of the study.

On the task page, participants were asked to roll a six-sided die privately (to exclude any detection risk) exactly 10 times and report the scores of the 10 die rolls. They could use a physical die, or simply search “Roll a die” on their browser to roll a virtual die. Then, participants were asked to self-report the 10 scores into a form field with 10 input boxes accepting numbers from 1 to 6. They also confirmed the total sum of reported die scores on a constant-sum double slider question that clearly showed how the decision-making participant’s self-reporting can directly affect their own and their paired participant’s potential additional earnings, while explicitly showing the male, female, or unmentioned sex signal of the paired participant.

Of course, participants could over-report the scores of their die rolls, and selfishly increase their earnings at the cost of their interaction partner without any detection, punishment, or reputational concerns. This creates a clear tradeoff between the moral motivation to report truthfully and the financial motivation to over-report scores and increase personal gain. Importantly, the question is, *how much* participants would over-report, and the amount may be systematically different when paired with male versus female targets.

Furthermore, having ten rounds of die-rolls increases sensitivity as individuals not only have a binary decision to report truthfully or dishonestly each round but can also vary *levels* of dishonesty, over a wider range across the 10 rounds.

Individual-level measures. After the dishonesty task, participants were asked to what extent they agreed with statements about the morality of their actions ("I believe my actions in the previous task are acceptable."), inference about the paired participant ("I believe the paired X would view my actions as fair."), and beliefs about cheating different sexes in general ("It is equally unacceptable to cheat men or women for money.", "It is more acceptable to cheat men than women for money.", and "It is more acceptable to cheat women than men for money.") presented on slider scales from 0 (Completely disagree) to 10 (Completely agree).

Participants also responded to three questions on sex, gender, and sexual orientation taken from Lefler and colleagues 2023: (1) What is your biological sex? (male, female, intersex, not listed [please specify], (2) What is your identity? (male, female, nonbinary/ fluid, queer/genderqueer, not listed [please specify], and (3) What is your sexual orientation? (heterosexual, gay or lesbian, bisexual, queer, pansexual, asexual, not listed [please specify]).

At the end of the survey, participants indicated their agreement to three exploratory measures ("My paired X would have cheated if they were in my shoes", "If I were to cheat my paired participant, I would <u>not</u> feel guilty", and "My paired X would have expected me to cheat them") presented on slider scales from 0 (Completely disagree) to 10 (Completely agree).

In measures in the survey, the "X" was replaced by a piped-in feature to show "participant", "male participant", or "female participant", depending on the experimental condition.

Results

Pre-registered analyses. The main outcome variable is the sum of the 10 reported die scores. For fair die rolls, the average reported die score should be 3.5 if everyone reports all of their rolls truthfully. So, the statistical expectation of the mean of the *sum* of the 10 reported die scores is 35. Then, should the mean reported sum of die scores be above 35, we can infer that some participants may have over-reported their die-rolls to cheat their paired participant.

A series of linear mixed-effects models were conducted, incorporating random effects for each country. This hierarchical or multi-level modeling approach accounts for the possibility that each participant's response is influenced not only by individual characteristics but also by the broader context of their country. By distinguishing between variance at the individual level and variance at the country level, the main outcome- the sum of the reported die rolls- can be estimated more precisely.

As pre-registered, the modeling approach was incremental: starting with only random effects in Model 1, adding fixed effects of the target's sex in Model 2, fixed effects of the decision-maker's sex in Model 3, and finally, their two-way interaction in Model 4. Please see Table 1 for a summary of the Results across models.

Table 1. Results from linear mixed effects models. Outcome variable: Total earnings from 10 rounds of self-reported die rolls.

	Model 1		**Model 2**		**Model 3**		**Model 4**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	36.72 (36.52 – 36.93)	**<0.001**	36.40 (36.04 – 36.75)	**<0.001**	36.80 (36.51 – 37.09)	**<0.001**	36.77 (36.27 – 37.27)	**<0.001**
Target [Male]			0.47 (-0.03 – 0.97)	0.064			-0.28 (-0.99 – 0.44)	0.449
Target [Neutral]			0.51 (0.01 – 1.01)	**0.045**			0.35 (-0.35 – 1.06)	0.321
Decision-maker [Female]					-0.15 (-0.56 – 0.26)	0.465	-0.74 (-1.44 – -0.03)	**0.040**
Decision-maker [Female] X Target [Male]							1.47 (0.47 – 2.46)	**0.004**
Decision-maker [Female] X Target [Neutral]							0.30 (-0.70 – 1.29)	0.561
Random Effects								
σ^2	34.24		34.21		34.24		34.13	
τ_{00}	0.00 Country		0.00 Country		0.00 Country		0.00 Country	
N	10 Country		10 Country		10 Country		10 Country	
Observations	3166		3166		3166		3166	
Marginal R^2 / Conditional R^2	0.000 / NA		0.002 / NA		0.000 / NA		0.005 / NA	

We registered the following hypotheses:

H1. Decision-makers will be more dishonest against male targets than female targets.

H2. Male decision-makers will be more dishonest against male (vs. female) targets than female decision-makers will other male (vs. female) targets.

First, Model 2 shows that male targets were not cheated significantly more than female targets, $b = .47$, $p = .064$, 95% CI [-.03, .97]. Rather, neutral targets were cheated significantly more than female targets, $b = .51$, $p = .045$, 95% CI [.01, 1.01].

Post hoc tests showed that compared to mean total earnings when the target was female ($M=36.49$, $SD=5.81$), mean total earnings was significantly higher when target was neutral ($M = 36.91$, $SD=6.02$), $t\ (3163) = 2.01$, $p = .045$, $d = .087$, 95% CI [.002, .173], and not significantly different when target was male ($M = 36.87$, $SD = 5.70$), $t\ (3163) = 1.85$, $p = .064$, $d = .081$, 95% CI [-.002, .166]. Comparing mean total earnings against the statistical expectation of £35, estimated dishonesty against female targets was 22% less than against neutral targets[21].

Because female targets were not cheated significantly less than male targets, we do not find statistical support for Hypothesis 1. Following our preregistration, we conducted a test of equivalence to examine if total earnings were not significantly different at the $d = .2$ effect size level. The results showed a non-significant TOST upper, $t\ (2108) = -1.08$, $p = .861$, and a significant TOST lower, $t\ (2108) = -2.68$, $p = .004$. As the lower TOST is non-significant at an alpha of .05, we cannot conclusively conclude null effects for Hypothesis 1.

Second, Model 4 shows a significant decision-maker [Female] X target [Male] interaction effect, $b = 1.47$, $p = .004$, 95% CI [.47, 2.46]. Interestingly, contrary to Hypothesis 2, the effect was driven by female decision-makers and not male decision-makers. For female decision-makers, compared to total earnings when target was female ($M=36.03$, $SD=5.68$), earnings was significantly higher when target was male ($M=37.22$, $SD=5.55$), $t\ (3160) = 3.34$, $p = .001$, $d = .204$, 95% CI [.084, .323], and not significantly different when target was neutral ($M=36.68$, $SD=5.65$), $t\ (3160) = 1.80$, $p = .072$ $d = .111$, 95% CI [-.010, .233]. This suggests that female decision-makers cheated male targets more than they cheated female

[21]Percentage = (36.91 – 36.42)/(36.91 – 35)

targets. Comparing mean total earnings against the statistical expectation of £35 from 10 die rolls, estimated dishonesty by female decision-makers against female targets was 53.6% less than against male targets[22].

For male decision-makers, compared to total earnings when target was female (M=36.76, SD=5.92), earnings was not significantly different when target was male (M=36.49, SD=5.85), t (3160) =-0.76, p = .449, d = -.047, 95% CI [-.169, .075], or neutral (M=37.12, SD=6.35), t (3160) = 0.99, p = 320, d = .061, 95% CI [-.059, .181]. Following our preregistration, a test of equivalence showed a non-significant TOST upper, t (1032) = -1.30, p = .097, and non-significant TOST lower, t (1032) = -0.21, p = .582. The upper TOST being non-significant at an alpha of 0.05 suggests that the evidence from the experiment does not conclusively support or reject Hypothesis 2.

Although Hypothesis 1 and 2 are neither supported nor conclusively rejected, there is strong empirical evidence of gender biases in dishonesty, nonetheless. Rather than male targets being cheated more than female targets, female targets are cheated less than neutral targets. This suggests that individuals, on average, find it less acceptable to cheat a target with a female signal. Further, rather than male decision-makers cheating female targets less than male targets, female decision-makers cheated female targets less than they did male targets. Please see Figure 1 for an illustration of comparisons across targets and decision-makers.

[22]Percentage = (37.22 – 36.403)/(37.22 – 35)

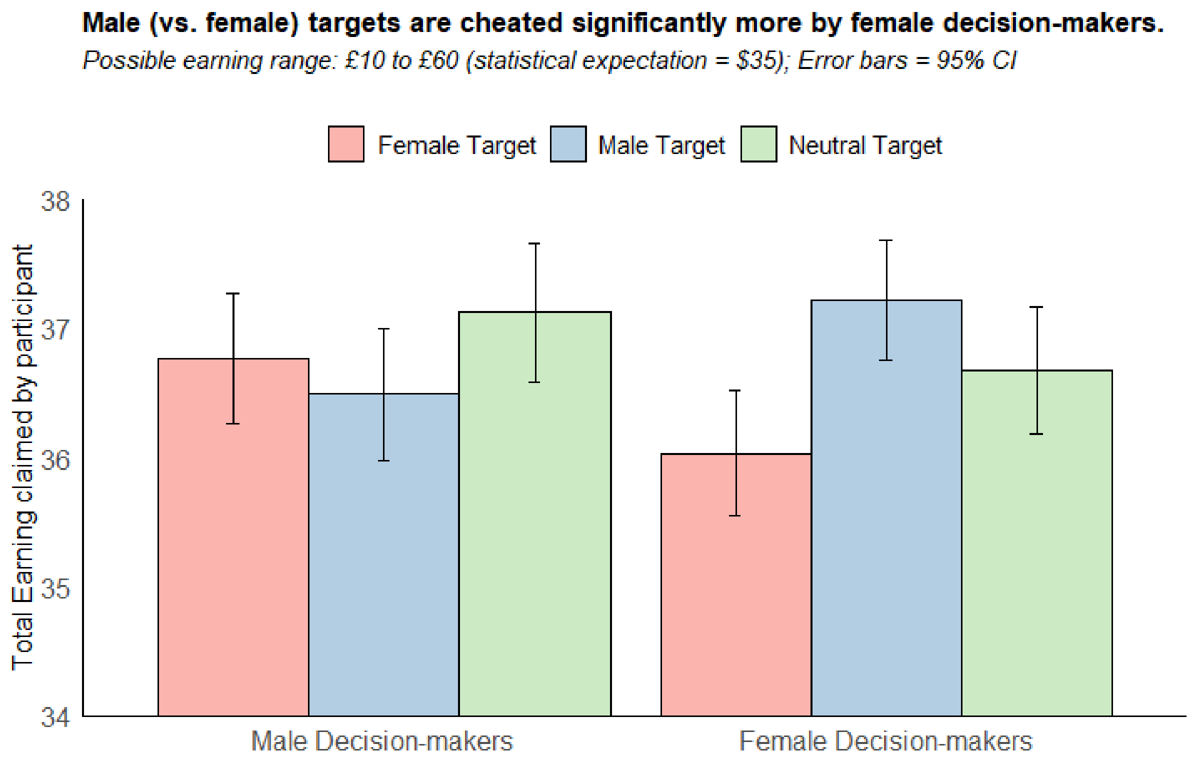


Figure 1. Mean total earnings across decision-maker sex and experimental conditions. The statistical expectation of total earnings, if participants reported truthfully, is £35, which is exceeded significantly in all cells. Error bars represent 95% confidence intervals.

Secondary analyses. To compare opinions of whether it is more acceptable to cheat men than women and vice versa, a paired samples t-test was conducted to compare agree to the statements "It is more acceptable to cheat men than women for money" and "It is more acceptable to cheat women than men for money". The results showed that on average, people express higher agreement with the statement about it being more acceptable to cheat men more than women ($M = 0.850$, $SD = 2.03$), than cheat women more than men ($M = 0.52$, SD 1.41), $t(3165) = 12.66$, $p < .001$, $d = .225$, 95% CI [.190, .260]. Although the mean agreements are quite low on the scale, please note that even in the presence of social desirability and the option to choose the anchor of "Completely disagree" to both statements, both male and female decision-makers express greater agreement to cheat men than women versus women than men (please see Figure 2 for an illustration).

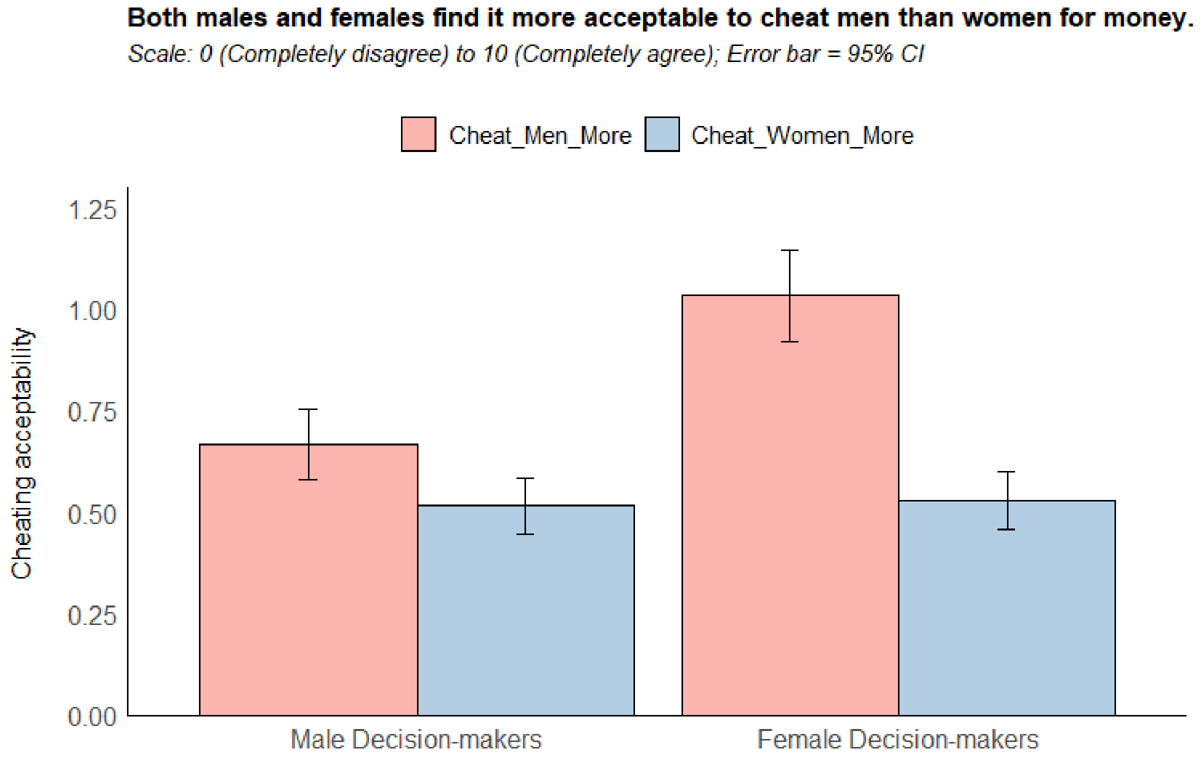


Figure 2. Mean reported agreement with the statements "It is more acceptable to cheat men than women for money" and "It is more acceptable to cheat women than men for money" across male and female participants. Both sexes report a greater acceptability of cheating men than women.

Exploratory analyses. Given that there was no conclusive support favoring or contradicting the pre-registered hypotheses, a series of exploratory analyses were conducted to examine the robustness of the results across datasets (e.g., including all complete responses without excluding those who failed the attention check). Please see Table S2 in Supplemental Materials for outputs from linear mixed effects models with country-level random effects intercepts, and fixed effects of target sex, decision-maker sex, and their interaction.

Surprisingly, we find that in the "All complete responses" dataset (i.e., not excluding responses from participants who failed the attention check), male targets are indeed cheated more than female targets, $b = .53$, $p = .036$, 95% CI [.05, 1.02]. Participants failing the attention check may have paid attention to the decision-making components of the study. So, an extra 212 responses make smaller effect sizes slightly more detectable. However, this result should be interpreted as merely speculative.

To assess the robustness of the interaction effect, I used a series of linear mixed-effects models across datasets as well. Please see Table S3 in Supplemental Materials for full outputs. The key trend here is the Decision-maker [Female] X Target [Male] interaction is significant across datasets, with no changes in the sign of the coefficient. Particularly interesting may be how the estimated coefficient for the interaction ($b = 1.81$) was markedly larger in the "Heterosexual = yes" dataset compared to other datasets where the same coefficient ranged from 1.39 to 1.47.

Finally, I also compared responses to exploratory measures across the cells using linear-mixed effects on the pre-registered dataset. Specifically, the goal was to explore variation in participant responses to what extent they would feel not guilty if they were to cheat their paired participant, they expected their pair would have cheated in their shoes, and their pair expected to be cheated. Comparing responses across the target sex alone (Table S4 in Supplemental Materials) and incorporating the interactions with decision-maker sex (Table S5 in Supplemental Materials) may help us gain insights into how individuals think about gender and dishonesty.

Seeing that the target's sex is the main predictor of interest and the absence of significant interaction effects (Table 5), a series of post hoc tests for each exploratory outcome measure. Because of the exploratory nature of the analysis, the p values reported in the following post hoc tests are Tukey corrected. Nonetheless, illustrations of differences across the decision-maker's sex are included in the form of Figures 3-5 for a more nuanced visualization.

First, participants were less likely to "not feel guilty" if they were to cheat a female paired participant (M=1.83, SD=2.76) than compared to both a neutral (M=2.13, SD=2.90), t (3163) = -2.72, p = .018, d = -.118, 95% CI [-.204, -.033], and male (M=2.17, SD=2.91), 2.42, p = .040, d = -.106, 95% CI [-.191, -.020] paired participant.

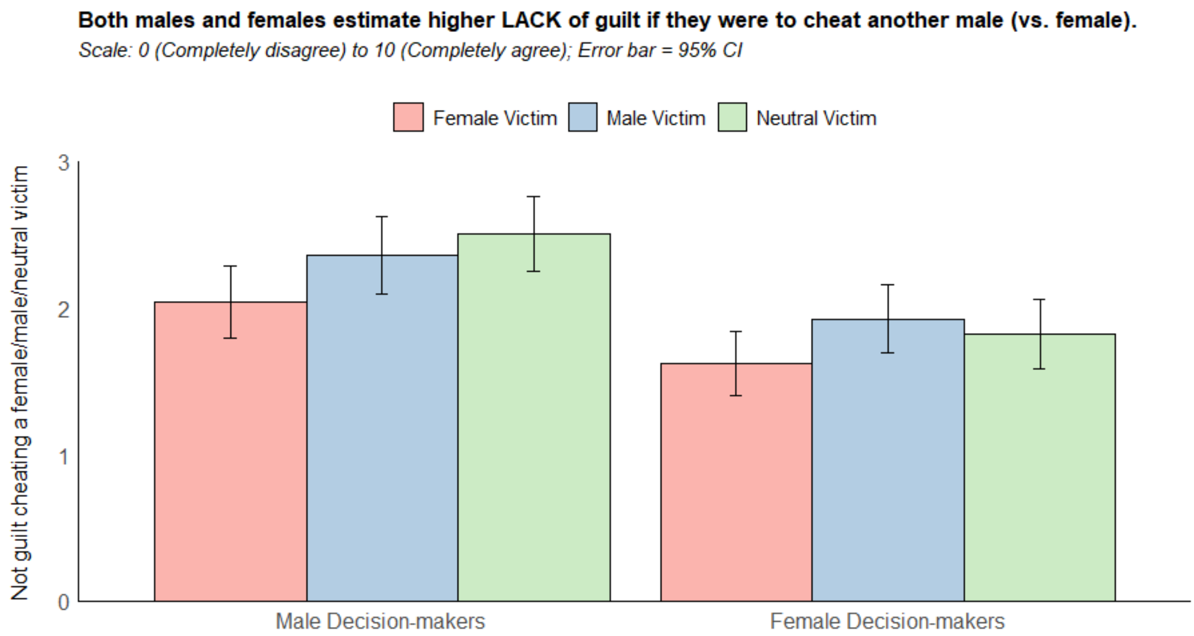


Figure 3. Mean reported agreement with the statement "If I were to cheat my paired male/female/participant, I would not feel guilty." across male and female decision-making participants. Both sexes report greater agreement when asked about cheating male than female targets.

Second, participants were less likely to expect that a paired female participant (M=3.76, SD=2.96) would cheat them than both a neutral (M=3.64, SD=2.87), t (3163) = -3.57, p = .001, d = -.156, 95% CI [-.241, -.070] and male (M=3.90, SD=2.98), t (3163) = -5.65, p < .002, d = -.246, 95% CI [-.332, -.161].

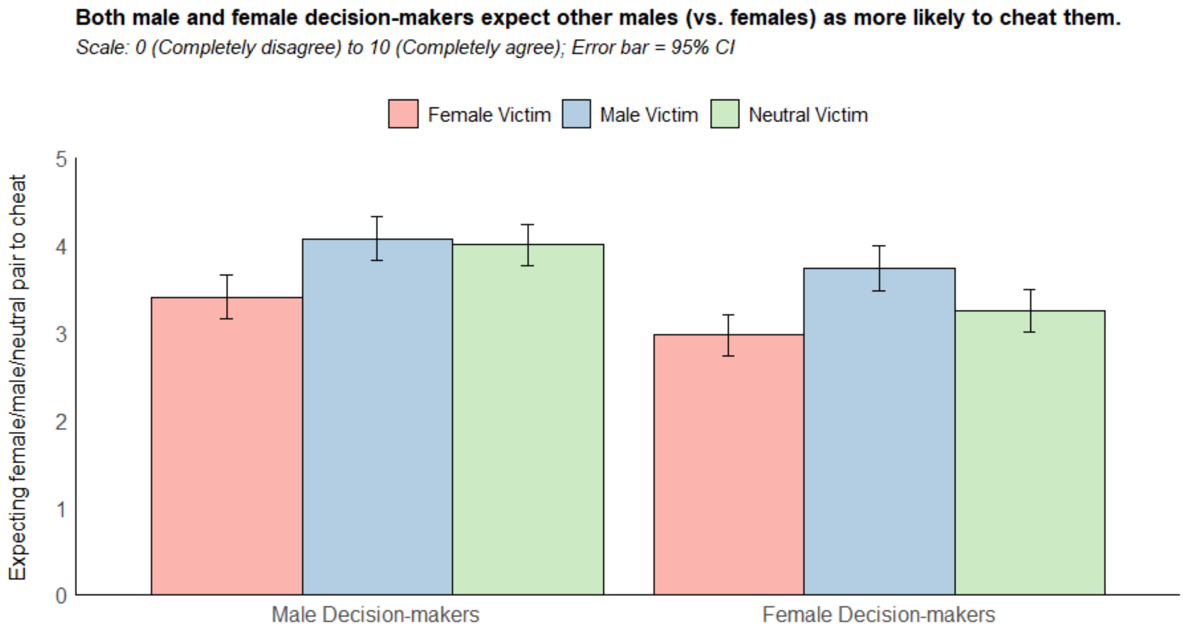


Figure 4. Mean reported agreement with the statement "My paired male/female/participant would have cheated if they were in my shoes." across male and female decision-making participants. Both sexes report greater agreement about paired male than female targets.

Finally, participants were less likely to think that a female paired participant (M=3.76, SD=2.96) expects to be cheated than both a neutral (M=4.19, SD=3.00), t (3163) = -3.24, p = .003, d = -.141, 95% CI [-.233, -.056], and male (M=4.21, SD=3.08), t (3163) = -3.39, p = .002, d = -.148, 95% CI [-226, -.062] paired participant.

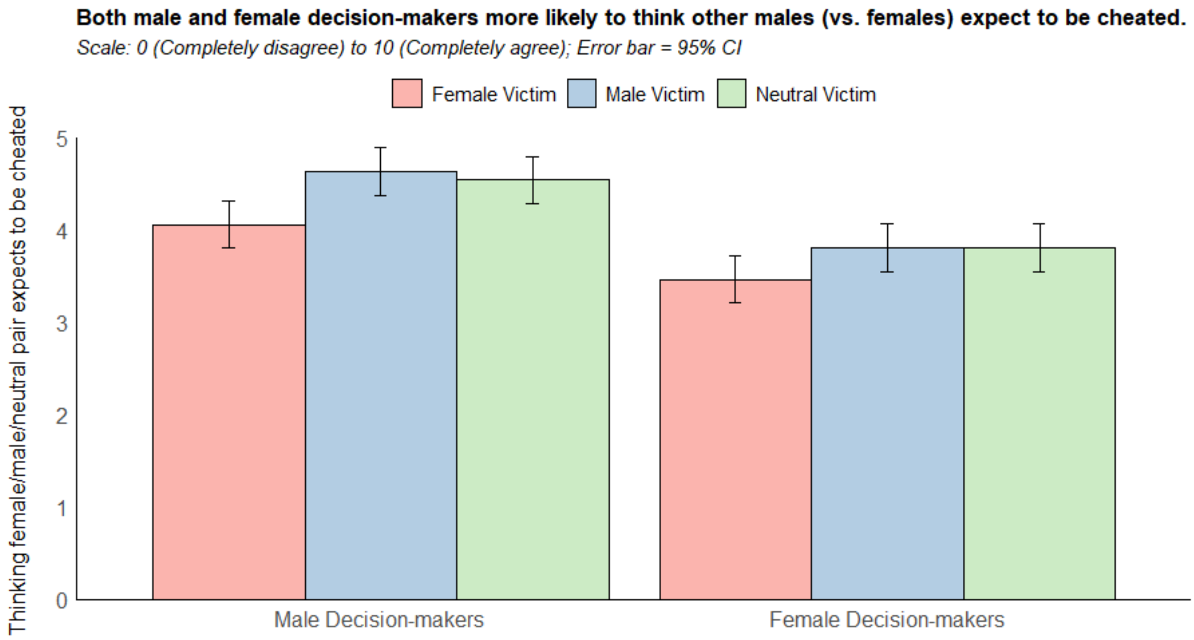


Figure 5. Mean reported agreement with the statement "My paired male/female/participant would have expected me to cheat." across male and female decision-making participants. Both sexes report greater agreement about paired male than female targets.

General Discussion

The current research examined two primary questions: (1) Are individuals more likely to cheat males or females? and (2) Are the effects symmetrical or asymmetrical across male and female decision-makers? A large-scale nine-country incentivized experiment (N = 3,166) tried to answer these questions.

Regarding the first question, the results showed that although individuals significantly cheated female targets less than unmentioned-sex targets (about 22% difference), the differences were not significant compared to male targets. This suggests that rather than a male target *increasing* dishonesty, as originally hypothesized in H2, the results indicate that a female signal of a target *reduces* dishonest behavior. Because participants could cheat under conditions of zero detection, punishment, and reputational risks, these differences support the view that individuals may indeed adopt higher moral standards when faced with the prospect of cheating a female target.

Regarding the second question, the effect was indeed asymmetrical: Gender bias in target-based dishonest behavior was driven by female decision-makers cheating other females less than they did males (about 53.6% difference). Dishonest behavior by male decision-makers did not differ across male and female targets. This finding aligns well with classical findings on how females are more favorable to other females, in the form of gendered ingroup favoritism (Rudman & Goodwin, 2004). Further, a review of economic experiments found substantial differences in social preferences of men and women, proposing the explanation that women are more sensitive to social cues (Croson & Gneezy, 2009). That may explain why female decision-makers may be more likely to consider the characteristics of the victim in potential dishonest behavior.

These specific findings first add to the nascent literature on how the characteristics of the victim affect dishonest behavior against them (Rotman et al., 2018; Soraperra et al., 2019; Yam & Reynolds, 2016). Second, the findings also extend the notion of a gender bias in moral typecasting (Reynolds et al., 2020) from differences in moral judgments of harm inflicted (also Graso et al., 2023), to behavioral evidence of gender-biased interpersonal dishonest behavior. More broadly, the current research adds to the growing literature on accounting for gender dynamics in moral behavior (Capraro, 2018; Childs, 2012; Grosch & Rau, 2017; Kastlunger et al., 2010; Ward & King, 2018), which has to date primarily examined gender differences from an agent/perpetrator/decision-maker perspective.

Practically, findings from the current research may be relevant to organizations, peer-to-peer platforms, and policymakers seeking to gain a deeper understanding of how gender may influence dishonest behavior. This is especially supported by the fact that dishonesty in lab-based settings such as the die-rolling paradigm has been shown to predict dishonest behavior in the real world such as fare evasion (Dai et al., 2018). Seeing how female targets are cheated less may imply that emphasizing the effect of dishonesty on females may help curb dishonest behavior such as in negotiations and transactions. This may be especially applicable to wider settings where parties have limited information about each other, and possess private information with economic potential such that one can use that to selfishly gain at the cost of another. Further, seeing how female decision-makers are especially less likely to cheat other females than males, while males are equally likely to cheat females and males, perhaps broad interventions may be effective in curbing male dishonesty. In that vein, female-female interactions may be less likely to include fraud and so auditing and surveillance resources could be diverted to other areas.

Of course, interpretations from this research have several limitations. First, although the study was conducted with participants from nine countries, the selection of countries was based on the availability of participants on the recruitment platform. So, future research aiming for more precise cross-cultural comparisons, may benefit from developing theoretical selection criteria when choosing countries to recruit participants. Second, although the study used the established die-rolling paradigm and had the advantage of comparing dishonest behavior in the absence of detection, punishment, and reputational consequences, it may be interesting nonetheless to test if the differences flip in other settings or different paradigms such as the "online matrix task" (Peer et al., 2024; Peer & Feldman, 2021).

Third, because the study's dishonesty measure only detects differences at the aggregate level and not the individual level, using honesty tasks that can detect cheating at the individual level (e.g., "spot the difference task" by Speer et al., 2020) may help increase confidence with lower sample size requirements. Finally, the current study was in many ways a "conservative test "of the hypotheses as participants merely made decisions toward each other, without actual interactions, which may reinforce biased dishonesty. So, future research may benefit from incorporating creative designs that incorporate communication into dishonesty research (Tønnesen et al., 2024), and perhaps extend how interpersonal dishonesty may vary when decision-makers and targets communicate before potential dishonesty occurs.

Despite the limitations, the current research is one of the first of its kind to study gender dynamics in moral decisions in the registered report format. For now, it can be concluded that female targets are cheated less than male targets, and the effect is driven by female decision-makers cheating female targets significantly less than male targets.

References

Agneman, G., & Chevrot-Bianco, E. (2023). Market Participation and Moral Decision-Making: Experimental Evidence from Greenland. *The Economic Journal*, *133*(650), 537–581. https://doi.org/10.1093/ej/ueac069

Arnestad, M., Studzinska, A., Nordmo, M., & Matthiesen, S. B. (2020). *#HeToo? Men trivialize cases of sexual harassment by a female aggressor toward a male victim, but women do not* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/q2zhg

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162. https://doi.org/10.1037/h0036215

Boghrati, R., & Berger, J. (2023). Quantifying cultural change: Gender bias in music. *Journal of Experimental Psychology: General*, *152*(9), 2591–2602. https://doi.org/10.1037/xge0001412

Cameron, J. J., & Stinson, D. A. (2019). Gender (mis)measurement: Guidelines for respecting gender diversity in psychological research. *Social and Personality Psychology Compass*, *13*(11), e12506. https://doi.org/10.1111/spc3.12506

Cappelen, A. W., Falch, R., & Tungodden, B. (2019). The Boy Crisis: Experimental Evidence on the Acceptance of Males Falling Behind. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3348981

Capraro, V. (2018). Gender differences in lying in sender-receiver games: A meta-analysis. *Judgment and Decision Making*, *13*(4), 345–355. https://doi.org/10.1017/S1930297500009220

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

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, *47*(2), 448–474. https://doi.org/10.1257/jel.47.2.448

Dai, Z., Galeotti, F., & Villeval, M. C. (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Management Science*, *64*(3), 1081–1100. https://doi.org/10.1287/mnsc.2016.2616

Dijker, A. J. M. (2010). Perceived vulnerability as a common basis of moral emotions. *British Journal of Social Psychology*, *49*(2), 415–423. https://doi.org/10.1348/014466609X482668

Dimant, E. (2023). Hate Trumps Love: The Impact of Political Polarization on Social Preferences. *Management Science*, mnsc.2023.4701. https://doi.org/10.1287/mnsc.2023.4701

Eagly, A. H. (2013). *Sex Differences in Social Behavior: A Social-role interpretation* (1st ed.). Psychology Press. https://doi.org/10.4324/9780203781906

Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*(3), 283–308. https://doi.org/10.1037/0033-2909.100.3.283

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

FeldmanHall, O., Dalgleish, T., Evans, D., Navrady, L., Tedeschi, E., & Mobbs, D. (2016). Moral Chivalry: Gender and Harm Sensitivity Predict Costly Altruism. *Social Psychological and Personality Science*, *7*(6), 542–551. https://doi.org/10.1177/1948550616647448

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

Georgeac, O. A. M., Rattan, A., & Effron, D. A. (2019). An Exploratory Investigation of Americans' Expression of Gender Bias Before and After the 2016 Presidential Election. *Social Psychological and Personality Science*, *10*(5), 632–642. https://doi.org/10.1177/1948550618776624

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, *145*(1), 1–44. https://doi.org/10.1037/bul0000174

Gibson, R., Tanner, C., & Wagner, A. F. (2013). Preferences for Truthfulness: Heterogeneity among and within Individuals. *American Economic Review*, *103*(1), 532–548. https://doi.org/10.1257/aer.103.1.532

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. https://doi.org/10.1037/0022-3514.70.3.491

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

Graso, M., Reynolds, T., & Aquino, K. (2023). Worth the Risk? Greater Acceptance of Instrumental Harm Befalling Men than Women. *Archives of Sexual Behavior*. https://doi.org/10.1007/s10508-023-02571-0

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520. https://doi.org/10.1037/a0013748

Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523. https://doi.org/10.1016/S1364-6613(02)02011-9

Grosch, K., & Rau, H. A. (2017). Gender differences in honesty: The role of social value orientation. *Journal of Economic Psychology*, *62*, 258–267. https://doi.org/10.1016/j.joep.2017.07.008

Jacobsen, C., Fosgaard, T. R., & Pascual-Ezama, D. (2018). WHY DO WE LIE? A PRACTICAL GUIDE TO THE DISHONESTY LITERATURE: WHY DO WE LIE? *Journal of Economic Surveys*, *32*(2), 357–387. https://doi.org/10.1111/joes.12204

Jampol, L., & Zayas, V. (2021). Gendered White Lies: Women Are Given Inflated Performance Feedback Compared With Men. *Personality and Social Psychology Bulletin*, *47*(1), 57–69. https://doi.org/10.1177/0146167220916622

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, *93*, 328–336. https://doi.org/10.1016/j.jebo.2013.04.003

Kajackaite, A., & Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior*, *102*, 433–444. https://doi.org/10.1016/j.geb.2017.01.015

Kastlunger, B., Dressler, S. G., Kirchler, E., Mittone, L., & Voracek, M. (2010). Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology*, *31*(4), 542–552. https://doi.org/10.1016/j.joep.2010.03.015

Kosakowska-Berezecka, N., Bosson, J. K., Jurek, P., Besta, T., Olech, M., Vandello, J. A., Bender, M., Dandy, J., Hoorens, V., Jasinskaja-Lahti, I., Mankowski, E., Venäläinen, S., Abuhamdeh, S., Agyemang, C. B., Akbaş, G., Albayrak-Aydemir, N., Ammirati, S., Anderson, J., Anjum, G., … Żadkowska, M. (2023). Gendered Self-Views Across 62 Countries: A Test of Competing Models. *Social Psychological and Personality Science*, *14*(7), 808–824. https://doi.org/10.1177/19485506221129687

Kray, L. J., Kennedy, J. A., & Van Zant, A. B. (2014). Not competent enough to know the difference? Gender stereotypes about women’s ease of being misled predict negotiator deception. *Organizational Behavior and Human Decision Processes*, *125*(2), 61–72. https://doi.org/10.1016/j.obhdp.2014.06.002

Lefler, E. K., Tabler, J., Abu-Ramadan, T. M., Stevens, A. E., Serrano, J. W., Shelton, C. R., & Hartung, C. M. (2023). Sex, Gender, and Sexual Orientation in Psychological Research: Exploring Data Trends & Researcher Opinions. *Psychological Reports*, 00332941231199959. https://doi.org/10.1177/00332941231199959

Martuza, J., Sjåstad, H., & Thorbjørnsen, H. (2023). *Business-Size Bias in Moral Concern: People are More Dishonest Against Big than Small Organizations* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/aujfy

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479. https://doi.org/10.1073/pnas.1211286109

Oh, D., Dotsch, R., Porter, J., & Todorov, A. (2020). Gender biases in impressions from faces: Empirical studies and computational models. *Journal of Experimental Psychology: General*, *149*(2), 323–342. https://doi.org/10.1037/xge0000638

Owuamalam, C. K., & Matos, A. S. (2022). Heterosexual men in Trump's America downplay compassion more for masculine (than for feminine) gay victims of hate crime: Why? *European Journal of Social Psychology*, *52*(2), 280–304. https://doi.org/10.1002/ejsp.2787

Peer, E., & Feldman, Y. (2021). Honesty pledges for the behaviorally-based regulation of dishonesty. *Journal of European Public Policy*, *28*(5), 761–781. https://doi.org/10.1080/13501763.2021.1912149

Peer, E., Mazar, N., Feldman, Y., & Ariely, D. (2024). How pledges reduce dishonesty: The role of involvement and identification. *Journal of Experimental Social Psychology*, *113*, 104614. https://doi.org/10.1016/j.jesp.2024.104614

Perales, F., Kuskoff, E., Flood, M., & King, T. (2023). Like Father, Like Son: Empirical Insights into the Intergenerational Continuity of Masculinity Ideology. *Sex Roles*, *88*(9–10), 399–412. https://doi.org/10.1007/s11199-023-01364-y

Prentice, D. A., & Carranza, E. (2002). What Women and Men Should Be, Shouldn’t be, are Allowed to be, and don’t Have to Be: The Contents of Prescriptive Gender Stereotypes. *Psychology of Women Quarterly*, *26*(4), 269–281. https://doi.org/10.1111/1471-6402.t01-1-00066

Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T., & Huguet, P. (2019). Committees with implicit biases promote fewer women when they do not believe gender bias exists. *Nature Human Behaviour*, *3*(11), 1171–1179. https://doi.org/10.1038/s41562-019-0686-3

Reynolds, T., Howard, C., Sjåstad, H., Zhu, L., Okimoto, T. G., Baumeister, R. F., Aquino, K., & Kim, J. (2020). Man up and take it: Gender bias in moral typecasting. *Organizational Behavior and Human Decision Processes*, *161*, 120–141. https://doi.org/10.1016/j.obhdp.2020.05.002

Ross, W. T., & Robertson, D. C. (2000). Lying: The Impact of Decision Context. *Business Ethics Quarterly*, *10*(2), 409–440. https://doi.org/10.2307/3857884

Rotman, J. D., Khamitov, M., & Connors, S. (2018). Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically. *Journal of Consumer Psychology*, *28*(2), 353–361. https://doi.org/10.1002/jcpy.1002

Rudman, L. A., & Goodwin, S. A. (2004). Gender Differences in Automatic In-Group Bias: Why Do Women Like Women More Than Men Like Men? *Journal of Personality and Social Psychology*, *87*(4), 494–509. https://doi.org/10.1037/0022-3514.87.4.494

Saad, G., & Gill, T. (2001). The effects of a recipient's gender in a modified dictator game. *Applied Economics Letters*, *8*(7), 463–466. https://doi.org/10.1080/13504850010005260

Shepherd, S., Kay, A. C., & Gray, K. (2019). Military veterans are morally typecast as agentic but unfeeling: Implications for veteran employment. *Organizational Behavior and Human Decision Processes*, *153,* 75–88. https://doi.org/10.1016/j.obhdp.2019.06.003

Shor, E., van de Rijt, A., & Fotouhi, B. (2019). A Large-Scale Test of Gender Bias in the Media. *Sociological Science*, *6*, 526–550. https://doi.org/10.15195/v6.a20

Soraperra, I., Weisel, O., & Ploner, M. (2019). Is the victim *Max (Planck)* or *Moritz* ? How victim type and social value orientation affect dishonest behavior. *Journal of Behavioral Decision Making*, *32*(2), 168–178. https://doi.org/10.1002/bdm.2104

Speer, S. P. H., Smidts, A., & Boksem, M. A. S. (2020). Cognitive control increases honesty in cheaters but cheating in those who are honest. *Proceedings of the National Academy of Sciences*, *117*(32), 19080–19091. https://doi.org/10.1073/pnas.2003480117

Thielmann, I., & Hilbig, B. E. (2019). No gain without pain: The psychological costs of dishonesty. *Journal of Economic Psychology*, *71*, 126–137. https://doi.org/10.1016/j.joep.2018.06.001

Tønnesen, M. H., Elbæk, C. T., Pfattheicher, S., & Mitkidis, P. (2024). Communication increases collaborative corruption. *Journal of Experimental Social Psychology*, *112,* 104603. https://doi.org/10.1016/j.jesp.2024.104603

Ward, S. J., & King, L. A. (2018). Gender Differences in Emotion Explain Women's Lower Immoral Intentions and Harsher Moral Condemnation. *Personality and Social Psychology Bulletin*, *44*(5), 653–669. https://doi.org/10.1177/0146167217744525

Yam, K. C., & Reynolds, S. J. (2016). The Effects of Victim Anonymity on Unethical Behavior. *Journal of Business Ethics*, *136*(1), 13–22. https://doi.org/10.1007/s10551-014-2367-5

Supplemental Materials for Article 1

For the manuscript

Does Conceptual Abstraction Moderate Whether Past Moral Deeds Motivate Consistency or Compensatory Behavior? A Registered Replication and Extension of Conway and Peetz (2012)

Table of Contents

1. Codebook of all measurements used in the study

Table S1. List of measures and codebook for the data file.				
Variable	# of items (a)	Item	Data File Column Name	Scale
Attention Check 1	1	Caution This text is about the following issue. In surveys like ours, sometimes participants do not carefully read the instructions and just click randomly to finish the survey. This leads to several random responses and that can compromise the results. To confirm that you read our instructions carefully, please select "A great deal" as your answer to the question "How much do you like sports?" below. How much do you like sports?	Atttention_Check1	Multiple Choice: A great deal A lot A moderate amount A little None at all
Participants' text response prompt		Please describe the event you recalled by typing in the text box below. Please provide as much detail as you can, and write at least a paragraph with complete sentences. Your responses are completely anonymous.	Response	

Willingness to Volunteer (WTV)	5	1. Volunteering is a worthwhile use of my time even if I do not get paid. 2. I am willing to volunteer for an organization I care about without financial compensation for me. 3. Even for an organization I care about, I am unwilling to work without getting paid (R). 4. Without some financial compensation, it is not worth doing volunteer work (R). 5. I am unlikely to undertake any type of work without being paid (R).	1. Vol1 2. Vol2 3. Vol3RC 4. VolRC4 5. VolRC5	Strongly disagree to Strongly Agree (7-point Likert scale)
Attention Check 2		PLEASE SELECT "Strongly disagree" here.	Attention_Check2	Strongly disagree to Strongly Agree (7-point Likert scale)
Willingness to Help (WTH) 4 scenarios, each with two questions, with the latter pertaining to the deservingness of help	8	1. (Cindy) Imagine you are a customer in the restaurant that Cindy visits for lunch. You notice that when the bill comes she is embarrassed to find that she does not have quite enough local currency to pay for her meal (they accept cash only). She apologizes to the waiter and asks for directions to the nearest bank machine, which turns out to be quite far away. You could save her the trouble and	The help items. 1. Cindy_Help 2. Susan_Help 3. Bill_Help 4. Jim_Help The deservingness items 1. Cindy_Deserve	Not at all to Very likely/much (7-point Likert scale)

		embarrassment by giving her the small amount she needs (approximately 50 cents). How likely would you be to give Cindy a small amount of money to save her the hassle of walking many blocks to a bank machine and back? How much does Cindy deserve your help? 2. (Susan) Imagine that you are one of Susan's co-workers. At some point, you are chatting and she asks a favor of you. She needs to deliver a large heavy parcel to the local post office. She does not have a car, but you do, so she asks if you would mind taking it for her on your way home. The post office is in the opposite direction as your home, so helping Susan will add at least 10-15 minutes to your evening commute. How likely would you be to agree to deliver Susan's parcel? How much does Susan deserve your help? 3. (Bill) Imagine you are one of Bill's neighbors. One day	2. Susan_Deserve 3. Bill_Deserve 4. Jim_Deserve	

		you see him working in his garden using a small hand-tool where a bigger one would be much easier to use. You happen to have the exact tool he needs, but you were planning to use it yourself later that day. Bill asks if you have any tools that would help make his gardening chores easier. How likely are you to offer your superior garden tool to Bill even though it would prevent you from using it on the same day? How much does Bill deserve your help? 4. (Jim) Imagine you work in the same building as Jim and sometimes see him in line to buy food at the building food court. One day, the line is particularly slow and you realize that you may not have long to eat your lunch before you must go back to work. Then Jim asks if he could move ahead of you in line because he has only a few minutes before his next meeting and might not get to buy lunch otherwise. How likely are you to let Jim move ahead of you in line even though it will make		

		your short lunch even shorter? How much does Jim deserve your help?		
Attention Check 3		Please indicate your agreement with the statement below. I swim across the Atlantic Ocean to get to work every day.	Attention_Check2.0	Multiple Choice: Strongly disagree Disagree Agree Strongly agree
Event Valence (manipulation check)	1	The event I wrote about made me feel good about myself.	Feel_Good	Completely disagree to completely agree (7-point Likert scale)
Event Distance (manipulation check)	1	The event I wrote about happened a long time ago.	Happened_Long	Completely disagree to completely agree (7-point Likert scale)
Specific Event Time (manipulation check)	1	Specifically, the event I wrote about happened	Happened_Specific	Multiple Choice: Within last week More than a year ago In between a week and a year
Feeling	20	Interested Excited Strong Enthusiastic Proud Alert Inspired Determined Attentive Active Distressed	Feeling_1 Feeling_2 Feeling_3 Feeling_4 Feeling_5 Feeling_6 Feeling_7 Feeling_8 Feeling_9 Feeling_10 Feeling_11	Very slightly or not at all to Extremely (5-point Likert scale)

		Upset Guilty Scared Hostile Irritable Ashamed Nervous Jittery Afraid	Feeling_12 Feeling_13 Feeling_14 Feeling_15 Feeling_16 Feeling_17 Feeling_18 Feeling_19 Feeling_20	
Age	1	What is your age in years	Age_Own	Blank box
Gender	1	What is your gender?	Gender_Own	Multiple Choice: Male Female Other/Prefer not to say
Education	1	What is the highest level of education you have completed?	Education	Multiple Choice: Some high school or less High school diploma or GED Some college, but no degree Associates or technical degree Bacherlor's degree Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, DDS etc.) Prefer not to say
Income	1		What is your total household income before taxes during the past 12 months?	Multiple Choice: Less than $25,000 $25,000-$49,999 $50,000-$74,999 $75,000-$99,999 $100,000-$149,999 $150,000 or more

				Prefer not to say
Zip Code	1	What is your US Zip Code?		Open-ended question.
Moral_Identity	10	Lastly, visualize a person who is caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind. Form a clear image of how that person would think, feel, and act and then indicate the extent to which you agree or disagree with the following statements. 1. It would make me feel good to be a person who has these characteristics. 2. Being someone who has these characteristics is an important part of who I am. 3. I often wear clothes that identify me as having these characteristics. 4. I would be ashamed to be a person who had these characteristics. 5. The types of things I do in my spare time (e.g., hobbies) clearly identify me as having these characteristics 6. The kinds of books and magazines that I read	1. Identity_1_I 2. Identity_2_I 3. Identity_3_S 4. Identity_4_I_R 5. Identity_5_S 6. Identity_6_S 7. Identity_7_I_R 8. Identity_8_S 9. Identity_9_S 10. Identity_10_I	Strongly disagree to Strongly agree (7-point Likert scale)

		identify me as having these characteristics. 7. Having these characteristics is not really important to me (R). 8. The fact that I have these characteristics is communicated to others by my membership in certain organizations. 9. I am actively involved in activities that communicate to others that I have these characteristics. 10. I strongly desire to have these characteristics.		

2. Experimental Manipulations

Table S2. Exact prompts for the recall tasks.	
Morality X Construal Level	Prompt
Moral/Recent	Please recall a time within the last week when you acted in such a way that you felt righteous or honorable. Perhaps you were loyal to a friend, were generous when you could have been selfish, were kind to someone for no particular reason, or caring toward someone who needed you.
Moral/Distant	Please recall a time over a year ago when you acted in such a way that you felt righteous or honorable. Perhaps you were loyal to a friend, were generous when you could have been selfish, were kind to someone for no particular reason, or caring toward someone who needed you.
Immoral/Recent	Please recall a time within the last week when you acted in such a way that you felt guilty or ashamed. Perhaps you were disloyal to a friend, were greedy when you should have shared, were mean to someone for no particular reason, or uncaring toward someone who needed you.
Immoral/Distant	Please recall a time over a year ago when you acted in such a way that you felt guilty or ashamed. Perhaps you were disloyal to a friend, were greedy when you should have shared, were mean to someone for no particular reason, or uncaring toward someone who needed you.
Neutral/Recent	Please recall a time over a year ago when you acted in such a way that you felt neutral or indifferent. Perhaps you were doing something by yourself (e.g., shopping for groceries, doing chores) and your actions did not affect anyone in any way.
Neutral/Distant	Please recall a time within the last week when you acted in such a way that you felt neutral or indifferent. Perhaps you were doing something by yourself (e.g., shopping for groceries, doing chores) and your actions did not affect anyone in any way.

3. Descriptive Statistics

Table S3A. Means (M) and standard deviations (SD) of the two manipulation checks across conditions.

Dependent Variable	Recalled Event Valence	Temporal Distance	Cell size	M	SD
Feel Good	Moral	Recent	447	5.86	1.22
		Distant	1226	5.79	1.33
		Recent + Distant	1673	5.81	1.30
	Immoral	Recent	455	1.89	1.31
		Distant	1211	1.78	1.39
		Recent + Distant	1666	1.81	1.32
	Neutral	Recent	1316	4.03	1.65
		Distant	436	3.91	1.62
		Recent + Distant	1752	4.00	1.64
Happened Long	Recent	Moral	447	4.51	1.39
		Immoral	455	4.29	1.38
		Neutral	1316	4.32	1.42
		Moral + Immoral + Neutral	2218	1.54	1.17
	Distant	Moral	1226	4.63	1.32
		Immoral	1211	4.37	1.39
		Neutral	436	4.26	1.47
		Moral + Immoral + Neutral	2218	4.07	1.86

Table S3B. Means (M) and standard deviations (SD) of the dependent variables across conditions.

Dependent Variable	Recalled Event Valence	Temporal Distance	Cell size	M	SD
Willingness to Volunteer	Moral	Recent	447	4.51	1.39
		Distant	1226	4.63	1.32
		Recent + Distant	1673	4.60	1.34

	Immoral	Recent	455	4.29	1.38
		Distant	1211	4.37	1.39
		Recent + Distant	1666	4.35	1.39
	Neutral	Recent	1316	4.32	1.42
		Distant	436	4.26	1.47
		Recent + Distant	1752	4.31	1.43
Willingness to help	Moral	Recent	447	5.69	0.89
		Distant	1226	5.68	1.01
		Recent + Distant	1673	5.68	0.98
	Immoral	Recent	455	5.56	0.99
		Distant	1211	5.63	1.02
		Recent + Distant	1666	5.60	1.01
	Neutral	Recent	1316	5.53	1.07
		Distant	436	5.48	1.07
		Recent + Distant	1752	5.50	1.07

Table S3C. Means (M) and standard deviations (SD) of the dependent variables across contrasts.				
Dependent Variable	Combined cell	Cell size	M	SD
Willingness to Volunteer	Moral/Distant + Immoral/ Recent	1681	4.54	1.34
	Moral/Recent + Immoral/ Distant	1658	4.41	1.39
	Neutral/Distant + Neutral/Recent	1752	4.31	1.43
Willingness to help	Moral/Distant + Immoral/Recent	1681	5.65	1.01
	Moral/Recent + Immoral/Distant	1658	5.64	.99
	Neutral/Distant + Neutral/Recent	1752	5.50	1.07

4. Sensitivity Analyses (power & minimum detectable effect sizes)

4.1. Minimum detectable effect sizes for the replication analyses (compared to C&P).

Figure S1.1. Minimum detectable effect size in C&P's original study for the interaction effect of event valence X event distance with 90% power: f = .326.

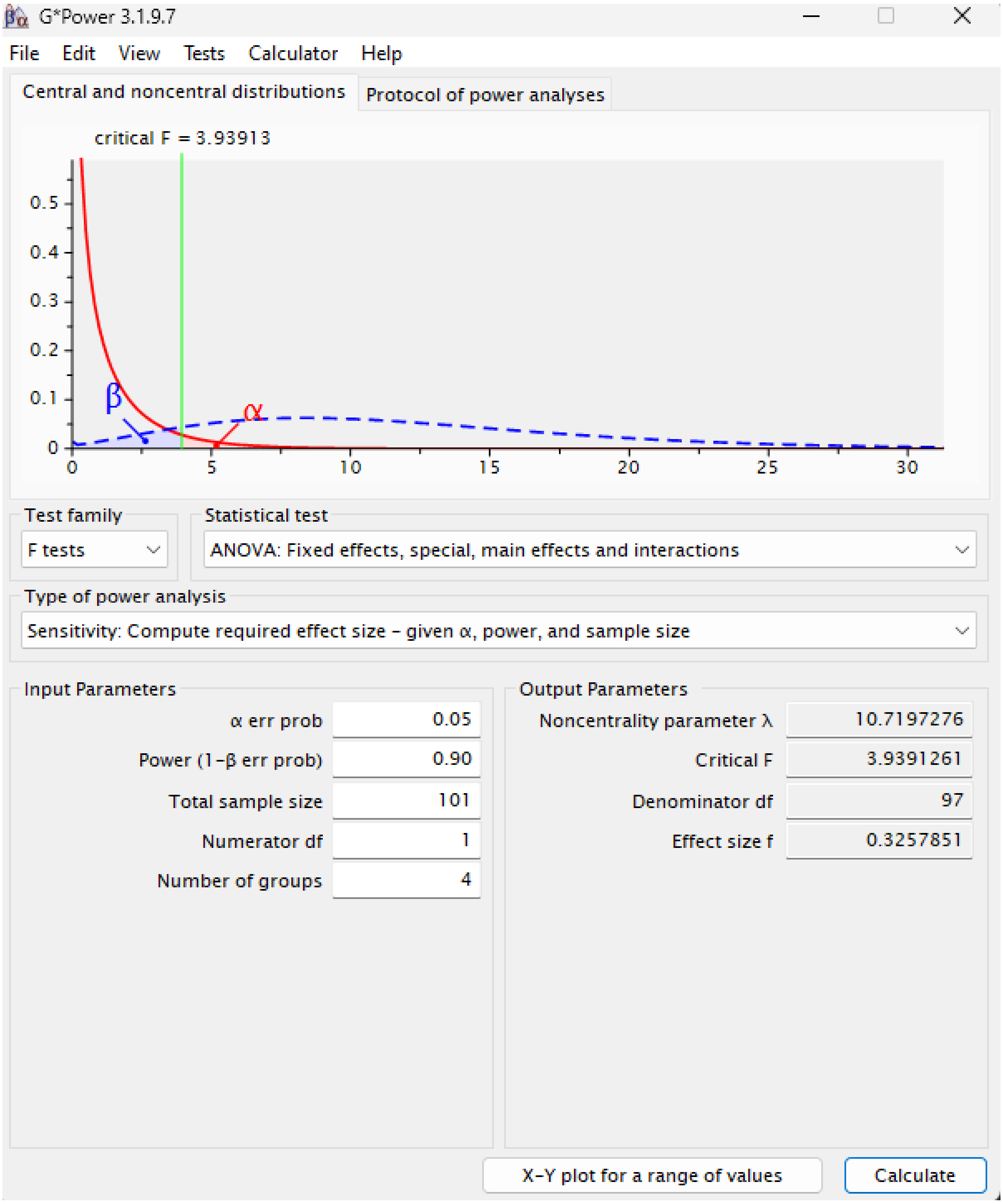


Figure S1.2. Minimum detectable effect size for the interaction effect of event valence X event distance with 90% power: f = .056.

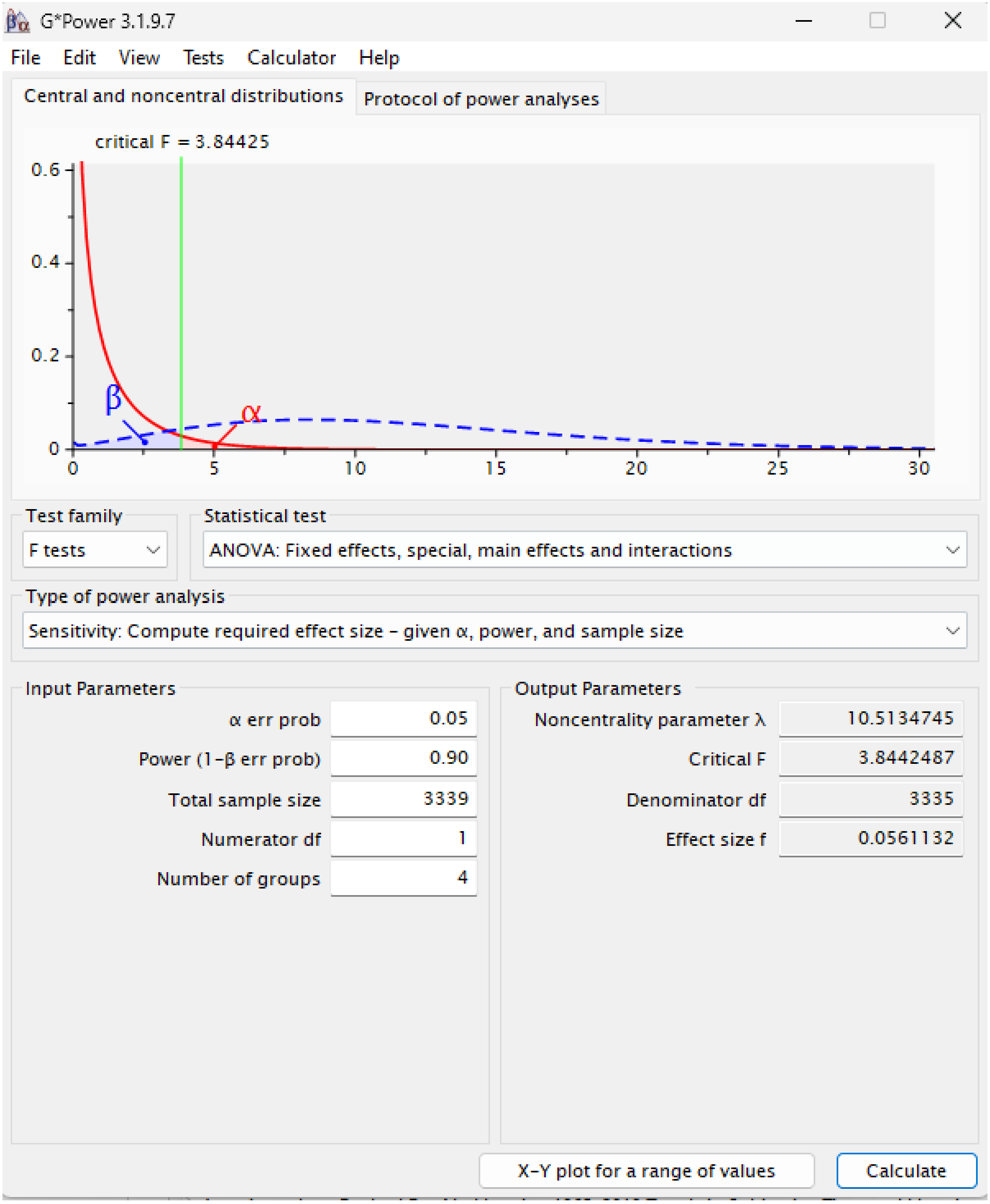


Figure S2.1. Minimum detectable effect size in C&P's original study for the planned contrast (H1: moral/distant + immoral/recent vs. moral/recent + immoral/distant): d = .651.

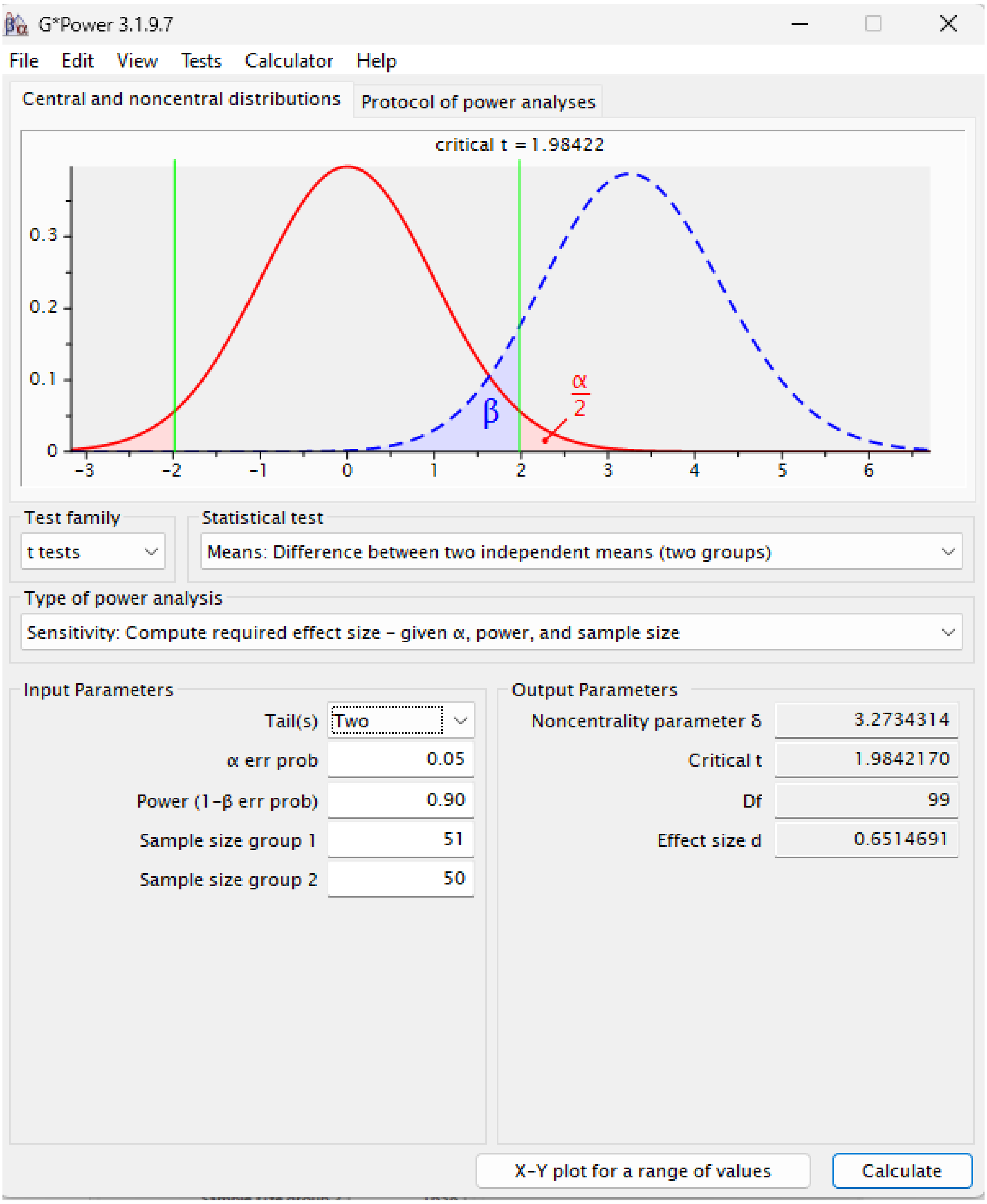
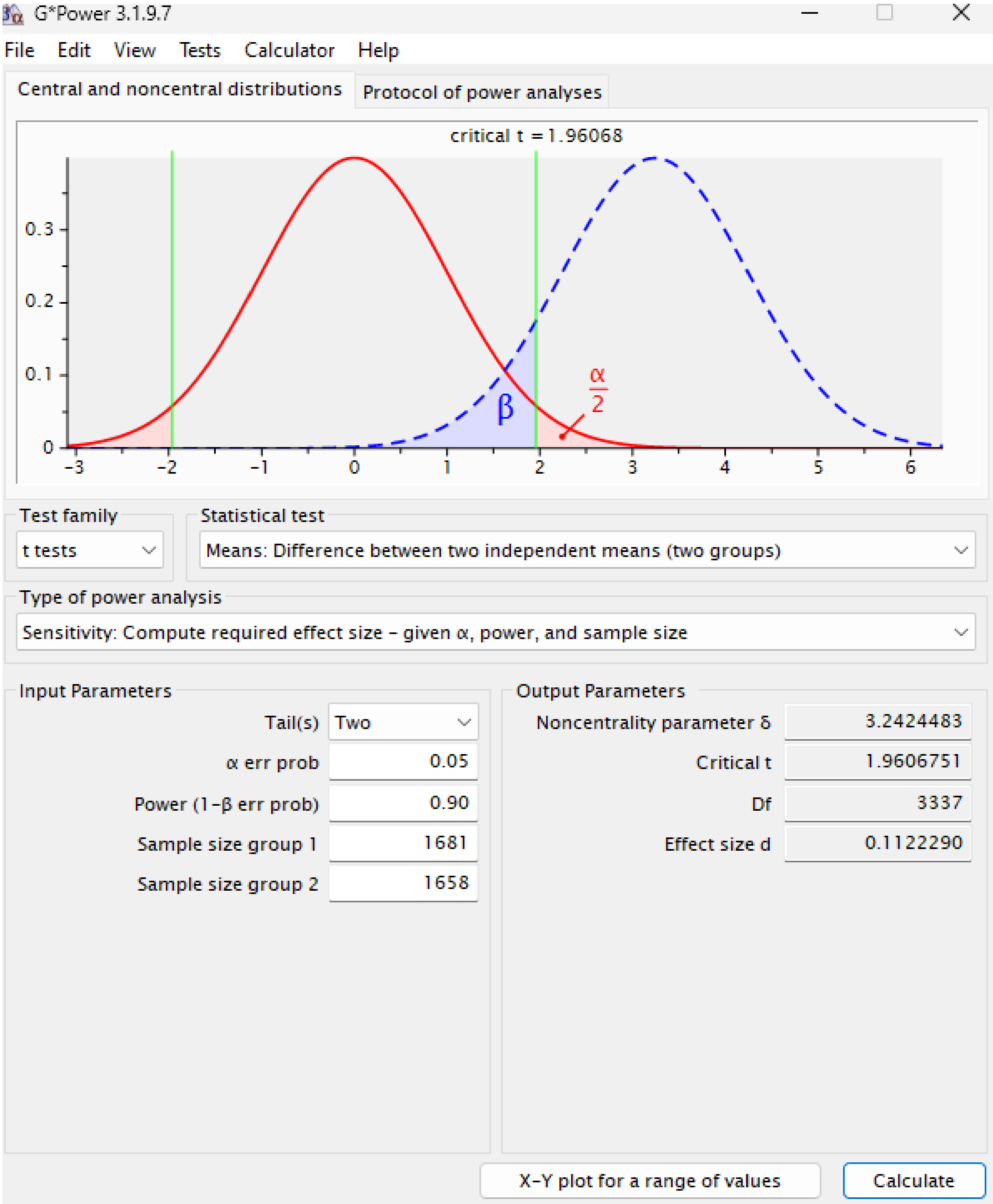


Figure S2.2. Minimum detectable effect size for the planned contrast (H1: moral/distant + immoral/recent vs. moral/recent + immoral/distant): d = .112.



4.2. Minimum detectable effect sizes for the extension analyses.

Figure S3. Minimum detectable effect size for the planned contrast (H2a: moral/distant + immoral/recent vs. neutral/recent + neutral/distant): d = .111

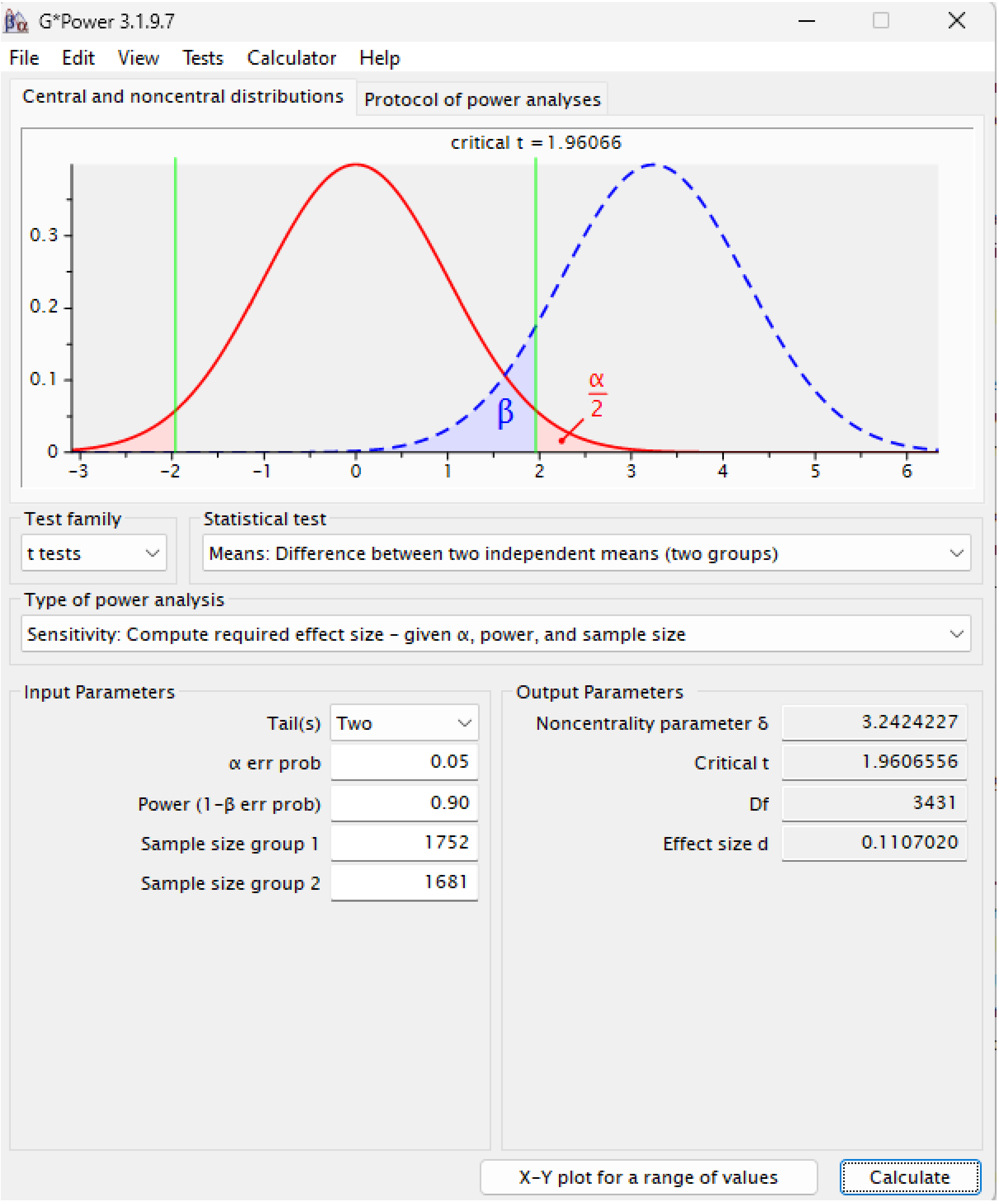


Figure S4. Minimum detectable effect size for the planned contrast (H2b: moral/recent + immoral/distant vs. neutral/recent + neutral/distant): d = .111

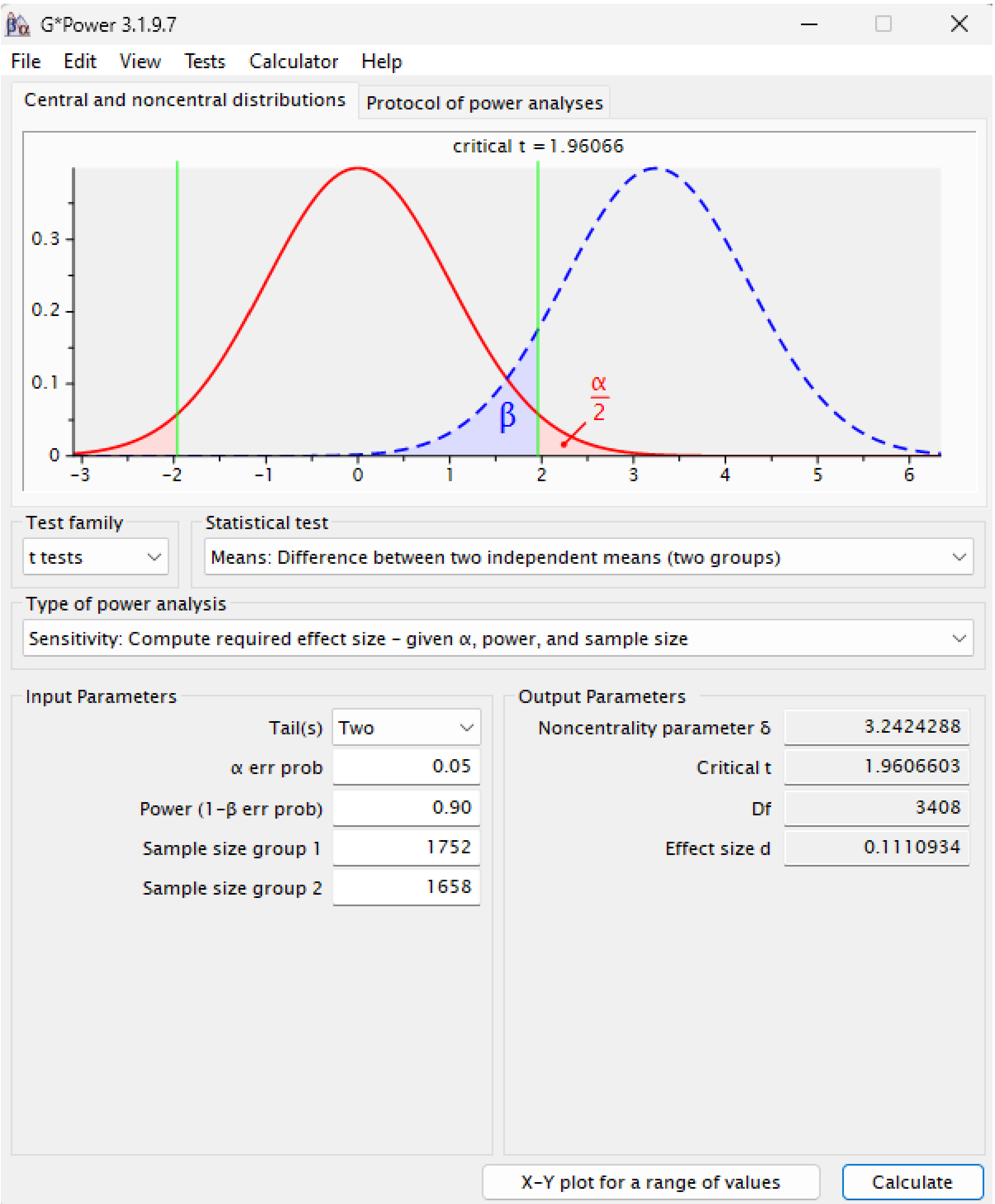


Figure S5. Minimum detectable effect size for moral licensing (H2c: moral/recent vs. neutral/recent): d =.200

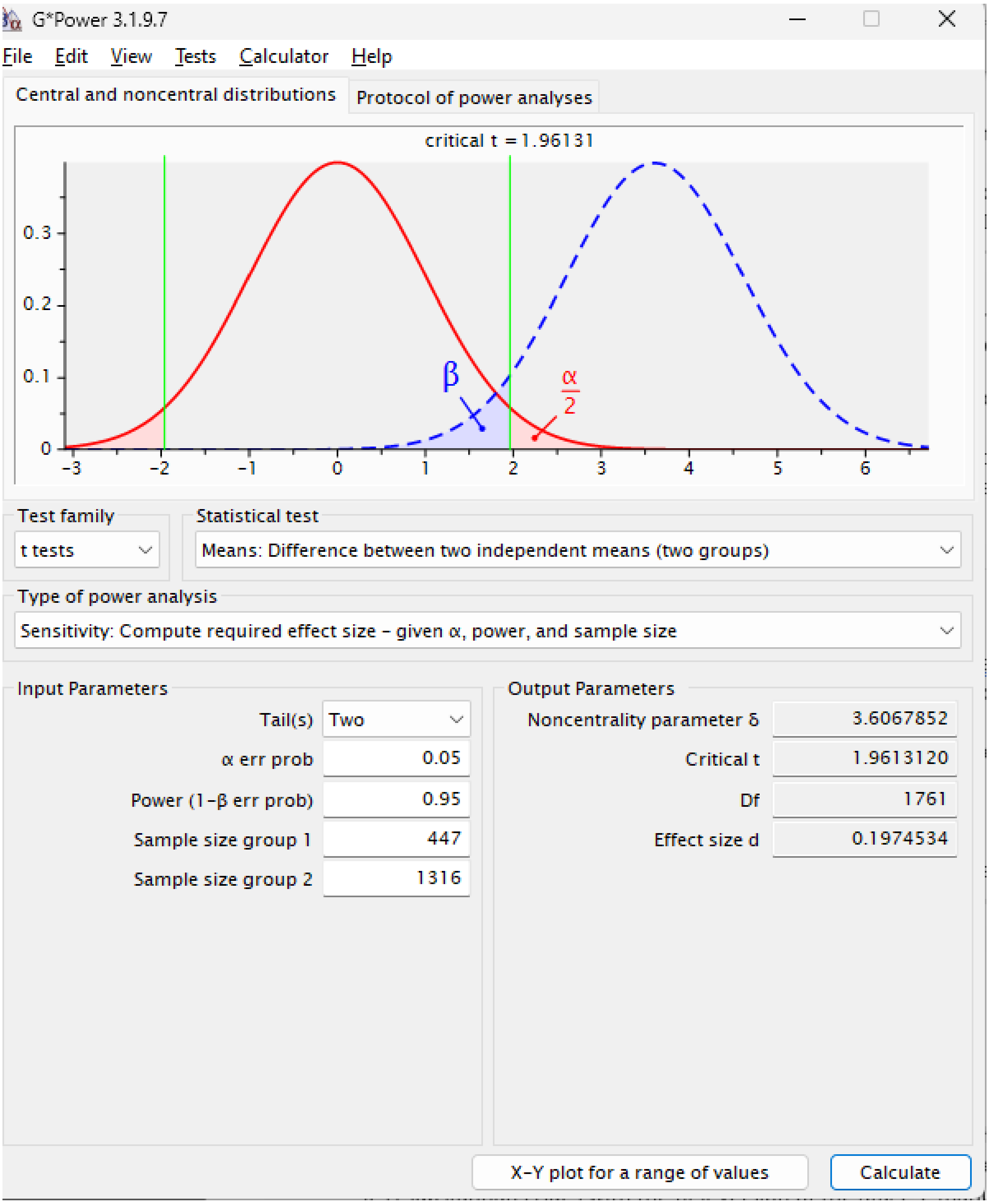


Figure S6. Minimum detectable effect size for moral cleansing (H2d: immoral/recent vs. neutral/recent): d = .196

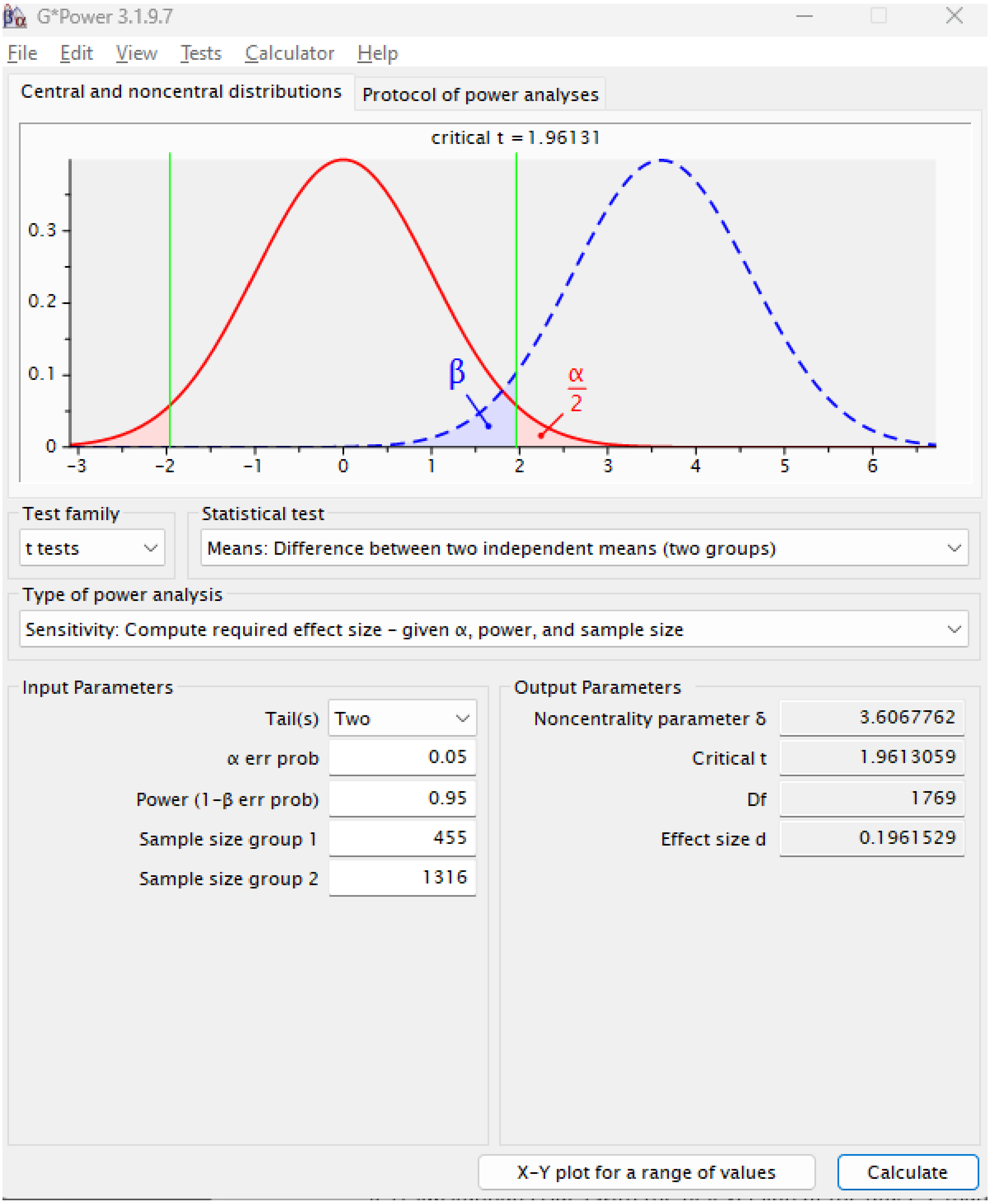


Figure S7. Minimum detectable effect size for positive moral consistency (moral/distant vs. neutral/distant): d = .201

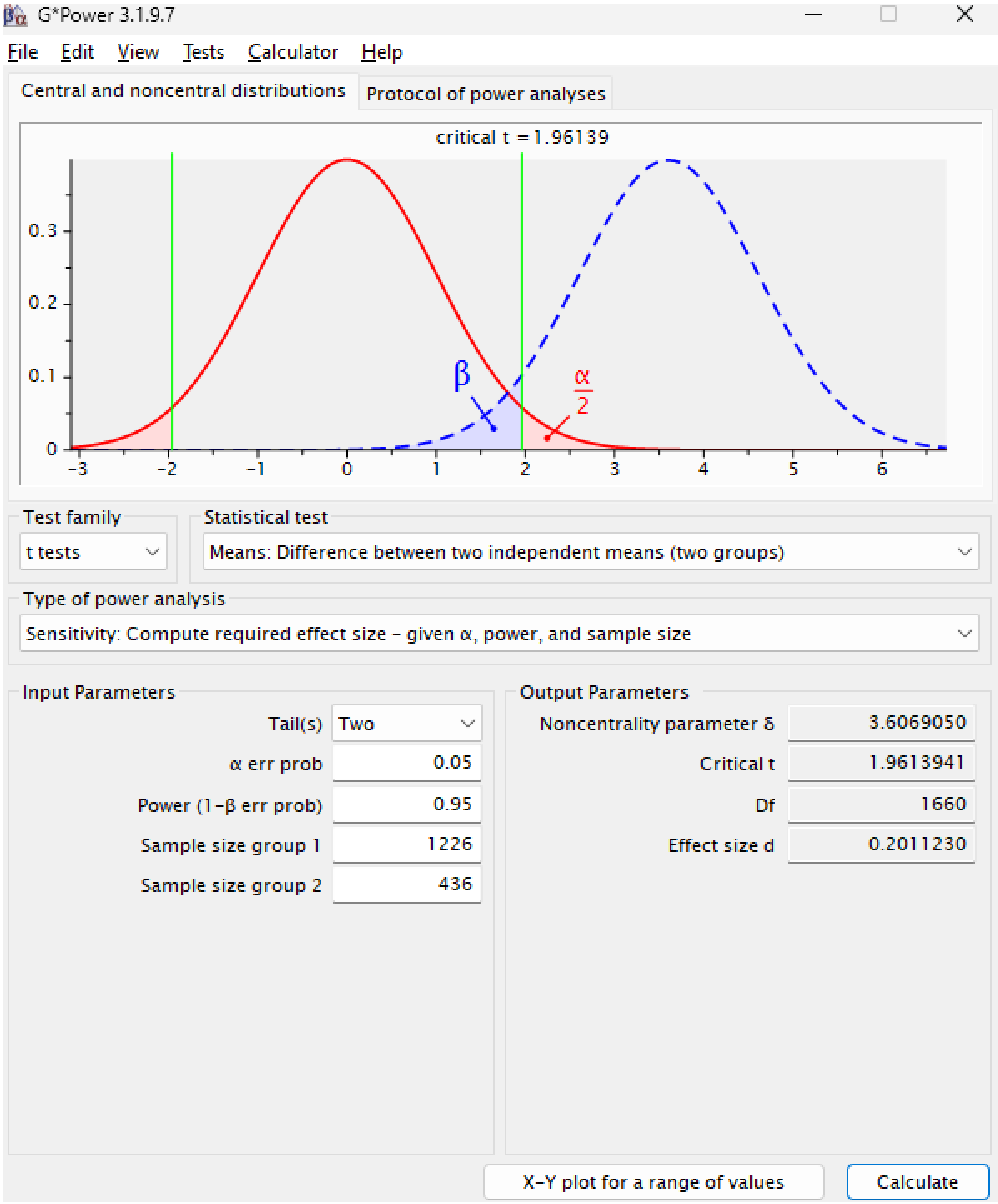
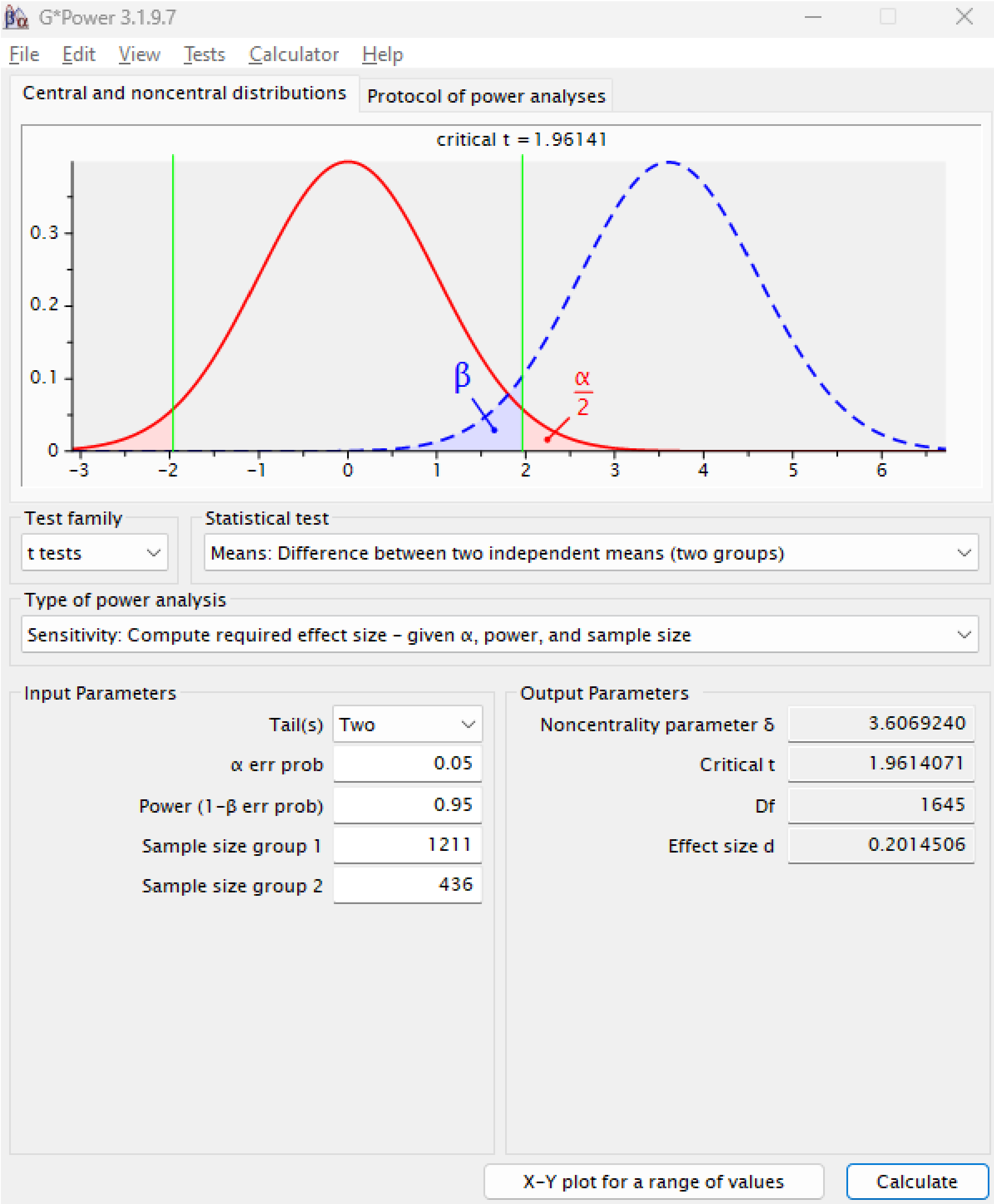


Figure S8. Minimum detectable effect size for negative moral consistency (immoral/distant vs. neutral/distant): d = .201



5. Comparing test statistics between the original and the replication

5.1. Manipulation checks

On a subset of the four original conditions of C&P, we first conducted a 2 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on recalled event positivity. The results revealed the expected effect of event valence on event positivity, F (1, 3335) = 6144.85, p < .001, $\eta^2{}_p = .648$. Event distance did not moderate the effect of event valence, F (1, 3335) = 0.079, p = .778, $\eta^2{}_p < .001$. Participants instructed to recall and write about a moral event (M = 5.81, SD = 1.30) reported their event to be significantly more positive than those in the immoral conditions (M = 1.81, SD = 1.31), t (3335) = 78.4, p_{Tukey} < 0.001, d = 3.06, 95% CI [2.95, 3.16].

Table S4A: Comparison of test statistics of the Event Valence manipulation check between C&P and the current study.

Predictor	Study	df	F	p	eta-squared
Valence	C&P	1	287.96	<.001	.75
	Current	1	6144.85	<.001	.648
Distance	C&P	NA			
	Current	1	2.97	.085	.00
Valence/Distance	C&P	1	.70	.406	.01
	Current	1	.08	.78	.00

A 2 (Event Valence: Moral vs. Immoral) X 2 (Event Distance: Recent vs. Distant) between-participants ANOVA on recalled event distance revealed the expected effect of event distance on perceived distance, F (1, 3335) = 1415.76, p < .001, $\eta^2{}_p = .298$. Event valence did not moderate the effect of event distance, F (1, 3335) = 0.061, p = .805, $\eta^2{}_p < .001$. Participants instructed to recall and write about a distant event (M = 4.06, SD = 1.87) perceived the event to be significantly more distant than participants in the recent conditions (M = 1.55, SD = 1.15), t (3335) = 37.6, p_{Tukey} < 0.001, d = 1.47, 95% CI [1.38, 1.55].

Table S4B: Comparison of test statics of the Event Distance manipulation check between C&P and the current study.

Predictor	Study	df	F	p	eta-squared
Valence	C&P	NA			
	Current	1	4.54	.033	.001
Distance	C&P	1	83.72	<.001	.47
	Current	1	1415.76	<.001	.298
Valence/Distance	C&P	1	1.43	.236	.02
	Current	1	.061	.805	<.001

5.2. Dependent variables

Next, we compare exact test statistics for the dependent variables between C&P and the current study.

Table S4C: Comparison of test statics of willingness-to-volunteer (WTV) between C&P and the current study.

Predictor	Study	df	F	p	eta-squared
Valence	C&P	1	.13	.716	<.01
	Current	1	20.33	<.001	.006
Distance	C&P	1	.07	.798	<.01
	Current	1	3.817	0.051	0.001
Valence X Distance	C&P	1	5.37	.023	.05
	Current	1	0.176	0.675	<.001

Table S4D: Comparison of test statics of willingness-to-help (WTH) between C&P and the current study.

Predictor	Study	df	F	p	eta-squared
Valence	C&P	1	1.26	.265	.01
	Current	1	5.67	.018	.002
Distance	C&P	1	.36	.550	<.01
	Current	1	.522	.470	<.001
Valence X Distance	C&P	1	5.97	.016	.06
	Current	1	1.052	.305	<.001

6. Statistical Considerations

6.1. Comparing test statistics between including versus excluding participants who responded incorrectly to the additional Event Distance manipulation check.

Around 20% of participants reported that the incident they recalled was "between one week and one year", thereby potentially undermining the Event Distance manipulation.

Table S5A. Responses to the specific manipulation check question for Event Distance.			
	within last week	in between a week and a year	more than a year ago
Recent (total = 2218)	2012; 90.7%	160; 7.2%	46; 2.1%
Distant (total = 2873)	190; 6.6%	851; 29.6%	1832; 63.8

So, we conducted all our analyses on a dataset where we excluded those participants. Overall, the results do not significantly change.

Table S5B. Comparing test statistics between including versus excluding participants who responded incorrectly to the additional Event Distance manipulation check.			
		Final Dataset	Dataset excluding wrong responders
Replication	Willingness to volunteer	Event Valence: F (1, 3335) = 20.332, p <.001, $\eta^2{}_p = .006$	Event Valence: F (1, 2325) = 17.80, p <.001, $\eta^2{}_p = .008$
		Event Distance: F (1, 3335) = 3.817, p = .051, $\eta^2{}_p = .001$	Event Distance: F (1, 2325) = .46, p = .497, $\eta^2{}_p < .001$
		Event Valence X Event Distance interaction: F (1, 3335) = .176, p = .675, $\eta^2{}_p < .001$.	Event Valence X Event Distance interaction: F (1, 2492) = .02, p = .886, $\eta^2{}_p < .001$.
		Moral vs. Immoral post hoc: t (2327) = 4.51, p < .001, d = .176 [.099, .252]	Moral vs. Immoral post hoc: t (2327) = 4.41, p < .001, d = .183 [.101, .264]
	Willingness to help	Event Valence: F (1, 3335) = 5.64, p = 018, $\eta^2{}_p = .002$	Event Valence: F (1, 2325) = 2.36, p = .125, $\eta^2{}_p = .001$
		Event Distance: F (1, 3335) = .519, p = .47, $\eta^2{}_p < .001$	Event Distance: F (1, 2325) = .015, p = .901, $\eta^2{}_p < .001$
		Event Valence X Event Distance interaction: F (1, 3335) = 1.052, p = .305, $\eta^2{}_p < .01$.	Event Valence X Event Distance interaction: F (1, 2325) = .884, p = .347, $\eta^2{}_p < .001$.

Extension	Willingness to volunteer	Event Valence: F (2, 5085) = 15.54, p < .001, $\eta^2{}_p$ = .006	Event Valence: F (2, 3838) = 11.67, p < .001, $\eta^2{}_p$ = .006
		Event Distance: F (1, 5085) = 1.16, p = .281, $\eta^2{}_p$ = .000.	Event Distance: F (1, 3838) = 0.501, p = .479, $\eta^2{}_p < .001$
		F (2, 5085) = 1.69, p = .184, $\eta^2{}_p$ = .001	Event Valence X Event Distance interaction: F (2, 3838) = 0.018, p = .982, $\eta^2{}_p < .001$
		Moral vs. Neutral post hoc: t (5085) = 5.15, p < .001, d = .201 [.124, .278]	Moral vs. Neutral post hoc: t (3838) = 4.21, p < .001, d = .191 [.102, .279]
		Immoral vs. Neutral post hoc: t (5085) = 0.730, p = .746, d = .028 [-.048, .105]	Immoral vs. Neutral post hoc: t (3838) = 0.207, p = .977, d = .009 [-.079, .097]
	Willingness to help	Event Valence: F (2, 5085) = 9.53, p < .001, $\eta^2{}_p$ = .004	Event Valence: F (2, 3838) = 8.61, p < .001, $\eta^2{}_p$ = .004
		Event Distance: F (1, 5085) = 1.20, p = .273, $\eta^2{}_p < .000$	Event Distance: F (1, 3838) = 0.020, p = .888, $\eta^2{}_p < .001$
		Event Valence X Event Distance interaction: F (2, 5085) = .555, p = .574, $\eta^2{}_p < .001$	Event Valence X Event Distance interaction: F (2, 3838) = 0.787, p = .455, $\eta^2{}_p < .001$
		Moral vs. Neutral post hoc: t (5085) = 4.36, p < .001, d = .171 [.094, .247]	Moral vs. Neutral post hoc: t (3838) = 4.10, p < .001, d = .186 [.097, .274]
		Immoral vs. Neutral post hoc: t (5085) = 2.05, p = .101, d = .080 [.003, .156]	Immoral vs. Neutral post hoc: t (3838) = 2.69, p = .020, d = .121 [.033, .209]

6.2 Robust linear regressions

Table S5C. Outputs from robust linear regressions.

	Helping Intention		Volunteering Intention	
Predictors	*Estimates*	*std. Beta*	*Estimates*	*std. Beta*
(Intercept)	5.58 *** (-0.01)	-0.01 (-0.07 – 0.04)	4.40 *** (-0.01)	-0.01 (-0.07 – 0.05)

Distance2 [Distant]	0.05	0.05	-0.03	-0.02
	(0.05)	(-0.06 – 0.16)	(-0.02)	(-0.14 – 0.09)
Valence [Moral]	0.15 **	0.14	0.19 *	0.14
	(0.14)	(0.05 – 0.24)	(0.14)	(0.03 – 0.25)
Valence [Immoral]	0.04	0.04	-0.06	-0.04
	(0.04)	(-0.07 – 0.14)	(-0.04)	(-0.15 – 0.07)
Distance2 [Distant] X	-0.02	-0.02	0.15	0.11
Valence [Moral]	(-0.02)	(-0.17 – 0.13)	(0.11)	(-0.05 – 0.27)
Distance2 [Distant] X	0.05	0.04	0.14	0.10
Valence [Immoral]	(0.04)	(-0.11 – 0.20)	(0.10)	(-0.06 – 0.26)
Observations	5091		5091	
R^2 / R^2 adjusted	0.005 / 0.004		0.009 / 0.008	

* *p<0.05* ** *p<0.01* *** *p<0.001*

7. Heterogeneity in prosocial intentions

It may be that participants' responses in willingness to volunteer and the four scenarios of willingness to help may have a common source of error terms. So, it may be worth examining them as five within-participant measures to get more precise estimates of the effects. Accordingly, we perform two linear mixed-effects models using the lme4 package in R (Bates et al., 2014) with fixed effects of event valence, event distance, and event valence X event distance, with random intercepts of participant ID and Prosociality Category (willingness to volunteer and the five scenarios).

We find only event valence to be a significant predictor. First, compared to the reference level of neutral event, participants reported higher prosocial intentions in the moral event condition (β = .11, SE = .04, p = .007). Second, compared to the reference level of moral event, participants reporter lower prosocial intentions in both the neutral (β = -.11, SE = .04, p = .007) and immoral (β = -.09, SE = .04, p = .019) conditions. Please see Tables S6A and S6B for full outputs from the models.

Table S6A. Outputs from linear mixed-effects model. We combined the composite measure of Volunteering intention and the Helping intentions of the four helping scenarios to have five measures of "Prosocial Intentions" per participant. (*Reference levels: Valence = Neutral, Distance = Recent)*

Predictors	*Estimates (SE)*	*std. Beta (SE)*	*Estimates CI*	*standardized CI*	*Statistic*	*p*
(Intercept)	5.28 (0.36)	-0.05 (0.22)	4.58 – 5.98	-0.48 – 0.38	14.85	**<0.001**
Valence (Moral)	0.17 (0.06)	0.11 (0.04)	0.05 – 0.30	0.03 – 0.18	2.68	**0.007**
Valence (Immoral)	0.02 (0.06)	0.01 (0.04)	-0.10 – 0.15	-0.06 – 0.09	0.37	0.710
Distance (Distant)	-0.03 (0.05)	-0.02 (0.03)	-0.13 – 0.08	-0.08 – 0.05	-0.52	0.600
Valence (Moral) X Distance (Distant)	0.04 (0.08)	0.03 (0.05)	-0.10 – 0.19	-0.06 – 0.12	0.58	0.561
Valence (Immoral) X Distance (Distant)	0.10 (0.07)	0.06 (0.05)	-0.05 – 0.25	-0.03 – 0.15	1.32	0.188

Random Effects

σ^2	1.52
τ_{00} participant ID	0.62
τ_{00} Prosocial Category	0.62
ICC	0.45
N Prosocial Category	5
N participant ID	5091
Observations	25455
Marginal R^2 / Conditional R^2	0.003 / 0.451

Table S6B. Outputs from linear mixed-effects model. We combined the composite measure of Volunteering intention and the Helping intentions of the four helping scenarios to have five measures of "Prosocial Intentions" per participant. (*Reference levels: Valence = Moral, Distance = Recent)*

Predictors	*Estimates (SE)*	*std. Beta*	*Estimates CI*	*standardized CI*	*Statistic*	*p*
(Intercept)	5.45 (0.36)	0.06 (0.22)	4.76 – 6.15	-0.37 – 0.49	15.34	**<0.001**
Valence (Neutral)	-0.17 (0.06)	-0.11 (0.04)	-0.30 – -0.05	-0.18 – -0.03	-2.68	**0.007**
Valence (Immoral)	-0.15 (0.06)	-0.09 (0.04)	-0.27 – -0.02	-0.17 – -0.01	-2.34	**0.019**
Distance (Distant)	0.02 (0.05)	0.01 (0.03)	-0.09 – 0.12	-0.05 – 0.07	0.30	0.766
Valence (Neutral) X Distance (Distant)	-0.04 (0.08)	-0.03 (0.05)	-0.19 – 0.10	-0.12 – 0.06	-0.58	0.561
Valence (Immoral) X Distance (Distant)	0.05 (0.07)	0.03 (0.05)	-0.09 – 0.20	-0.06 – 0.12	0.73	0.463
Random Effects						
σ^2	1.52					

τ_{00} Participant ID	0.62
τ_{00} Prosocial Category	0.62
ICC	0.45
N Prosocial Category	5
N Participant ID	5091
Observations	25455
Marginal R^2 / Conditional R^2	0.003 / 0.451

8. ANCOVAs with the importance of moral identity internalization and symbolization as controls.

8.1 Willingness-to-volunteer: Direct outputs from jamovi

ANCOVA - Volunteering_Intention

	Sum of Squares	df	Mean Square	F	p	$\eta^2 p$
Distance2	3.11	1	3.11	1.94	0.164	0.000
Valence	52.83	2	26.42	16.50	< .001	0.006
Moral_Internalization	1303.57	1	1303.57	813.98	< .001	0.138
Moral_Symbolization	71.54	1	71.54	44.67	< .001	0.009
Distance2 ✻ Valence	16.08	2	8.04	5.02	0.007	0.002
Residuals	8140.26	5083	1.60			

Post Hoc Comparisons - Distance2 ✻ Valence

Comparison									
Distance2	Valence		Distance2	Valence	Mean Difference	SE	df	t	p Tukey
Recent	Moral	-	Recent	Immoral	0.1917	0.0843	5083	2.274	0.205
		-	Recent	Neutral	0.1075	0.0693	5083	1.550	0.632
		-	Distant	Moral	-0.1829	0.0699	5083	-2.615	0.094
		-	Distant	Immoral	0.0870	0.0701	5083	1.242	0.816
		-	Distant	Neutral	0.2266	0.0852	5083	2.660	0.084
	Immoral	-	Recent	Neutral	-0.0842	0.0689	5083	-1.223	0.826
		-	Distant	Moral	-0.3746	0.0695	5083	-5.391	< .001
		-	Distant	Immoral	-0.1046	0.0696	5083	-1.503	0.662
		-	Distant	Neutral	0.0349	0.0848	5083	0.412	0.998

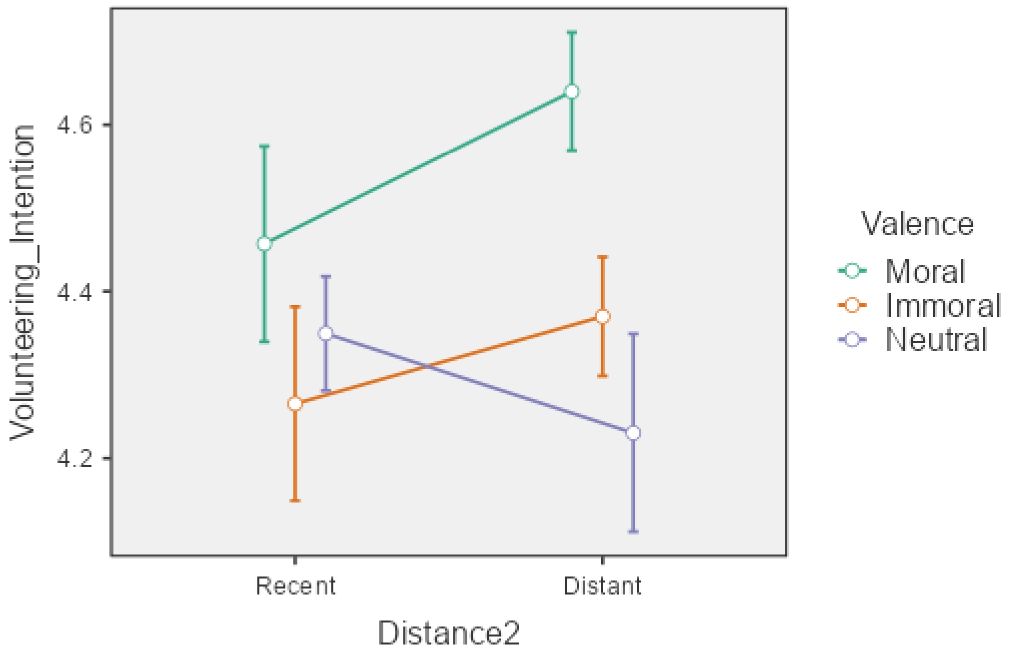
Post Hoc Comparisons - Distance2 ✻ Valence

Comparison									
Distance2	**Valence**		**Distance2**	**Valence**	**Mean Difference**	**SE**	**df**	**t**	**p Tukey**
	Neutral	-	Distant	Moral	-0.2904	0.0502	5083	-5.780	< .001
		-	Distant	Immoral	-0.0204	0.0504	5083	-0.406	0.999
		-	Distant	Neutral	0.1191	0.0699	5083	1.703	0.530
Distant	Moral	-	Distant	Immoral	0.2700	0.0513	5083	5.265	< .001
		-	Distant	Neutral	0.4095	0.0706	5083	5.803	< .001
	Immoral	-	Distant	Neutral	0.1396	0.0707	5083	1.975	0.357

Note. Comparisons are based on estimated marginal means

Estimated Marginal Means

Distance2 ✻ Valence



8.2 Willingness-to-help: Direct outputs from jamovi

ANCOVA - Helping_Intention

	Sum of Squares	**df**	**Mean Square**	**F**	**p**	**η^2p**
Distance2	1.572	1	1.572	1.735	0.188	0.000
Valence	17.373	2	8.686	9.587	< .001	0.004
Moral_Internalization	612.489	1	612.489	676.034	< .001	0.117
Moral_Symbolization	8.139	1	8.139	8.983	0.003	0.002
Distance2 ✻ Valence	0.810	2	0.405	0.447	0.639	0.000
Residuals	4605.211	5083	0.906			

Post Hoc Comparisons - Distance2 ✻ Valence

Comparison									
Distance2	**Valence**		**Distance2**	**Valence**	**Mean Difference**	**SE**	**df**	**t**	**p_{Tukey}**
Recent	Moral	-	Recent	Immoral	0.1149	0.0634	5083	1.812	0.458
		-	Recent	Neutral	0.1587	0.0521	5083	3.043	0.028
		-	Distant	Moral	-0.0239	0.0526	5083	-0.454	0.998
		-	Distant	Immoral	0.0347	0.0527	5083	0.659	0.986
		-	Distant	Neutral	0.1429	0.0641	5083	2.231	0.224
	Immoral	-	Recent	Neutral	0.0438	0.0518	5083	0.847	0.959
		-	Distant	Moral	-0.1388	0.0523	5083	-2.655	0.085
		-	Distant	Immoral	-0.0802	0.0524	5083	-1.531	0.644
		-	Distant	Neutral	0.0281	0.0638	5083	0.440	0.998
	Neutral	-	Distant	Moral	-0.1826	0.0378	5083	-4.832	< .001
		-	Distant	Immoral	-0.1240	0.0379	5083	-3.271	0.014

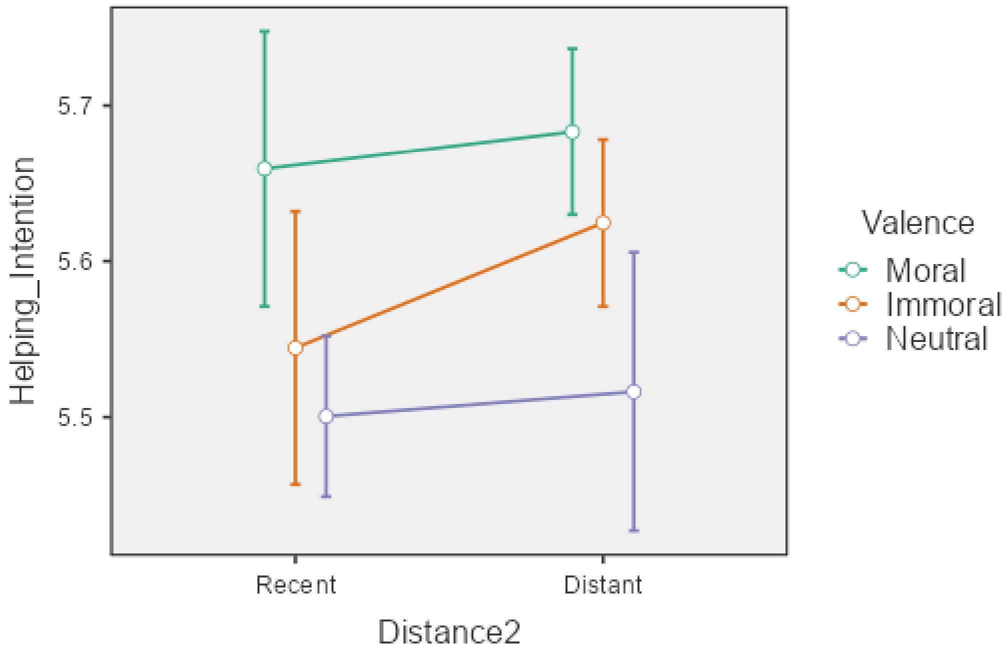
Post Hoc Comparisons - Distance2 ✻ Valence

Comparison									
Distance2	Valence		Distance2	Valence	Mean Difference	SE	df	t	p Tukey
		-	Distant	Neutral	-0.0158	0.0526	5083	-0.300	1.000
Distant	Moral	-	Distant	Immoral	0.0586	0.0386	5083	1.519	0.652
		-	Distant	Neutral	0.1668	0.0531	5083	3.143	0.021
	Immoral	-	Distant	Neutral	0.1082	0.0532	5083	2.036	0.322

Note. Comparisons are based on estimated marginal means

Estimated Marginal Means

Distance2 ✻ Valence



9. References

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

The jamovi project. (2022). *Jamovi.* (2.3) [Computer software]. https://www.jamovi.org

Qualtrics (Version 2022). (2022). Qualtrics. https://www.qualtrics.com

R Core Team. (2022). *R: A language and environment for statistical computing.* (4.1.1). R Foundation for Statistical Computing. https://www.r-project.org/

Supplemental Materials for Article 2

For the manuscript

Business-Size Bias in Moral Concern:
People are More Dishonest Against Big than Small Organizations

The supplemental materials are included to detail all measures, additional analyses, and data visualizations that support the current research.

Table of Contents

1. Pilots

Pilot 1A

The goal of Pilot 1A was to test our intuition on whether there were differences in normative judgments regarding the importance of being truthful to big vs. small businesses across a range of industries.

Method

In a two-condition between-participants design, we asked 300 Prolific participants (Mage=32.2, SD = 10.8, 49.7% female) “(As a customer, how) How important is it to be truthful towards the following?” on an 11-point scale from 0 (Not at all) to 10 (Very much).

Participants were presented with 8 targets in a matrix table: 1. Big banks, 2. Local banks, 3. Big tech companies, 4. Tech startups, 5. Large insurance companies, 6. Local insurance companies, 7. Large online retailers, and 8. Small online retailers. The order of presenting these eight was randomly varied across participants. Different signals of sizes (big: big, large; small: small, startup, local) were used to make the big vs. small distinctions less obvious.

Results

First, we conducted a repeated-measures ANOVA with the eight targets as within-participant variables and customer framing (yes vs. no) as the between-participant variable. The results revealed that only the within-subject variable of target was significant, $F\ (7, 2065) = 86.4$, $p < .001$, partial $\eta 2 = .227$. neither the interaction of customer framing and target, $F\ (7, 2065) = 1.11$, $p = .335$, partial $\eta 2 = .004$, nor the between-participant variable of customer framing, $F\ (7, 295) = .261$, $p = .610$, partial $\eta 2 = .001$.

Tukey’s post hoc tests revealed that across industries, participants reported it was less important to be truthful to the big (vs. small) target in each pair (e.g., large online retailers vs. small online retailers) were significantly different, with all t’s > 5.1 and all Tukey′s p < .001, except the insurance pair, $t(295) = 2.96$, Tukey’s $p = .065$. Further, a paired sample t-test of the difference between means of big vs. small firms showed that on average, participants reported it was less important to be truthful to big ($M = 7.11$, $SD = 2.40$) than small ($M = 7.84$, $SD = 1.81$) businesses, $t\ (296) = -9.01$, $p < .001$, $d = .523$, 95% CI [,401, .644]. For detailed pairwise comparisons between big vs. small businesses across industries, please see Table SA.

Table SA. Pairwise comparisons between big vs. small businesses across industries

Industry	Big M (SD)	Small M (SD)	Paired sample t-tests
Bank	7.89 (2.67)	8.41 (1.97)	t (296) = -5.10, p < .001, d = -.296, 95% CI [-.412, -.179]
Tech	6.02 (3.13)	6.86 (2.70)	t (296) = -7.14, p < .001, d = -.414, 95% CI [-.532, -.295]
Insurance	8.98 (2.48)	8.35 (2.09)	t (296) = -2.96, p = .003, d = -.172, 95% CI [-.286, -.057]
Retailer	6.42 (2.90)	7.72 (2.32)	t (296) = -9.29, p < .001, d = -.539, 95% CI [-.660, -.417

Pilot 1B

The goal of Pilot 1B was to test our intuition on whether there were differences in the moral acceptability of dishonest behavior toward big vs. small businesses.

Method

We asked a nationally representative sample of 498 Prolific participants based in the USA (Mage = 45.1, SD = 16.0, 51.0% female) to what extent they thought it was acceptable to conceal or falsify information for personal benefit against big vs. small businesses.

First, depending on the conceal or falsify conditions, participants were shown the text, “We may sometimes conceal (falsify) information for our own benefit. For all questions, assume that it is impossible to detect any concealing (falsifying) and so there can be no punishment.” Then, they were asked to confirm they understood.

Then, depending on the customer framing (yes vs. no) and dishonesty (conceal vs. falsify) conditions, participants were asked “(As a customer, which) Which is more okay to conceal information from/ falsify information to for our own benefit?”.

Participants were presented with three options. In the banking industry conditions, the options were: (1) A large bank, (2) A local bank, and (3) Both equally not okay. In the tech industry conditions, the options were: (1) A big tech company, (2) A tech startup, and (3) Both equally not okay. Then, they were asked to write their reasons for their choice in 1-2 sentences in a text entry response box.

Results

A series of Chi-square tests showed that choices were not significantly different across customer framing (yes vs. no) conditions, chi-square (2, 498) = 1.92, p = .383. The choices were marginally significant across the target (industry: bank vs. tech), chi = (2, 498) = 4.85, p = .089.

Table SB. Number and percentage of participants choosing a participant potential victim as acceptable to cheat.

	Bank	Tech
Big	65 (26.7%)	65 (25.5%)
Small	0 (0%)	5 (2.0%)
Neither	178 (73.3%)	185 (72.5%)

A chi-square test of proportions of big vs. small choices, merged across industries and customer framing, showed that people 26.1% of people stated it acceptable to cheat a big business whereas only 1% of people stated it is acceptable to cheat a small business, suggesting a highly significant difference of 25.1 percentage points (95% CI [17.56, 33.13], chi (135) = 36.17, p < .001.

2. All measurements used in the studies

Table S1A. List of measures – Study 1a			
Variable	# of items (a)	Item	Scale
Return Likelihood	1	How likely would you be to return the shoes by expressing dissatisfaction?	-100 (No way) to 100 (For sure)
Profitableness	1	How profitable does Nimbus seem?	-100 (Not at all profitable) to 100 (Very profitable)
Powerfulness	1	How powerful does Nimbus seem?	0 (Not at all powerful) to 100 (Very powerful)
Feeling	1	To what extent does Nimbus seem capable of emotions, feelings, and experiences?	0 (Not at all capable) to 100 (Completely capable)
Morality	1	To what extent does Nimbus seem "morally bad" or "morally good"?	-100 (Completely bad) to 100 (Completely good)
Gender, Age		Pulled from Prolific's records.	

Table S1B. List of measures – Study 1b			
Variable	# of items (a)	Item	Scale
Return Likelihood	1	How likely would you be to return the shoes by expressing dissatisfaction?	-100 (No way) to 100 (For sure)
Likability	1	How likable does Nimbus seem?	-100 (Not at all likable) to 100 (Very likable)
Vulnerability	1	How vulnerable does Nimbus seem?	-100 (Not at all vulnerable) to 100 (Very vulnerable)
Gender, Age		Pulled from Prolific's records.	

Table S1C. List of measures – Study 1c			
Variable	Number of items (Cronbach's alpha)	Item	Scale
Return Likelihood	1	How likely would you be to return the shoes by expressing dissatisfaction?	-100 (No way) to 100 (For sure)

Likability	1	How likable does Nimbus seem?	-100 (Not at all likable) to 100 (Very likable)
Perceived Vulnerability	3 (a = .652)	How vulnerable does Nimbus seem?	-100 (Not at all vulnerable) to 100 (Very vulnerable)
		How profitable does Nimbus seem?	-100 (Not at all profitable) to 100 (Very profitable)
		How powerful does Nimbus seem?	-100 (Very powerless) to 100 (Very powerful)
Morality	1	How "morally bad" or "morally good" does Nimbus seem?	-100 (Completely bad) to 100 (Completely good)
Humanness	1	How human does Nimbus seem?	-100 (Not at all human) to 100 (Completely human)
Harm-Benefit	1	How much might Nimbus "harm society" or "benefit society"?	-100 (Mostly harms) to 100 (Mostly benefits)
Gender, Age		Pulled from Prolific's records.	

Table S2A. List of measures – Study 2a			
Self-reported correct guess 1	1	Did you guess correctly? (You can receive £X {where X = 1-6 depending on the randomly generated die roll} from Mike's store (EveryDayMart).	Binary choice: YES/ NO
Self-reported correct guess 2	1	Did you guess correctly? Mike's store (EveryDayMart) will pay the bonus to 10 participants (randomly chosen) for this extra round as well.	Binary choice: YES/ NO
Variable	# of items (a)	Item	Scale
Writing Clarity	1	To what extent do you think the sample text is Clear	1 (Not at all) to 7 (Very much)
		Well-written	
		Precise	
		Without errors	

Organizational Perceptions page 1	1	To what extent does Mike's store (EveryDayMart) seem Caring	1 (Not at all) to 7 (Very much)
		Ethical	
		Profitable	
		Likable	
Organizational Perceptions page 2	1	To what extent does Mike's store (EveryDayMart) seem Empathetic	1 (Not at all) to 7 (Very much)
		Harmful	
		Vulnerable	
		Trustworthy	
Actions	1	To what extent would you be willing to visit/ buy from/ refer Mike's Store (EveryDayMart) Visit	1 (Not at all) to 7 (Very much)
		Buy from	
		Refer	
Morality	1	To what extent does Nimbus seem "morally bad" or "morally good"?	-100 (Completely bad) to 100 (Completely good)
Beliefs about others (First roll)	1	What percentage of participants do you think would state that they guessed correctly?	0 to 100
Beliefs about others (Second roll)	1	What percentage of participants do you think would state that they guessed correctly on this extra round?	Guess Percentage beliefs (First roll)
Gender, Age		Pulled from Prolific's records.	

Table S2B. List of measures – Study 2b			
Self-reported correct guess		Did you guess correctly? (You can receive £X {where X = 1-6 depending on the randomly generated die roll} from Indigo Hotels (Indigo Hotels & Resorts).	Binary choice: YES/ NO.
Variable	# of items (a)	Item	Scale
Writing Clarity	1	To what extent is Indigo's "About Us" text unclear or clear?	-5 (Very unclear) to 5 (Very clear)
Morality	1	To what extent does Indigo Hotel (Indigo Hotels & Resorts) seem unethical or ethical?	-5 (Very unethical) to 5 (Very ethical)
Size	1	To what extent does Indigo Hotel (Indigo Hotels & Resorts) seem small or large?	-5 (Very small) to 5 (Very large)
Weak-Strong	1	To what extent does Indigo Hotel (Indigo Hotels & Resorts) seem weak or strong?	-5 (Very weak) to 5 (Very strong)
Rating	1	On a scale of 0-5, what is Indigo's rating on Google?	0 (Terrible) to 5 (Excellent)
Upgrade	1	Out of 100 guests at Indigo Hotel (Indigo Hotels & Resorts), how many ask for room upgrades?	0 to 100
Beliefs about guests	5 (a = .784)	Out of 100 guests at Indigo Hotel (Indigo Hotels & Resorts), how many take home the following items from their stays? Towels	0 to 100
		Bathrobes	
		Hair dryers	
		Hangers	
		Toiletries	

Beliefs about other participants	1	What percentage of participants do you think would state that they guessed correctly?	Guess percentage beliefs
Gender	1	What is your gender?	Categorical choice: Female/ Male/ Other/ Prefer not to say
Age	1	How old are you in years?	Open-ended question with only a number within the range of 18 to 150 accepted as content.

Table S2C. List of measures – Study 2c			
Self-reported correct guess	1	Did you guess correctly? (You can receive £X {where X = 1-6 depending on the randomly generated die roll} from Little Hotel Parkside (Grand Hotels & Resorts).	Binary choice: YES/ NO.
Variable	# of items (a)	Item	Scale
Writing Clarity	1	To what extent is the "About Us" text clear?	-5 (Very unclear) to 5 (Very clear)
Capability	1	To what extent does Little Hotel Parkside (Grand Hotels & Resorts) seem capable?	-5 (Very unethical) to 5 (Very ethical)
Likability	1	To what extent does Little Hotel Parkside (Grand Hotels & Resorts) seem likable?	-5 (Very small) to 5 (Very large)
Powerfulness	1	To what extent does Little Hotel Parkside (Grand Hotels & Resorts) seem powerful?	-5 (Very weak) to 5 (Very strong)
Rating	1	On a scale of 0-5, what is Little Hotel Parkside (Grand Hotels & Resorts)' rating on Google?	0 (Terrible) to 5 (Excellent)
Upgrade		Out of 100 guests at Little Hotel Parkside (Grand Hotels & Resorts), how many ask for room upgrades?	0 to 100
Beliefs about guests	6 (a = .800)	Out of 100 guests at Little Hotel Parkside (Grand Hotels &	0 to 100

		Resorts), how many take home the following items from their stays? Towels	
		Bathrobes	
		Hair dryers	
		Hangers	
		Toiletries	
		Hangers	
Beliefs about other participants	1	What percentage of participants do you think would state that they guessed correctly?	0 to 100
Gender	1	What is your gender?	Categorical choice: Female/ Male/ Other/ Prefer not to say
Age	1	How old are you in years?	Open-ended question with only a number within the range of 18 to 150 accepted as content.

Table S3A. List of measures – Study 3a			
Variable	# of items (a)	Item	Scale
Underbilling error reaction	1	What would you do? (There are no wrong answers. We're just interested in people's decisions).	Binary choice: 0 (report the billing error to a Lil (Colossal Mart) employee, 1 (leave without reporting the error)
Sureness	1	How sure are you that you would report the billing error to a Lil (Colossal Mart) employee/ leave Lil (Colossal Mart) without reporting the billing error?	0 (Not at all) to 100 (completely)

DeNiAL Framework (Skowronek, 2022) inspired measures	4 (a = .882)	To what extent do you agree with the following statements? (There are no wrong answers. We are just interested in your opinion). Lil (Colossal Mart) deserves to be notified of underbilling errors.	0 (Not at all) to 100 (completely)
		One is responsible for reporting underbilling errors to Lil (Colossal Mart).	
		Underbilling errors can harm Lil (Colossal Mart).	
		One is obliged to report underbilling errors to Lil (Colossal Mart).	
Anticipated Guilt	1	Report: I'd feel guilty if I didn't report the billing error to a Lil (Colossal Mart) employee. Leave: I wouldn't feel guilty if I were to leave Lil (Colossal Mart) without reporting the billing error.	0 (Not at all) to 100 (completely)
Perceived vulnerability	3 (a = .862)	To what extent does Lil (Colossal Mart) seem Rich	0 (Not at all) to 100 (completely)
		Powerful	
		Vulnerable	

Perceived morality	3 (a = .875)	To what extent does Lil (Colossal Mart) seem Virtuous	
		Prosocial	
		Ethical	
Beliefs about other participants	1	Out of 100 participants, how many would you guess would choose to report the billing error to a Lil (Colossal Mart) employee/ leave Lil (Colossal Mart) without reporting the billing error? 10 participants with the closest guesses will get a £2 bonus.	0 to 100
Gender	1	What is your gender?	Categorical choice: Female/ Male/ Other/ Prefer not to say
Age	1	How old are you in years?	Open-ended question with only a number within the range of 18 to 150 accepted as content.
Self-service checkout familiarity	1	Have you used self-service checkouts before?	Categorical choice: Yes/ No, but I'm familiar with it/ No, and I'm not familiar with it

Table S3V. List of measures – Study 3b			
Variable	# of items (a)	Item	Scale
Underbilling error reaction	1	What would you do? (There are no wrong answers. We're just interested in people's decisions).	Binary choice: 0 (report the billing error to a Lil (Colossal Mart) employee, 1 (leave without reporting the error)

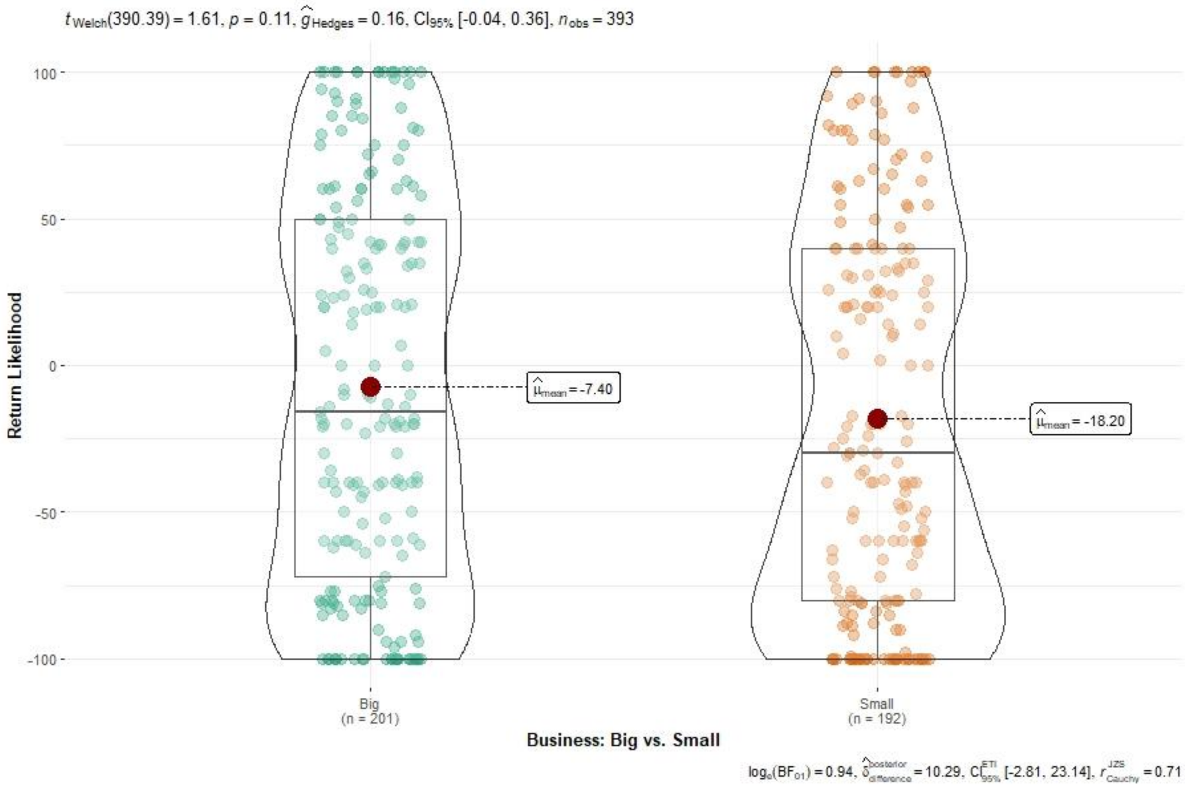
Sureness	1	How sure are you that you would report the billing error to a Lil (Colossal Mart) employee/ leave Lil (Colossal Mart) without reporting the billing error?	0 (Not at all) to 100 (completely)
Perceived vulnerability	3 (a = .712)	To what extent does Lil (Colossal Mart) seem Profitable	0 (Not at all) to 100 (completely)
		Powerful	
		Vulnerable	
Perceived morality	3 (a = .839)	To what extent does Lil (Colossal Mart) seem Virtuous	
		Prosocial	
		Moral	
Beliefs about other participants	1	Imagine that 100 participants are taking this survey and are faced with the same decision. How many of them do you think would choose to leave ${e://Field/Mart} Mart without reporting the billing error?	0 to 100
Gender	1	What is your gender?	Categorical choice: Female/ Male/ Other/ Prefer not to say
Age	1	How old are you in years?	Open-ended question with only a number within the range of 18 to 150 accepted as content.
Self-service checkout familiarity	1	Have you used self-service checkouts before?	Categorical choice: Yes/ No, but I’m familiar with it/ No, and I’m not familiar with it

3. Visualization of main results across studies.

Main package by Patil (2021)

Study 1a

Figure S1A.1. Effect of business size on fraudulent return intentions in Study 1a.



Study 1b

Figure S1B.1. Effect of business size on fraudulent return intentions in Study 1b.

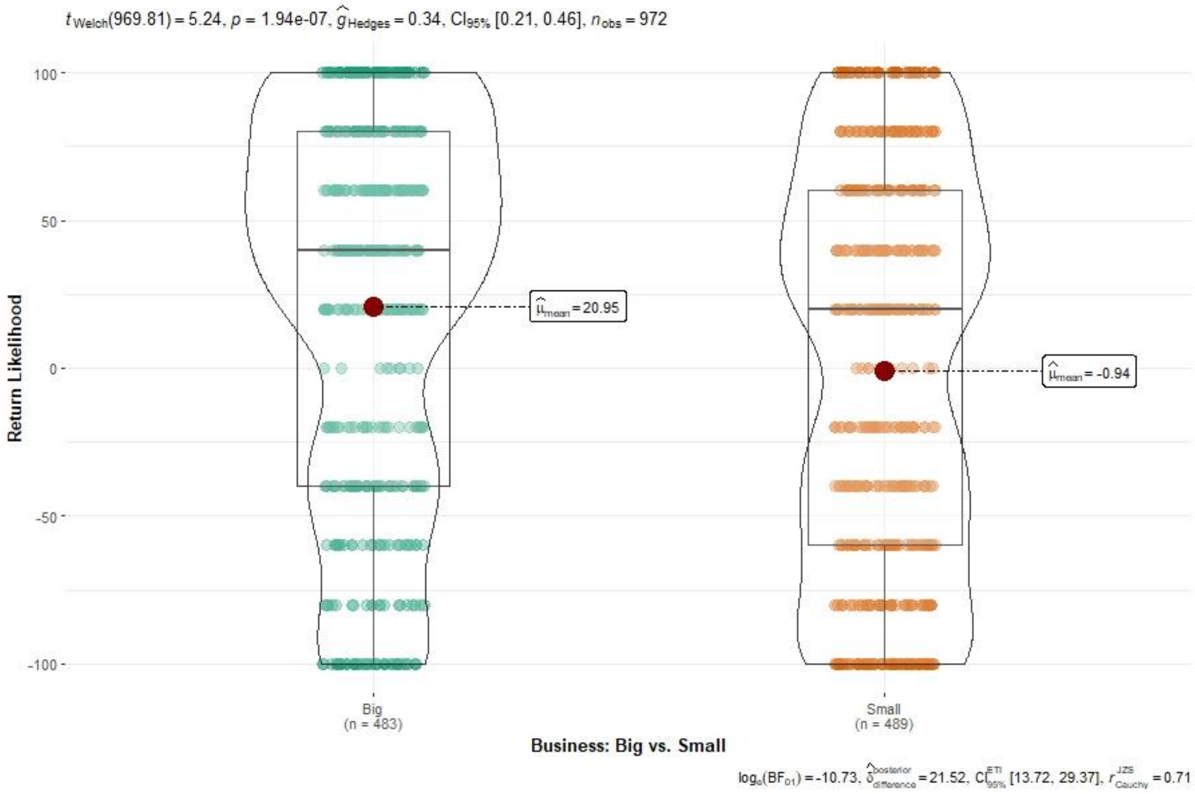
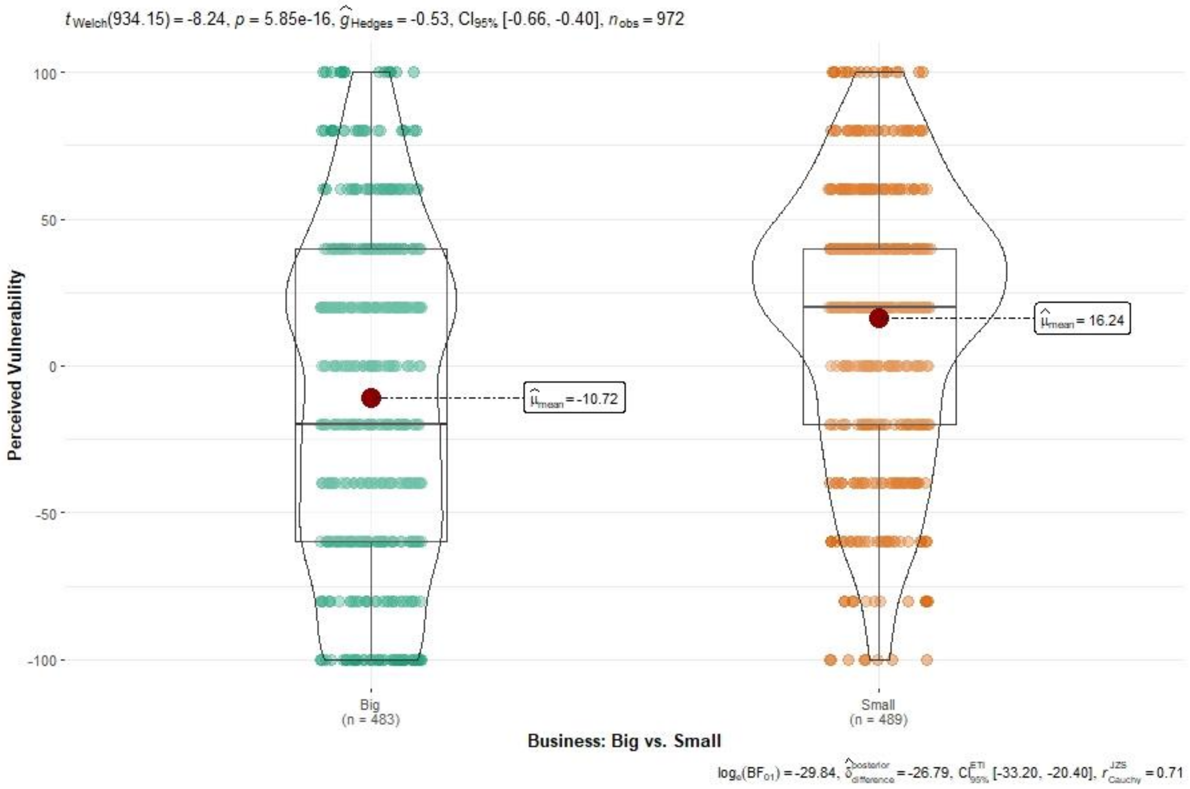
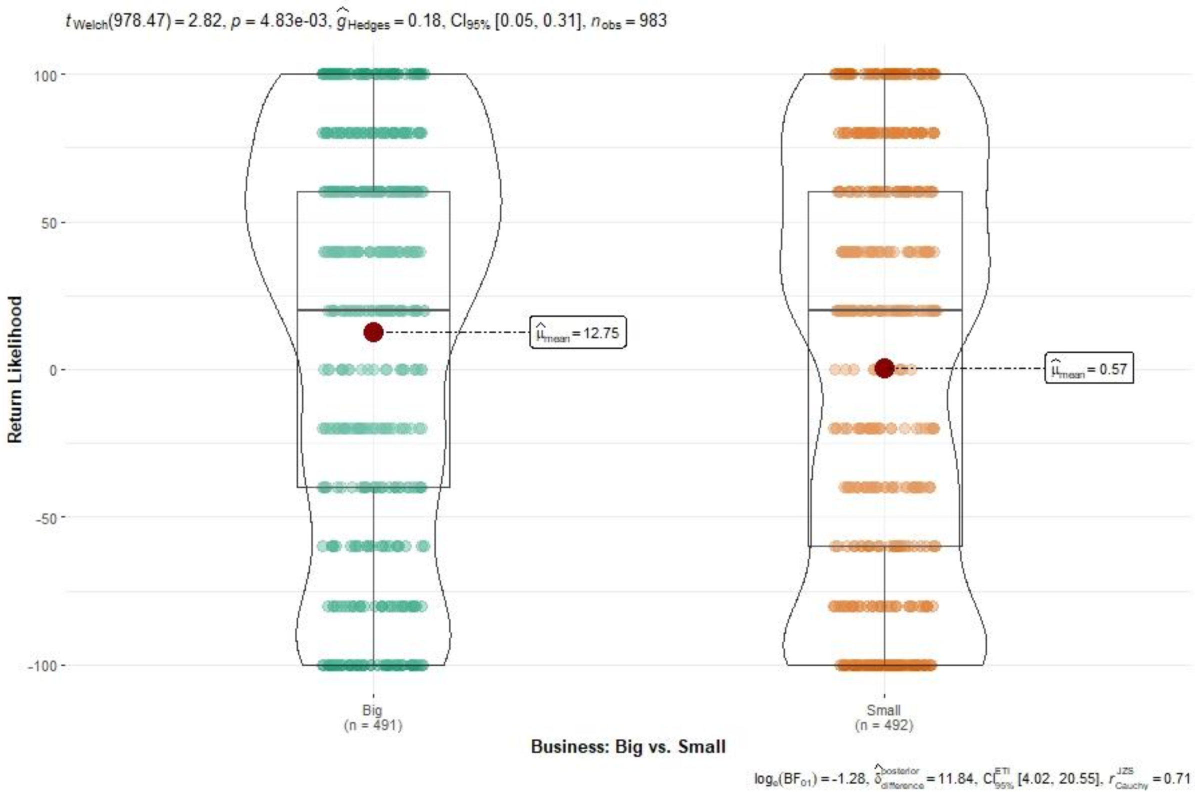


Figure S1B.2. Effect of business size on perceived vulnerability in Study 1b.



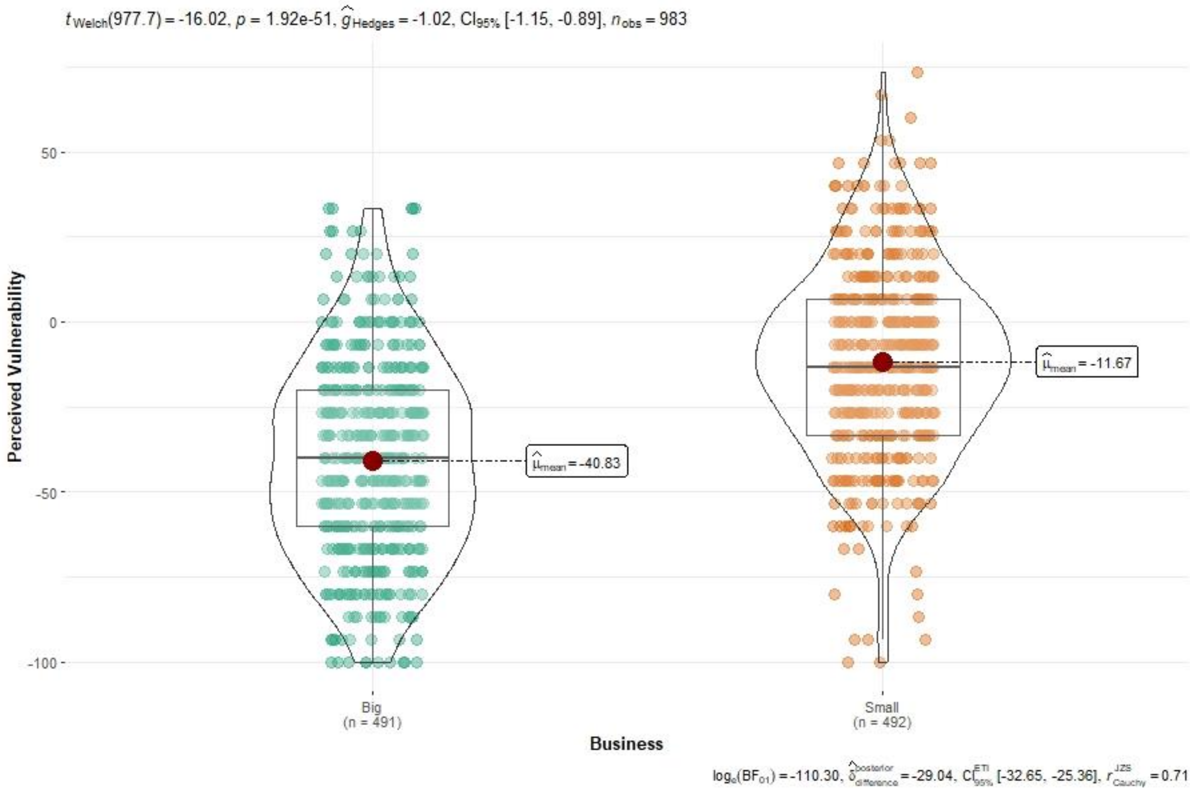
Study 1c

Figure S1C.1. Effect of business size on fraudulent return intentions in Study 1c



.

Figure S1C.2. Effect of business size on perceived vulnerability in Study 1c.



Meta-analysis Studies 1a-c

Dishonest Intentions

Random-Effects Model (k = 3)

	Estimate	se	Z	p	CI Lower Bound	CI Upper Bound
Intercept	0.236	0.0586	4.03	< .001	0.121	0.351
	.	.	.	.	.	.

Note. Tau² Estimator: Restricted Maximum-Likelihood

Heterogeneity Statistics

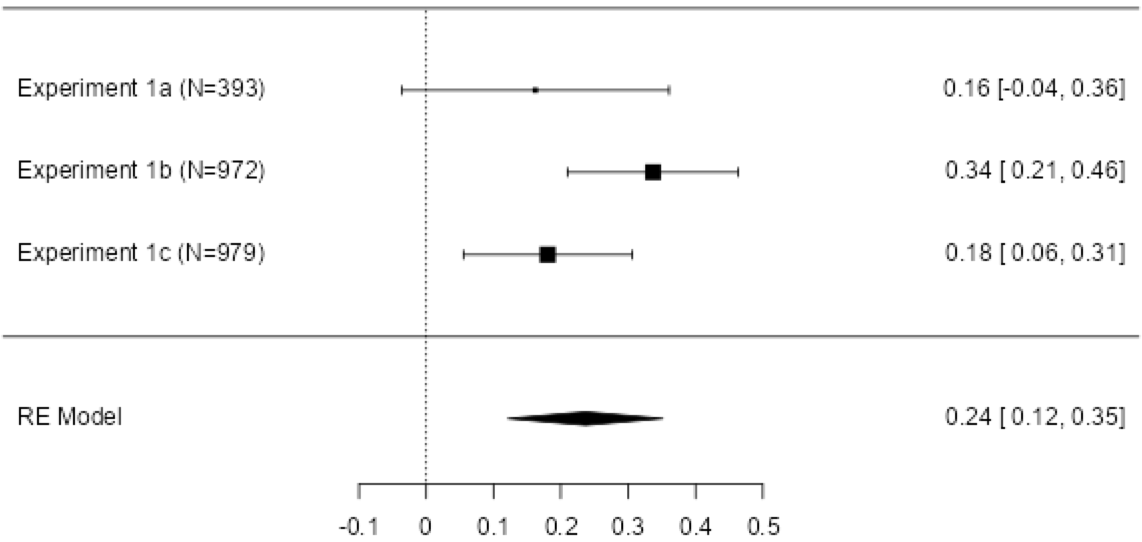
Tau	Tau²	I²	H²	R²	df	Q	p
0.069	0.0048 (SE= 0.0104)	46.53%	1.870	.	2.000	3.688	0.158

The analysis was carried out using the standardized mean difference as the outcome measure. A random-effects model was fitted to the data. The amount of heterogeneity (i.e., tau²), was estimated using the restricted maximum-likelihood estimator (Viechtbauer 2005). In addition to the estimate of tau², the Q-test for heterogeneity (Cochran 1954) and the I² statistic are reported. In case any amount of heterogeneity is detected (i.e., tau² > 0, regardless of the results of the Q-test), a prediction interval for the true outcomes is also provided. Studentized residuals and Cook's distances are used to examine whether studies may be outliers and/or influential in the context of the model. Studies with a studentized residual larger than the 100 x (1 - 0.05/(2 X k))th percentile of a standard normal distribution are considered potential outliers (i.e., using a Bonferroni correction with two-sided alpha = 0.05 for k studies included in the meta-analysis). Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances are considered to be influential. The rank correlation test and the regression test, using the standard error of the observed outcomes as predictor, are used to check for funnel plot asymmetry.

A total of k=3 studies were included in the analysis. The observed standardized mean differences ranged from 0.1623 to 0.3367, with the majority of estimates being positive (100%). The estimated average standardized mean difference based on the random-effects model was \hat{\mu} = 0.2364 (95% CI: 0.1215 to 0.3512). Therefore, the average outcome differed significantly from zero (z = 4.0324, p < 0.0001). According to the Q-test, there was no significant amount of heterogeneity in the true outcomes (Q(2) = 3.6883, p = 0.1582, tau² = 0.0048, I² = 46.5285%). A 95% prediction interval for the true outcomes is given by 0.0587 to 0.4140. Hence, even though there may be some heterogeneity, the true outcomes of the studies are generally in the same direction as the estimated average outcome. An examination of the studentized residuals revealed that none of the studies had a value larger than ± 2.3940 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential.

Neither the rank correlation nor the regression test indicated any funnel plot asymmetry (p = 1.0000 and p = 0.5535, respectively).

Forest Plot

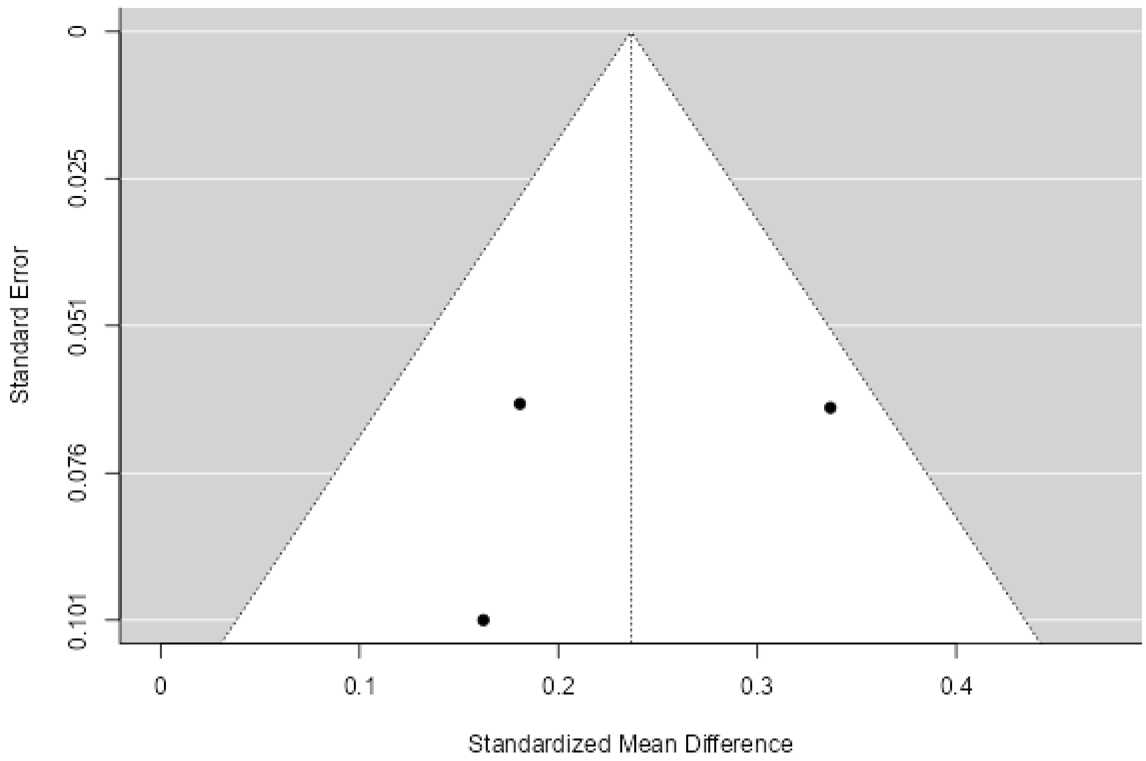


Publication Bias Assessment

Test Name	value	p
Fail-Safe N	32.000	< .001
Begg and Mazumdar Rank Correlation	0.333	1.000
Egger's Regression	-0.593	0.553
Trim and Fill Number of Studies	0.000	.

Note. Fail-safe N Calculation Using the Rosenthal Approach

Funnel Plot

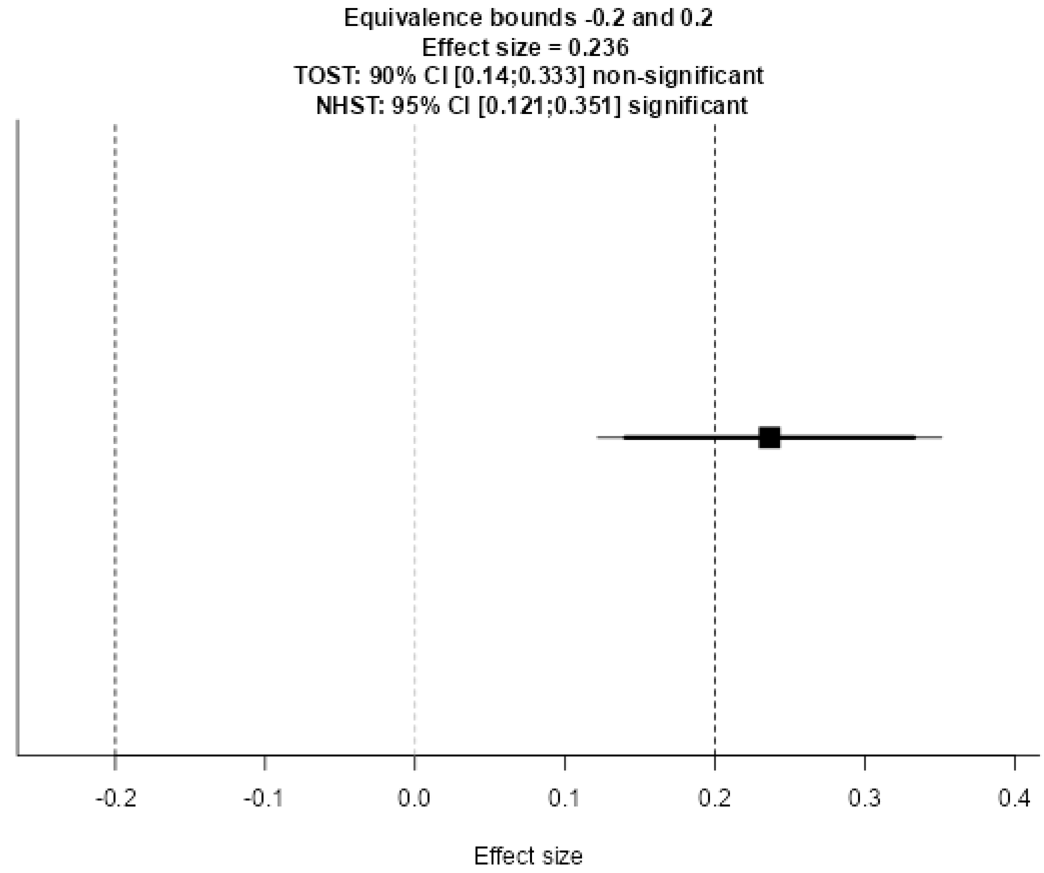


Two One-Sided Tests Equivalence Testing

Z-Value Lower Bound	P-Value Lower Bound	Z-Value Upper Bound	P-Value Upper Bound	Lower Limit Confidence Interval TOST	Upper Limit Confidence Interval TOST	LL_CI_ZTEST	Z-Value Upper Limit Confidence Interval TOST
7.444	< .001	0.620	0.732	0.140	0.333	0.121	0.351

The equivalence test was non-significant, Z = 0.620, p = 0.732, given equivalence bounds of -0.200 and 0.200 and an alpha of 0.05. The null hypothesis test was significant, Z = 4.032, p = 0.0000552, given an alpha of 0.05. Based on the equivalence test and the null hypothesis test combined, we can conclude that the observed effect is statistically different from zero and statistically not equivalent to zero.

Equivalence Test Plot



Equivalence bounds -0.2 and 0.2
Effect size = 0.236
TOST: 90% CI [0.14;0.333] non-significant
NHST: 95% CI [0.121;0.351] significant
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
Effect size

Study 2a

Figure S2A.1. Effect of business size on the proportion of reported correct guesses of the score of the first die roll in Study 2a.

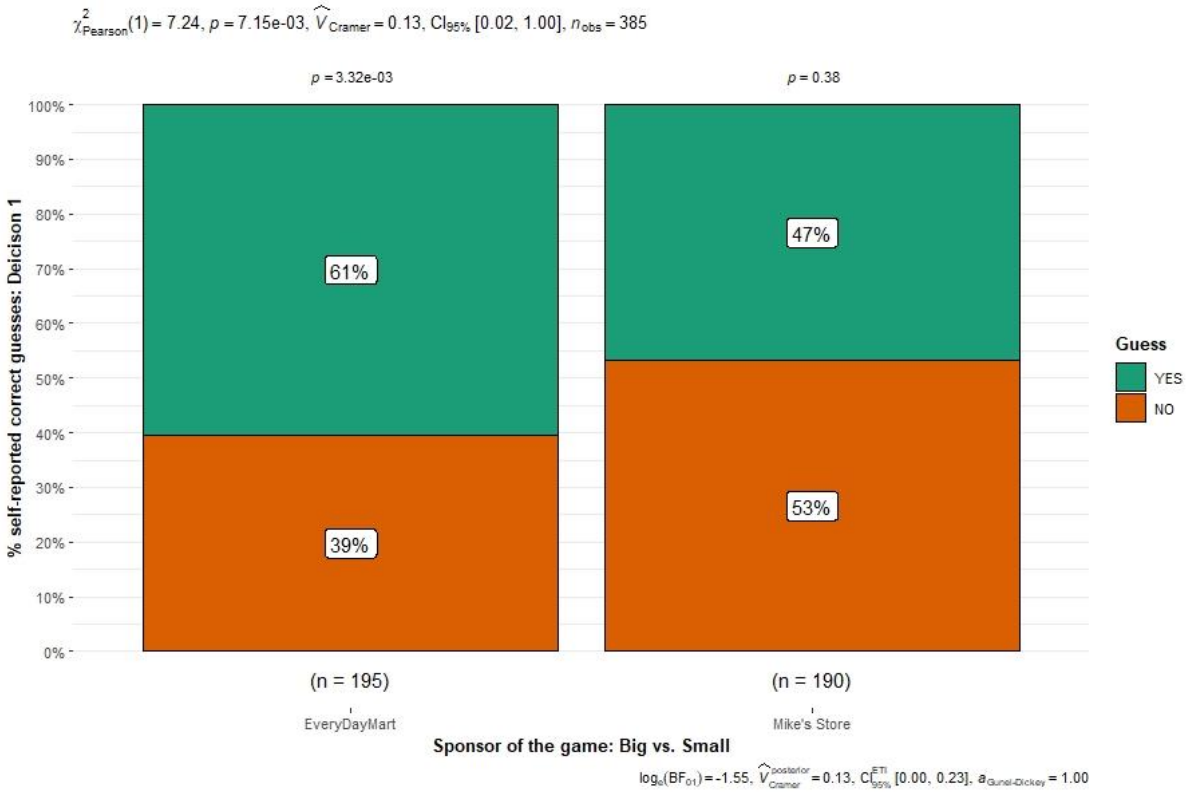
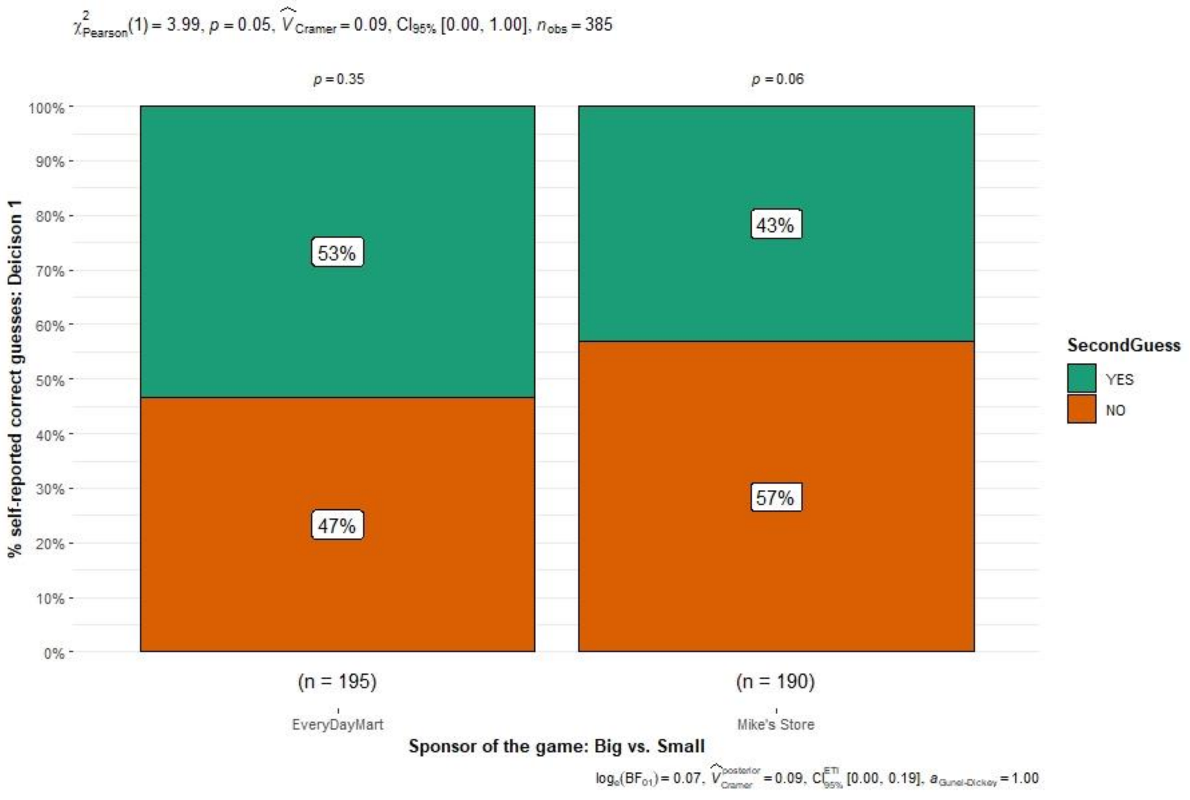


Figure S2A.2. Effect of business size on the proportion of reported correct guesses of the score of the second die roll in Study 2a.



Study 2b

Figure S2B.1. Effect of business size on the proportion of reported correct guesses of the score of the die roll in Study 2b.

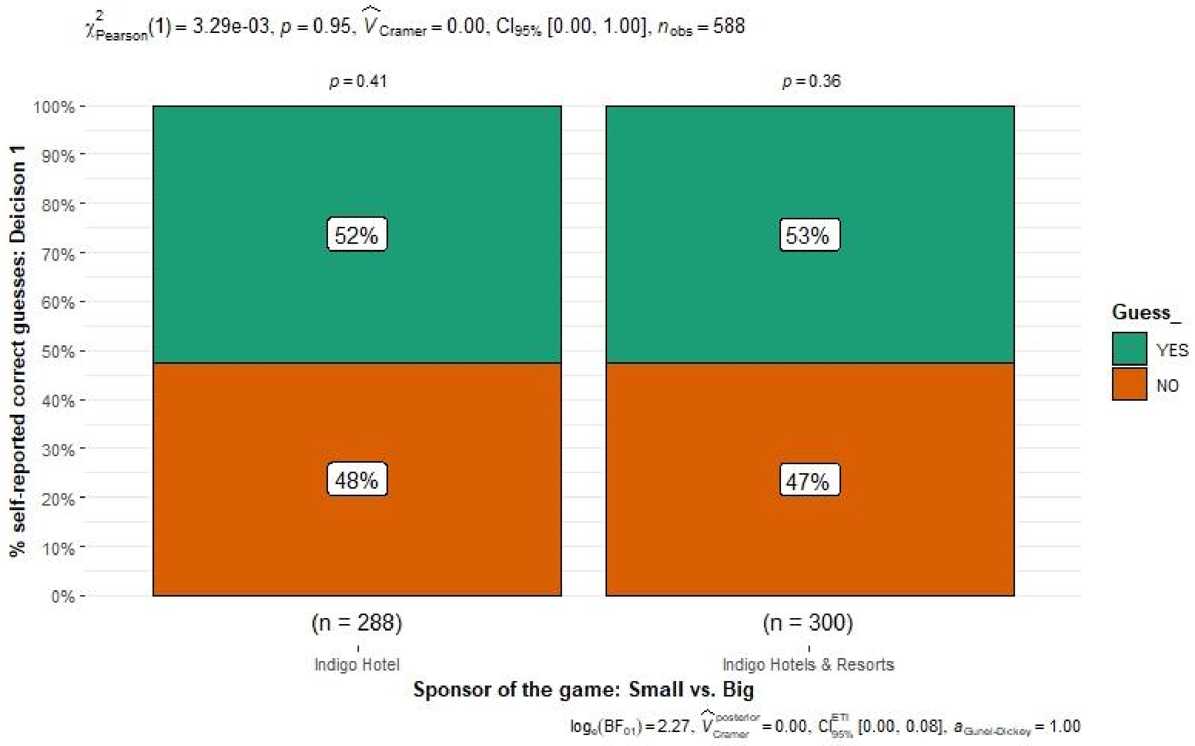


Figure S2B.1. Differences in beliefs about others' dishonesty toward big vs. small businesses in Study 2b.

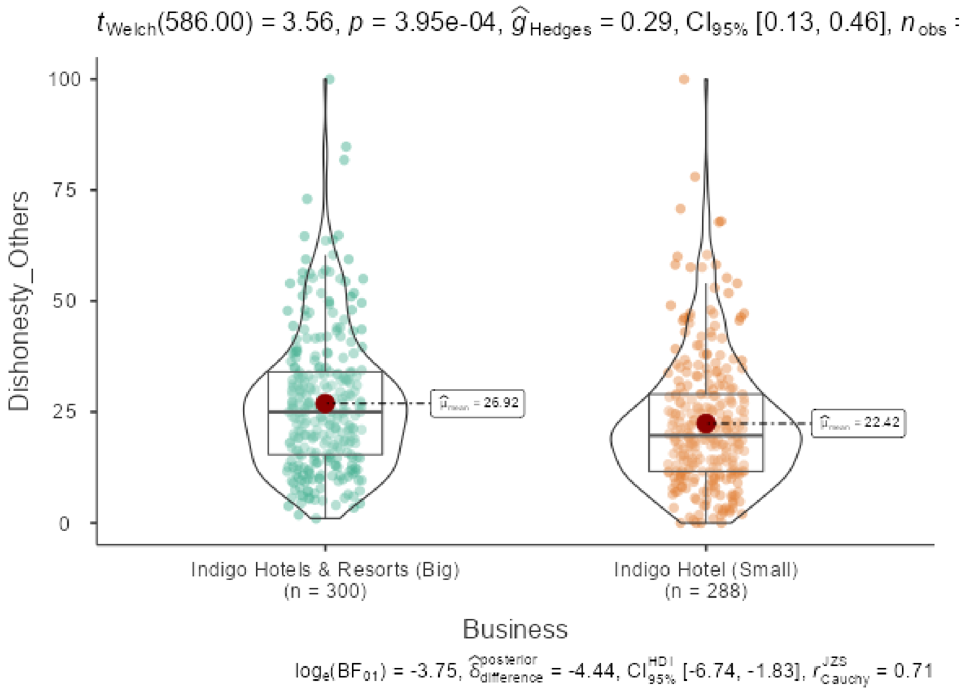
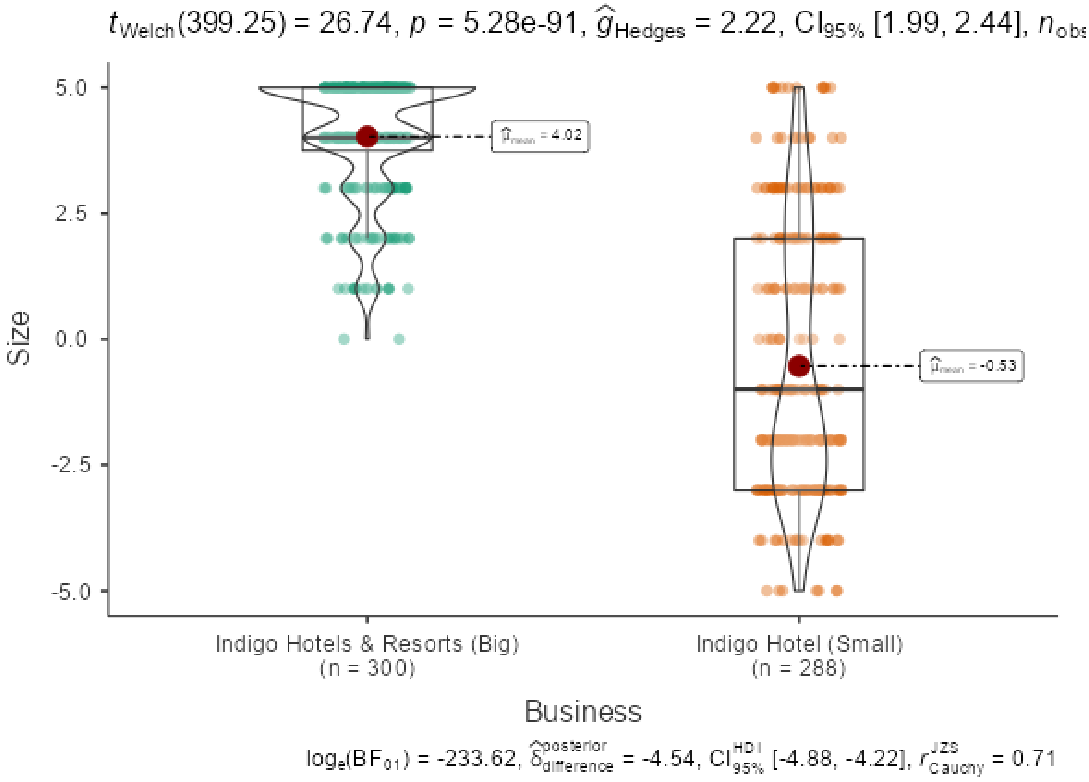


Figure S2B.1. Differences in business Size Perceptions in Study 2b.



Study 2c

Figure S2C.1. Effect of business size on the proportion of reported correct guesses of the score of the die roll in Study 2c.

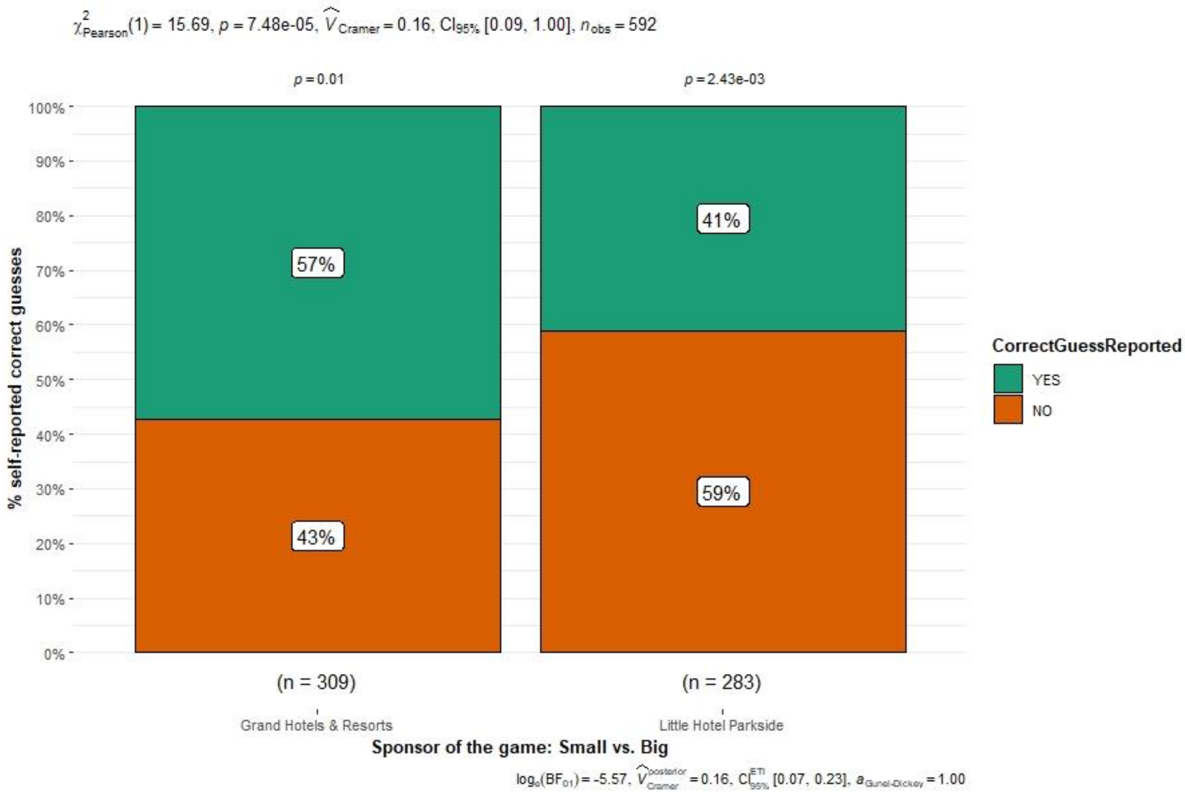
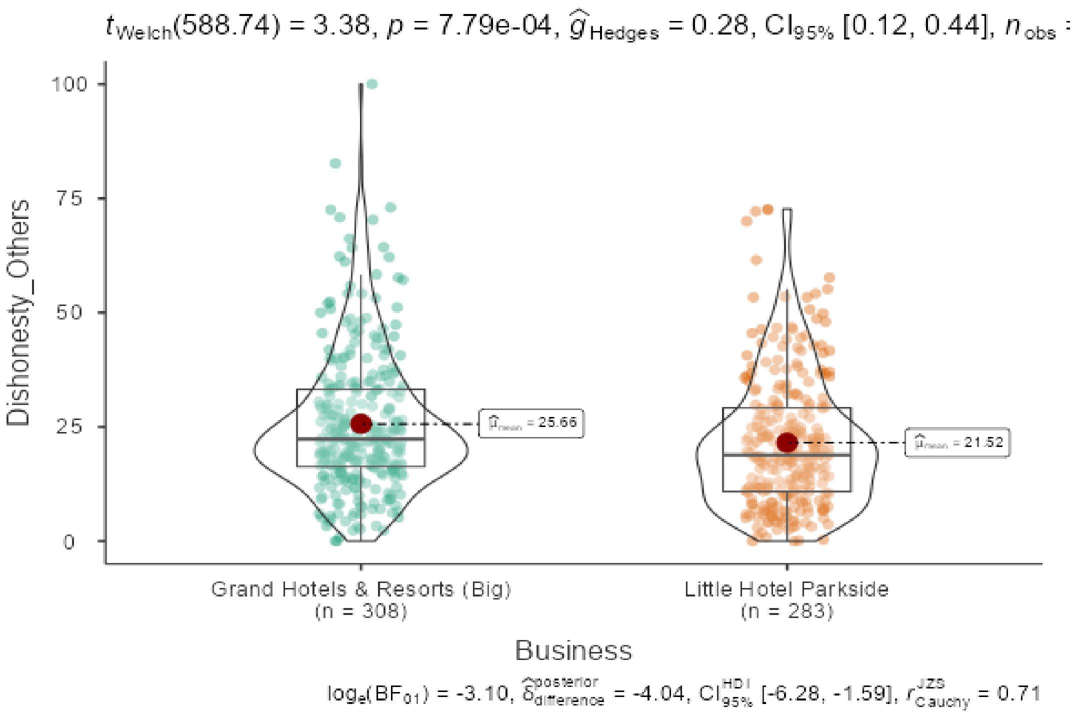


Figure S2C.1. Differences in beliefs about others' dishonesty toward big vs. small businesses in Study 2c.



Meta-analysis Studies 2a-c

Behavioral Dishonesty

Random-Effects Model (k = 4)

	Estimate	se	Z	p	CI Lower Bound	CI Upper Bound
Intercept	0.448	0.170	2.63	0.008	0.115	0.781
	.	.	.	.	.	.

Note. Tau² Estimator: Restricted Maximum-Likelihood

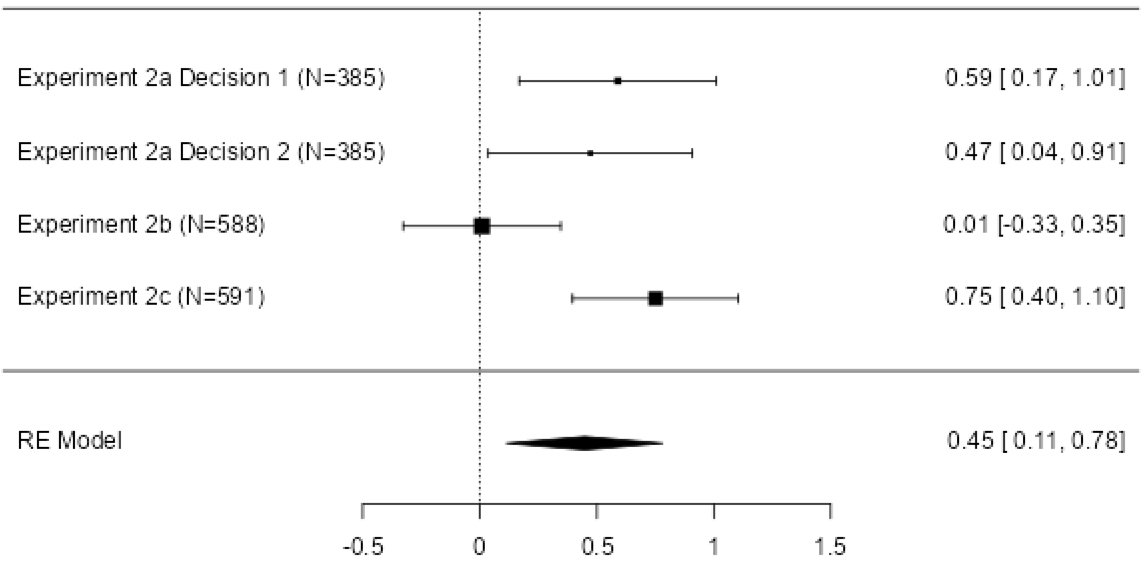
Heterogeneity Statistics

Tau	Tau²	I²	H²	R²	df	Q	p
0.277	0.0769 (SE= 0.0944)	66.85%	3.017	.	3.000	9.697	0.021

The analysis was carried out using the log odds ratio as the outcome measure. A random-effects model was fitted to the data. The amount of heterogeneity (i.e., tau²), was estimated using the restricted maximum-likelihood estimator (Viechtbauer 2005). In addition to the estimate of tau², the Q-test for heterogeneity (Cochran 1954) and the I² statistic are reported. In case any amount of heterogeneity is detected (i.e., tau² > 0, regardless of the results of the Q-test), a prediction interval for the true outcomes is also provided. Studentized residuals and Cook's distances are used to examine whether studies may be outliers and/or influential in the context of the model. Studies with a studentized residual larger than the 100 x (1 - 0.05/(2 X k))th percentile of a standard normal distribution are considered potential outliers (i.e., using a Bonferroni correction with two-sided alpha = 0.05 for k studies included in the meta-analysis). Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances are considered to be influential. The rank correlation test and the regression test, using the standard error of the observed outcomes as predictor, are used to check for funnel plot asymmetry.

A total of k=4 studies were included in the analysis. The observed log odds ratios ranged from 0.0103 to 0.7507, with the majority of estimates being positive (100%). The estimated average log odds ratio based on the random-effects model was \hat{\mu} = 0.4479 (95% CI: 0.1146 to 0.7811). Therefore, the average outcome differed significantly from zero (z = 2.6337, p = 0.0084). According to the Q-test, the true outcomes appear to be heterogeneous (Q(3) = 9.6972, p = 0.0213, tau² = 0.0769, I² = 66.8497%). A 95% prediction interval for the true outcomes is given by -0.1898 to 1.0855. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative. An examination of the studentized residuals revealed that one study (Study 2b (N=588)) had a value larger than ± 2.4977 and may be a potential outlier in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry (p = 1.0000 and p = 0.5575, respectively).

Forest Plot



Publication Bias Assessment

Fail-Safe N Analysis (File Drawer Analysis)

Fail-safe N	p
27.000	< .001

Note. Fail-safe N Calculation Using the Rosenthal Approach

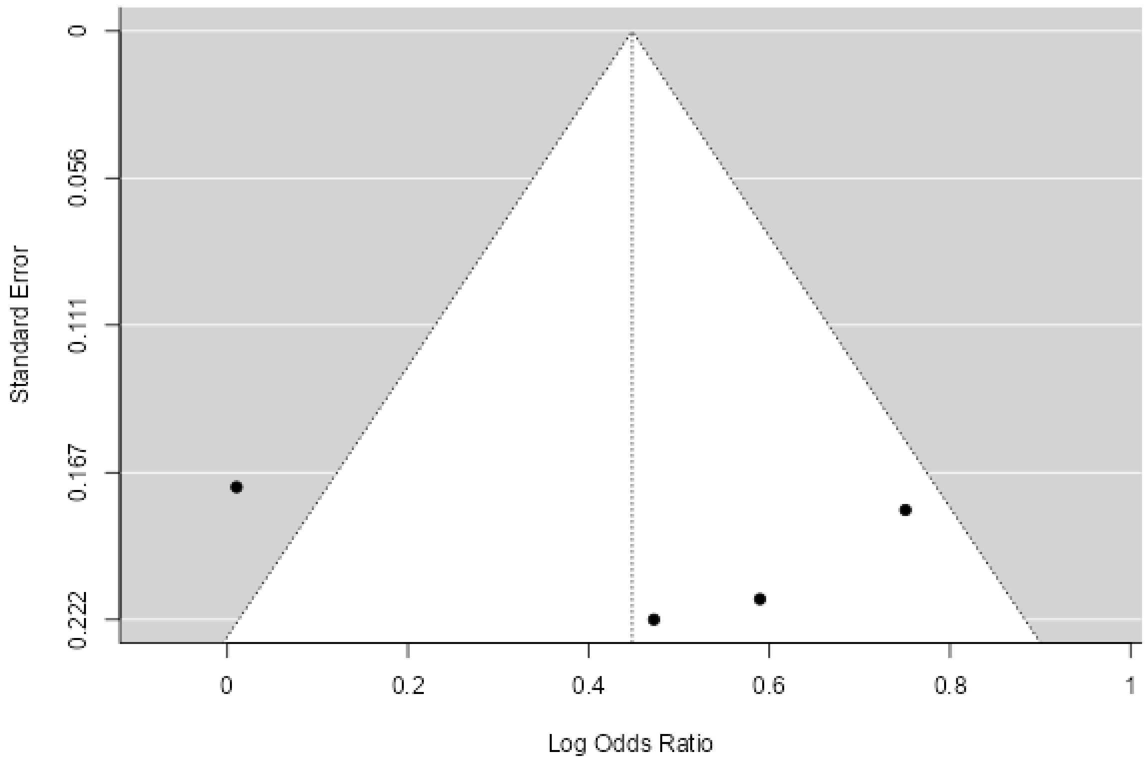
Rank Correlation Test for Funnel Plot Asymmetry

Kendall's Tau	p
0.000	1.000

Regression Test for Funnel Plot Asymmetry

Z	p
0.586	0.558

Funnel Plot



Study 3a

Figure S3A.1 Effect of business size on the proportion of people choosing to report an underbilling error or leave without reporting.

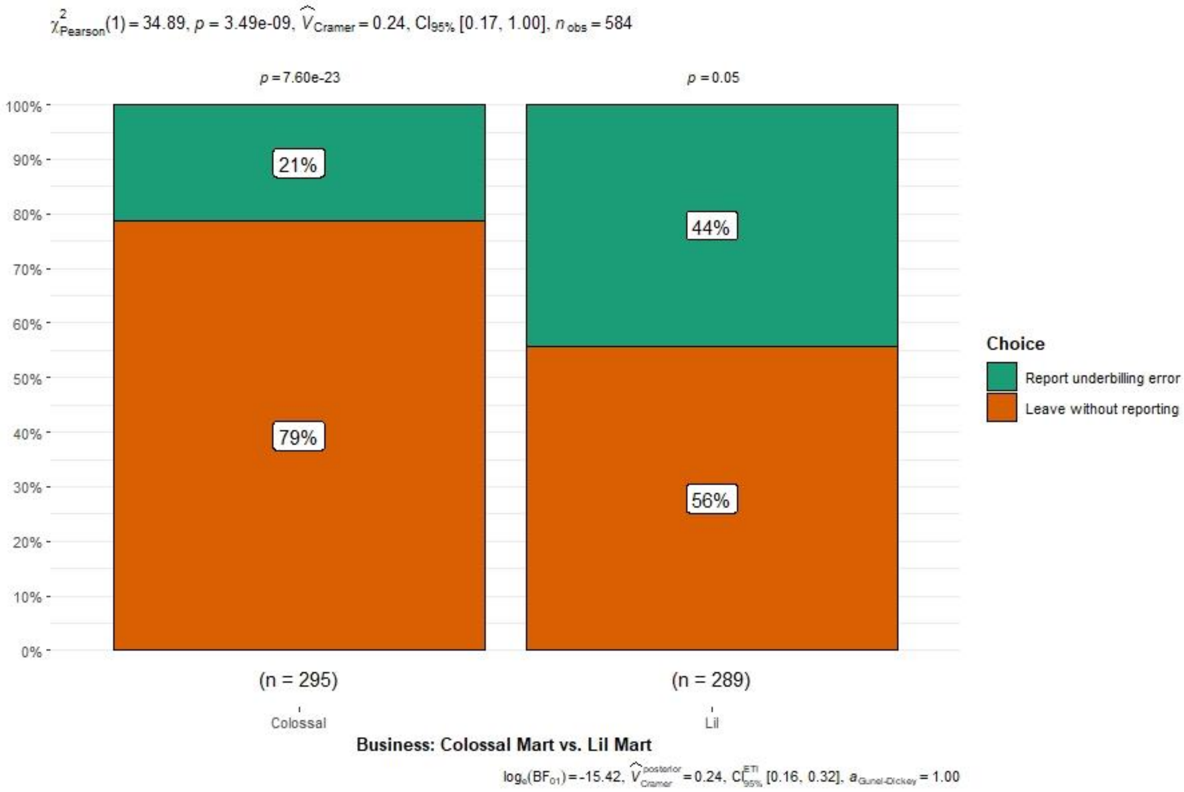


Figure S3B.2 Effect of business size on Decision (1 = Honest, -1 = Dishonest) X Sureness.

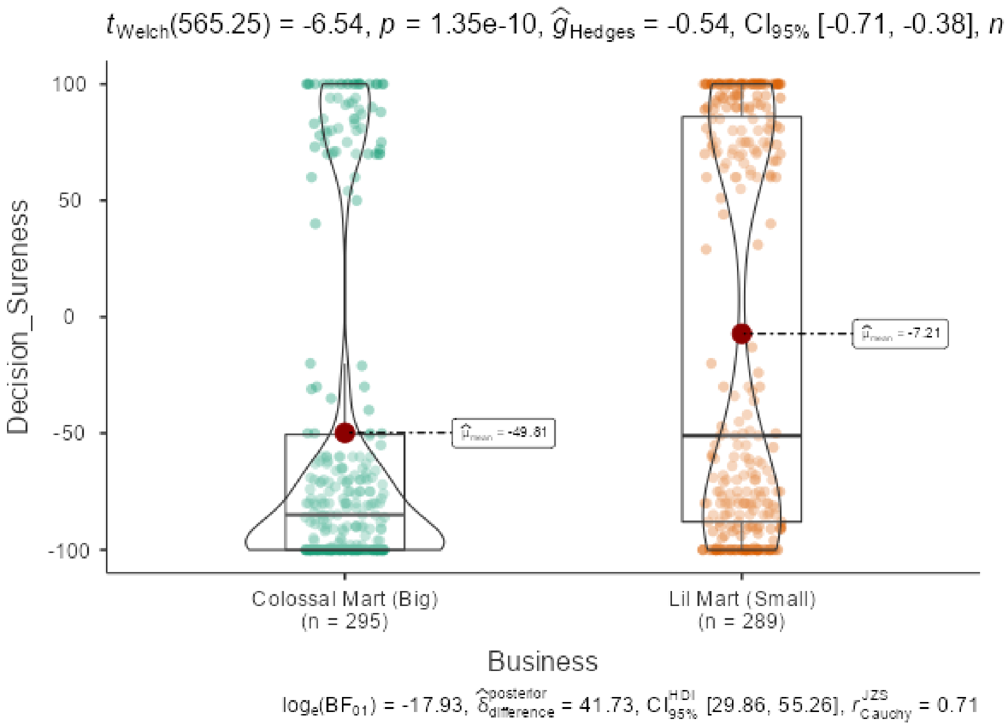


Figure S3A.3. Effect of business size on perceived vulnerability in Study 3a.

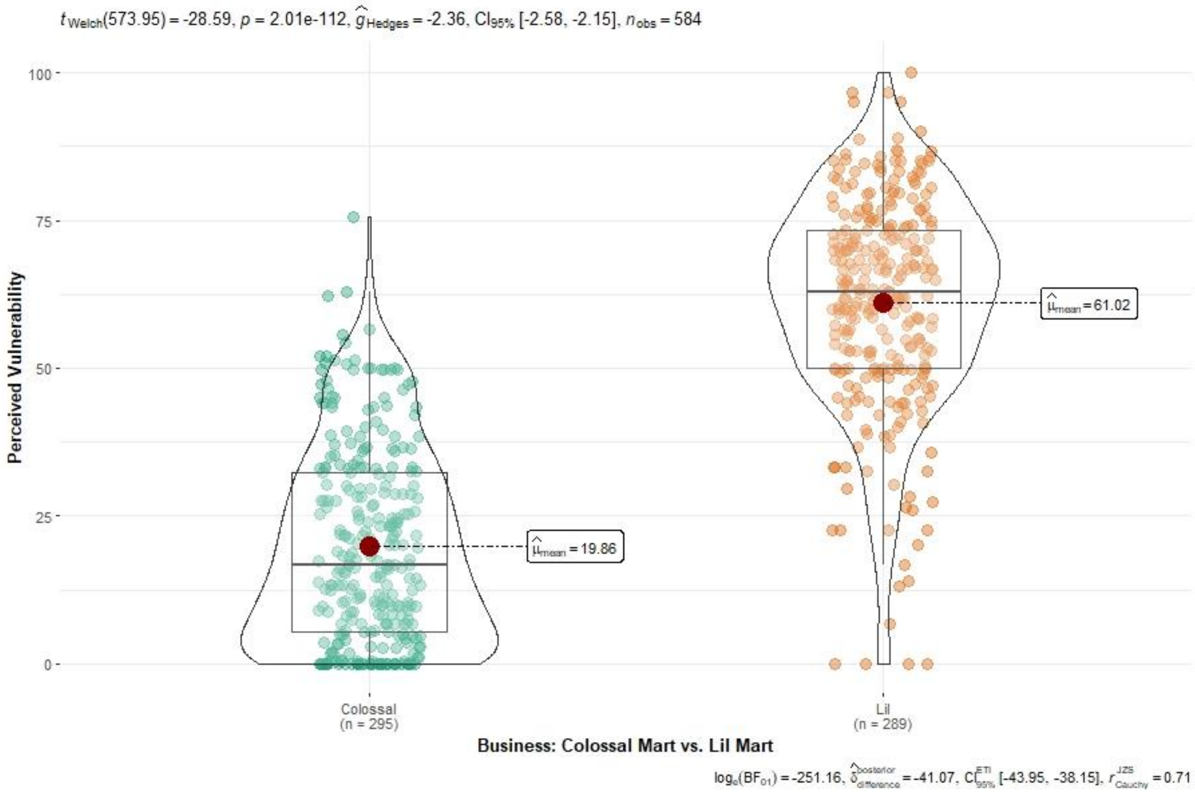
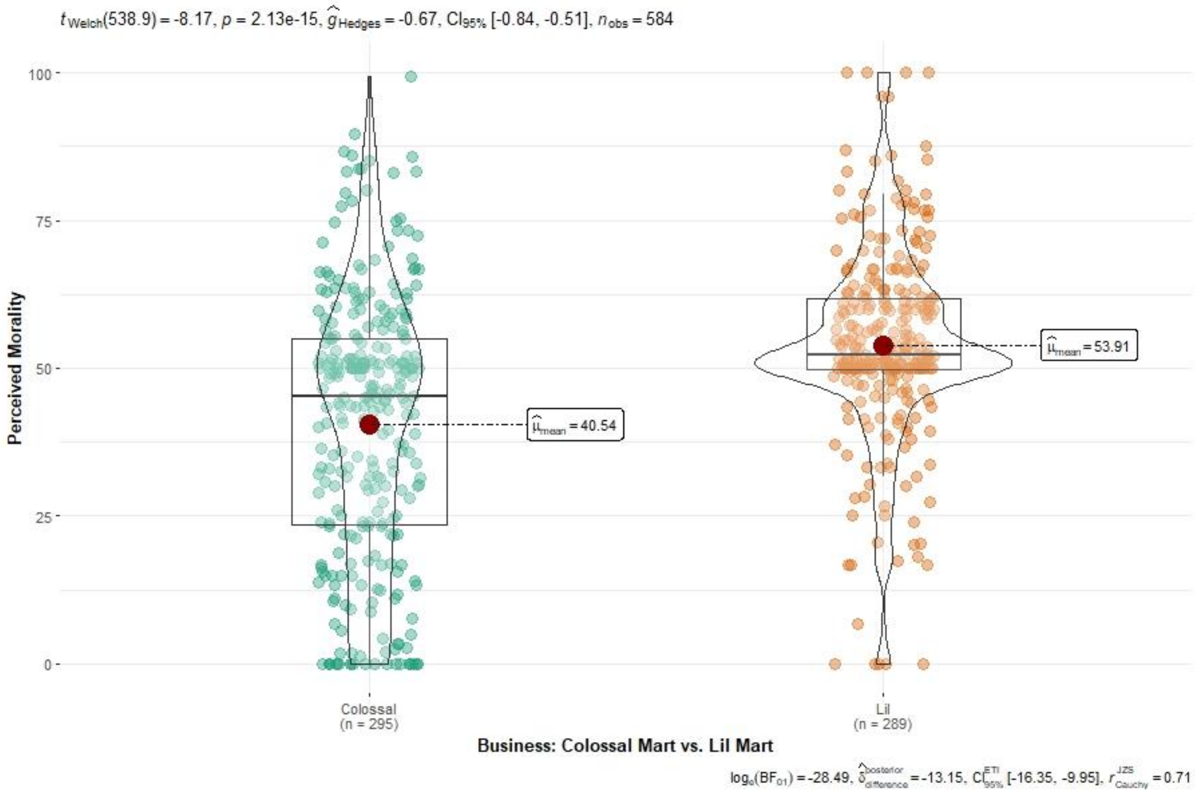


Figure S3A.4. Effect of business size on perceived morality in Study 3a.



Study 3b

Figure S3B.1 Effect of business size on the proportion of people choosing to report an underbilling error or leave without reporting.

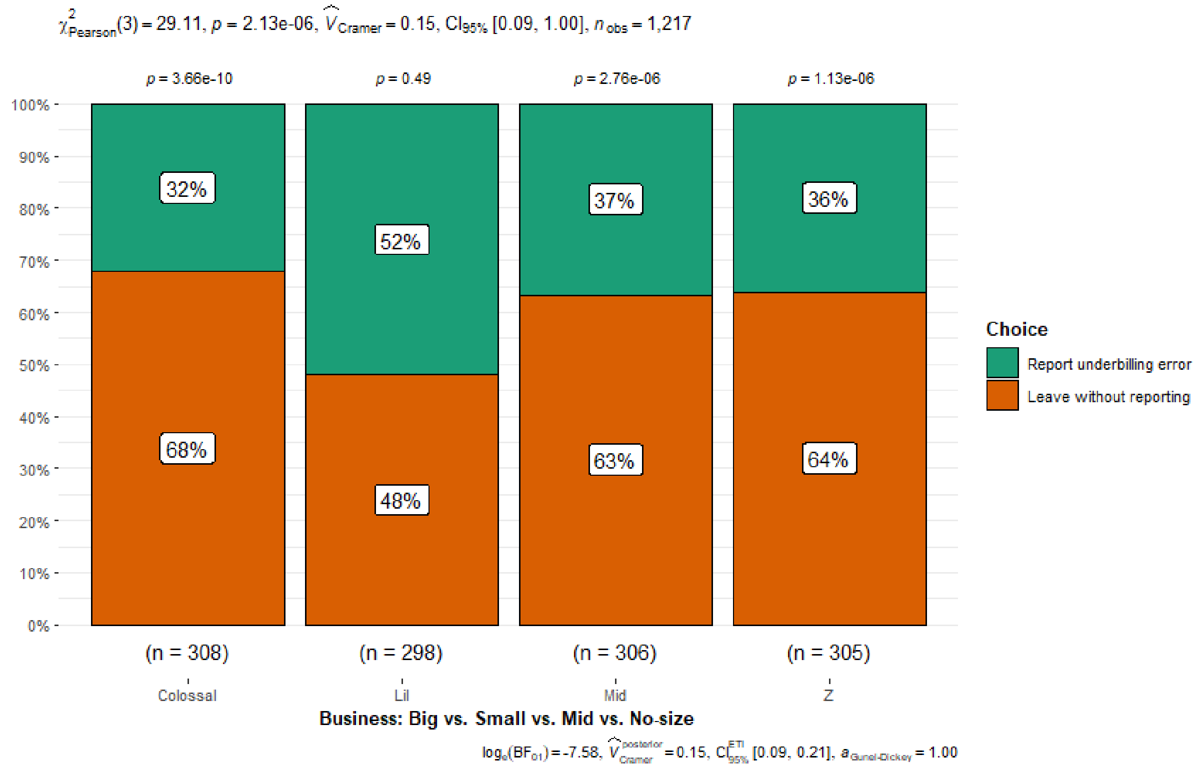


Figure S3B.2 Effect of business size on Decision (1 = Honest, -1 = Dishonest) X Sureness.

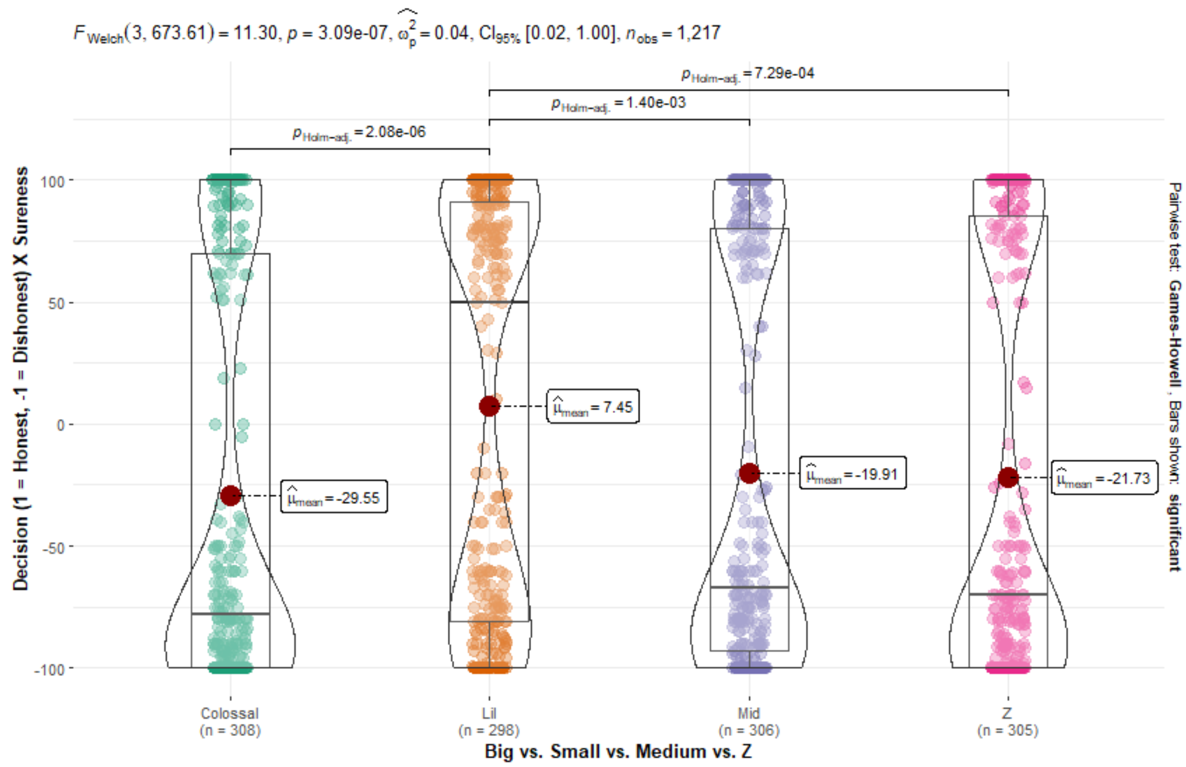


Figure S3B.3. Effect of business size on perceived vulnerability in Study 3b.

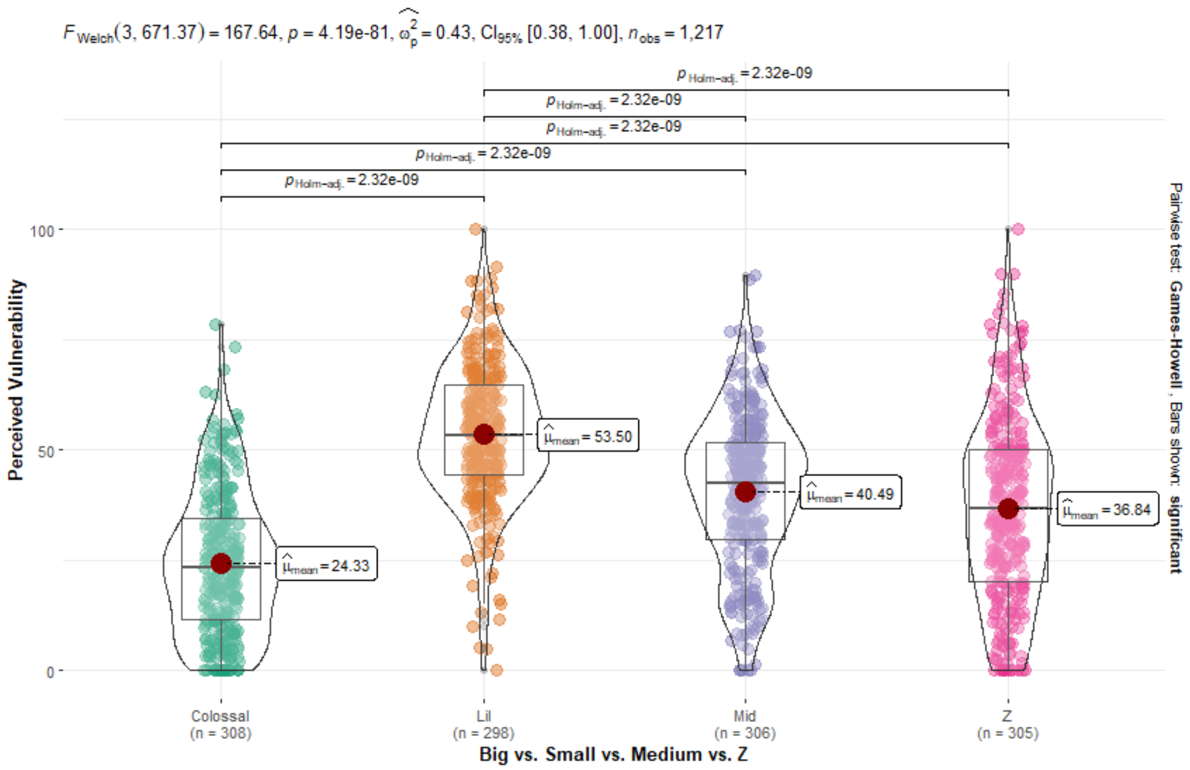


Figure S3B.4. Effect of business size on perceived morality in Study 3b.

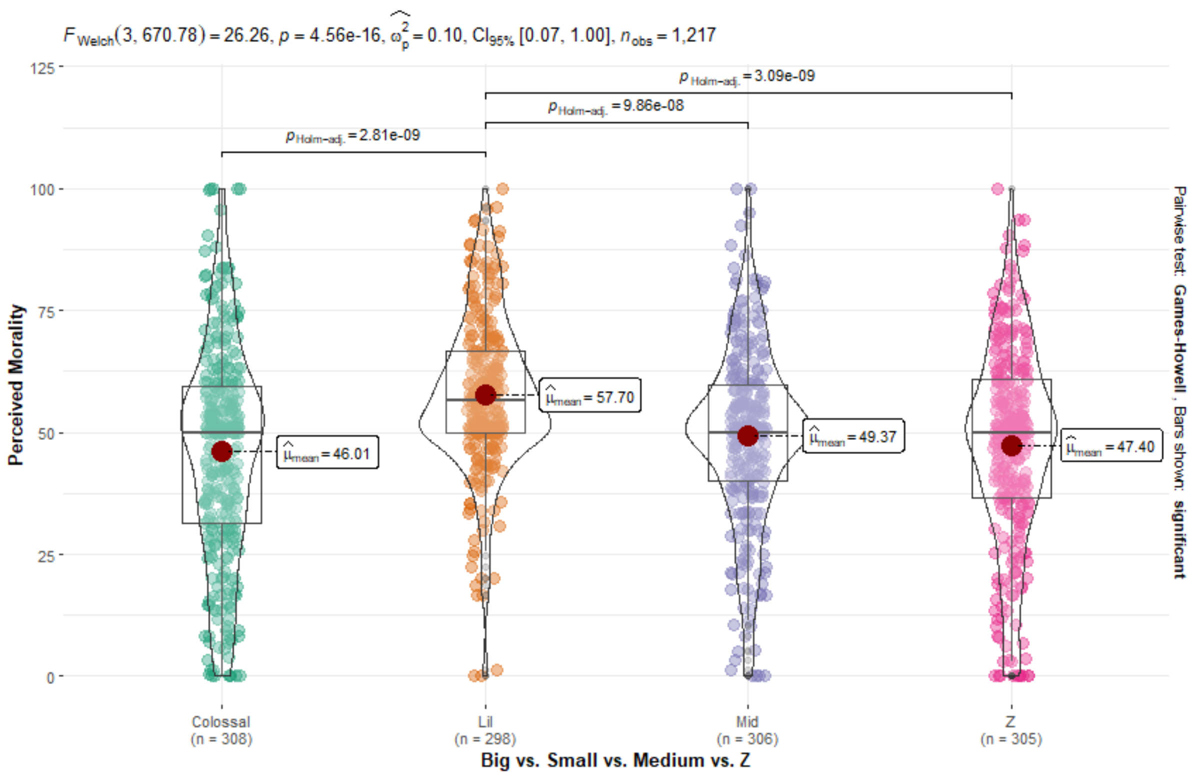
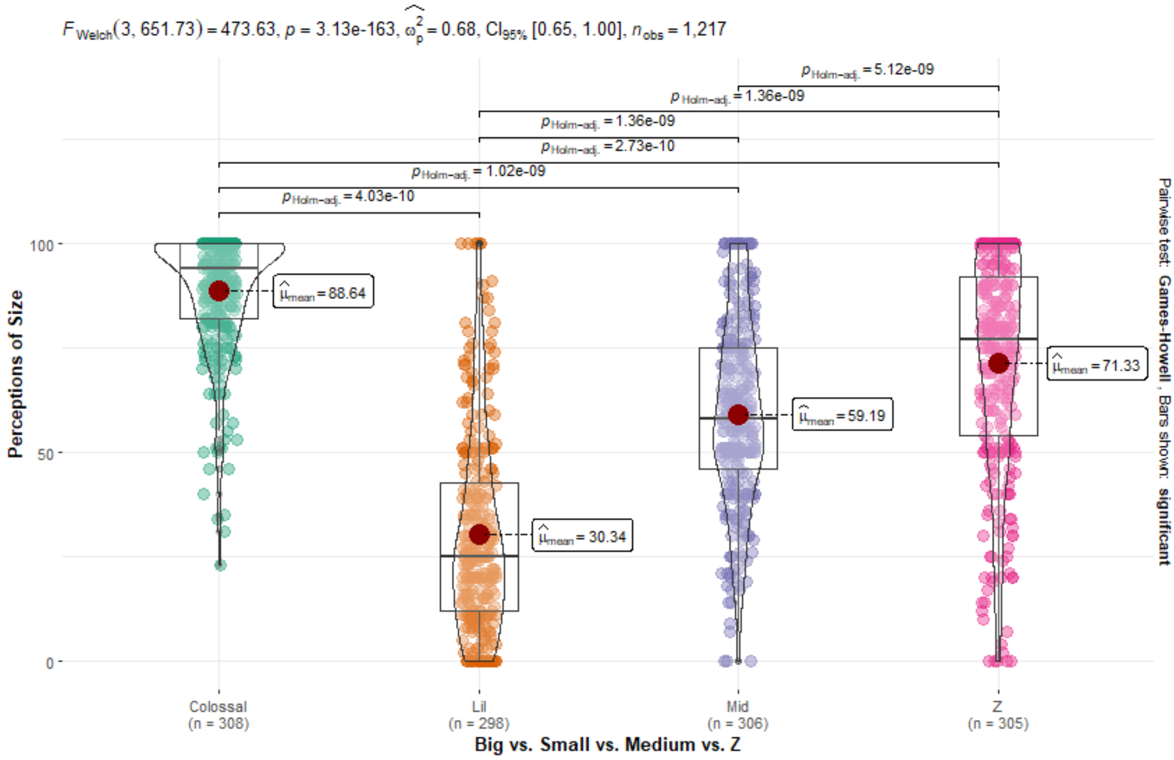


Figure S3B.4. Effect of business size on perceived size in Study 3b.



Output of Meta-analysis of all eight experiments

Meta-Analysis

Random-Effects Model (k = 9)

	Estimate	se	Z	p	CI Lower Bound	CI Upper Bound
Intercept	0.306	0.0604	5.06	< .001	0.187	0.424
	.	.	.	.	.	.

Note. Tau² Estimator: Restricted Maximum-Likelihood

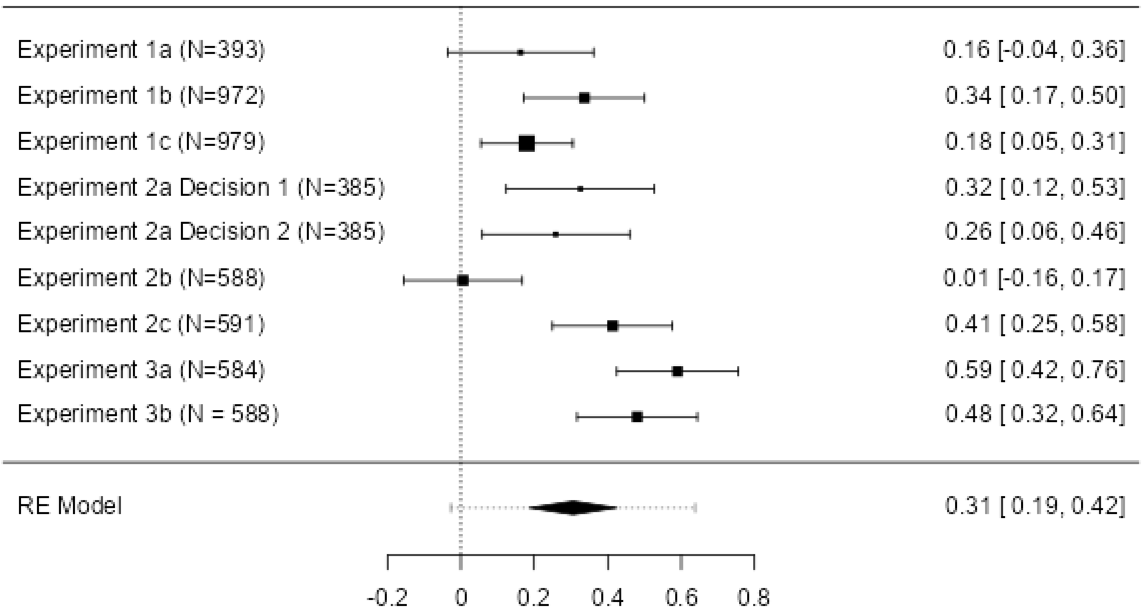
Heterogeneity Statistics

Tau	Tau²	I²	H²	R²	df	Q	p
0.159	0.0252 (SE= 0.0164)	77.54%	4.452	.	8.000	36.721	< .001

Model Fit Statistics and Information Criteria

	log-likelihood	Deviance	AIC	BIC	AICc
Maximum-Likelihood	3.179	21.126	-2.358	-1.963	-0.358
Restricted Maximum-Likelihood	2.389	-4.779	-0.779	-0.620	1.621

Forest Plot

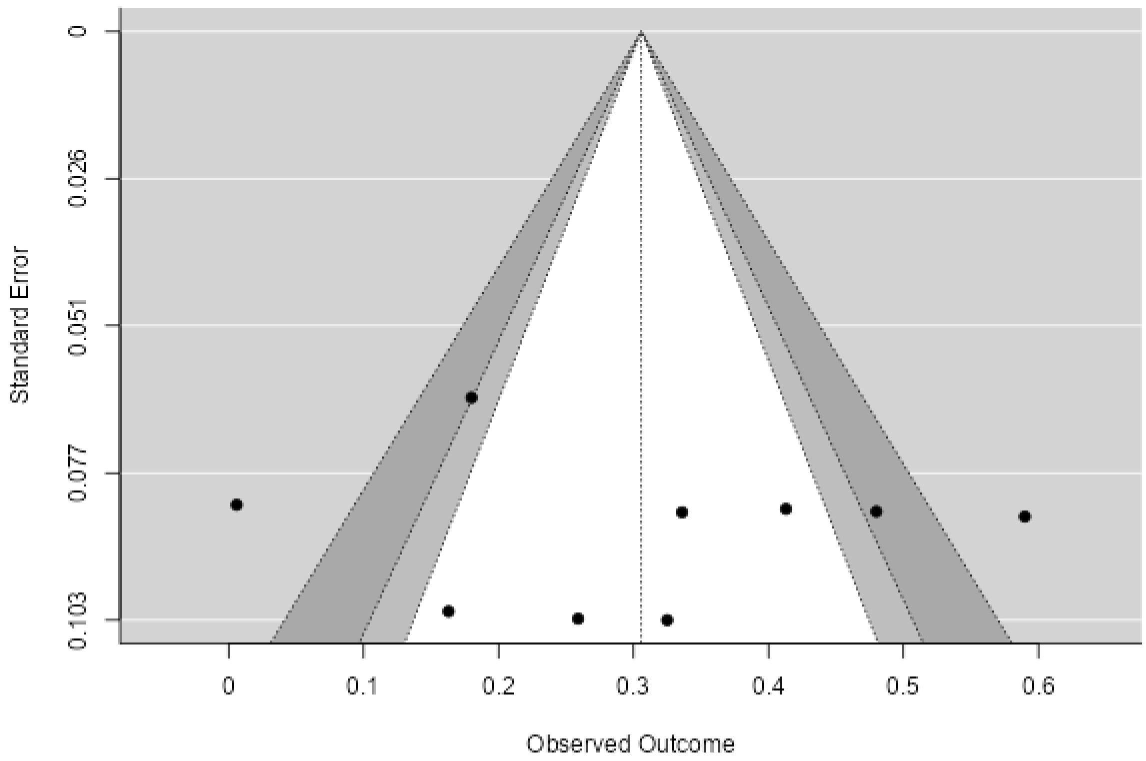


Publication Bias Assessment

Test Name	value	p
Fail-Safe N	367.000	< .001
Kendalls Tau	0.167	0.612
Egger's Regression	0.076	0.939

Note. Fail-safe N Calculation Using the Rosenthal Approach

Funnel Plot



Two One-Sided Tests Equivalence Testing

Z-Value Lower Bound	P-Value Lower Bound	Z-Value Upper Bound	P-Value Upper Bound	LL_CI_TOST	UL_CI_TOST	LL_CI_ZTEST	UL_CI_ZTEST
8.375	< .001	1.752	0.960	0.206	0.405	0.187	0.424

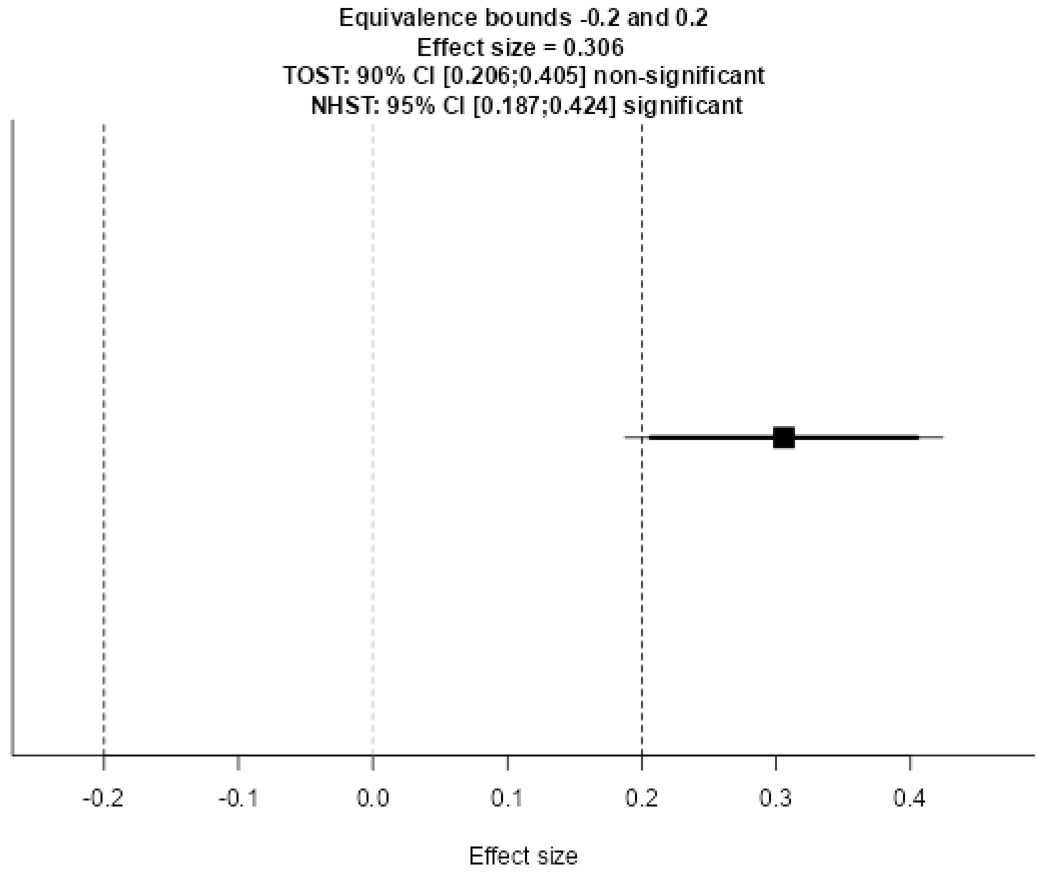
Two One-Sided Tests Equivalence Testing: Text Summary

The equivalence test was non-significant, Z = 1.752, p = 0.960, given equivalence bounds of -0.200 and 0.200 and an alpha of 0.05.

The null hypothesis test was significant, Z = 5.064, p = 0.000000411, given an alpha of 0.05.

Based on the equivalence test and the null hypothesis test combined, we can conclude that the observed effect is statistically different from zero and statistically not equivalent to zero.

Equivalence Test Plot


Equivalence bounds -0.2 and 0.2
Effect size = 0.306
TOST: 90% CI [0.206;0.405] non-significant
NHST: 95% CI [0.187;0.424] significant
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
Effect size

4. Experimental manipulations of business size: big vs. small.

Figure S1A.2. Business-size manipulation in Study 1a: BIG.

Study 1a

In this study, we are going to ask you a some questions about a hypothetical scenario.

Imagine you are attending an out-of-town wedding. Having already arrived at the destination, you realize you forgot to bring your dress shoes. After some searches, you find Nimbus- a big multinational fashion company that sells all kinds of apparel. Nimbus has 1000+ stores and the company owns many popular brands.

You decide to buy the dress shoes from Nimbus and wear the shoes all weekend. The shoes fit OK, and you did not experience much discomfort. However, you have similar dress shoes back at home. You find out that Nimbus accepts returns from dissatisfied customers within 30-days of purchase. And the return process is quick and hassle-free.

Click to edit validation

★ What is a real company that comes to mind which may be similar to Nimbus?

Figure S1A.3. Business-size manipulation in Study 1a: SMALL.

In this study, we are going to ask you some questions about a hypothetical scenario.

Imagine you are attending an out-of-town wedding. Having already arrived at the destination, you realize you forgot to bring your dress shoes. After some searches, you find Nimbus- a local store that sells clothes, shoes, and accessories. Nimbus is family-run from a small shop downtown and they sell unique fashion items.

You decide to buy the dress shoes from Nimbus and wear the shoes all weekend. The shoes fit OK, and you did not experience much discomfort. However, you have similar dress shoes back at home. You find out that Nimbus accepts returns from dissatisfied customers within 30-days of purchase. And the return process is quick and hassle-free.

* What is a real business that comes to mind which may be similar to Nimbus?

Study 1b

Figure S1B.3. Business-size manipulation in Study 1b: BIG.

In this study, we are going to ask you some questions about a hypothetical scenario.

Imagine you are attending an out-of-town wedding. Having already arrived at the destination, you realize you forgot to bring your dress shoes. After some searches, you find Nimbus- a big multinational fashion company that sells all kinds of apparel. Nimbus has 1000+ stores and the company owns many popular brands.

You decide to buy the dress shoes from Nimbus and wear those to the wedding. The dress shoes cost $144.95. The shoes fit okay, and you did not experience much discomfort.

However, you already have similar dress shoes at home. You find out that Nimbus accepts returns from dissatisfied customers within 30-days of purchase. And the return process is quick and hassle-free.

* What is a real company that comes to mind which may be similar to Nimbus?

Figure S1B.4. Business-size manipulation in Study 1b: BIG.

In this study, we are going to ask you some questions about a hypothetical scenario.

Imagine you are attending an out-of-town wedding. Having already arrived at the destination, you realize you forgot to bring your dress shoes. After some searches, you find Nimbus- a local store that sells clothes, shoes, and accessories. Nimbus is family-run from a small shop downtown and they sell unique fashion items.

You decide to buy the dress shoes from Nimbus and wear those to the wedding. The dress shoes cost $144.95. The shoes fit okay, and you did not experience much discomfort.

However, you already have similar dress shoes at home. You find out that Nimbus accepts returns from dissatisfied customers within 30-days of purchase. And the return process is quick and hassle-free.

* What is a real shop that comes to mind which may be similar to Nimbus?

Study 1c

Figure S1C.3. Business-size manipulation in Study 1c: BIG.

In this study, we are going to ask you some questions about a hypothetical scenario.

Imagine you are attending an out-of-town wedding. Having already arrived at the destination, you realize you forgot to bring your dress shoes. After some searches, you find Nimbus- a big multinational fashion company that sells all kinds of apparel. Nimbus has 1000+ stores and the company owns many popular brands.

You decide to buy the dress shoes from Nimbus and wear those to the wedding. The dress shoes cost $144.95. The shoes fit okay, and you did not experience much discomfort.

However, you already have similar dress shoes at home. You find out that Nimbus accepts returns from dissatisfied customers within 30-days of purchase. And the return process is quick and hassle-free.

* What is a real company that comes to mind which may be similar to Nimbus?

Figure S1C.4. Business-size manipulation in Study 1c: SMALL.

In this study, we are going to ask you some questions about a hypothetical scenario.

Imagine you are attending an out-of-town wedding. Having already arrived at the destination, you realize you forgot to bring your dress shoes. After some searches, you find Nimbus- a local store that sells clothes, shoes, and accessories. Nimbus is family-run from a small shop downtown and they sell unique fashion items.

You decide to buy the dress shoes from Nimbus and wear those to the wedding. The dress shoes cost $144.95. The shoes fit okay, and you did not experience much discomfort.

However, you already have similar dress shoes at home. You find out that Nimbus accepts returns from dissatisfied customers within 30-days of purchase. And the return process is quick and hassle-free.

* What is a real shop that comes to mind which may be similar to Nimbus?

Study 2a

Figure S2A.3. Business-size manipulation in Study 2a: BIG.

We want to know about your perceptions of a large American retailer. For privacy reasons, we cannot disclose the actual name. Imagine it's called EveryDayMart.

Below is a sample text from their page's "About Us" section.

EveryDayMart is one of the biggest retailers in the country. The company has over 5,000 stores across all states, and employs more than 1.4 million workers.

EveryDayMart offers groceries, electronics, home categories, and many other items at competitive prices. Visit to shop for necessities and more!

* Can you think of a real retailer that might have a similar description?

Figure S2A.4. Business-size manipulation in Study 2a: SMALL.

We want to know about your perceptions of an American corner store. For privacy reasons, we cannot disclose the actual name. Imagine it's called Mike's Store.

Below is a sample text from their page's "About Us" section.

Mike's Store is a small corner store. The store operates on a vibrant street downtown and is run by a local family with the help of part-time workers.

Mike's Store offers groceries, homemade food, and some household items at competitive prices. Visit to shop for necessities and more!

* Can you think of a real store that might have a similar description?

Study 2b

Figure S2B.4. Business-size manipulation in Study 2b: BIG.

We want to know about your perceptions of an American hotel chain. For privacy reasons, we changed the name to Indigo Hotels and Resorts.

Below is a sample text from their "About Us" section.

With more than 8 000 hotels and over 14 000 rooms, Indigo Hotels & Resorts is one of the world's largest hotel chains. As an industry leader, Indigo's's portfolio has 38 brands and 8 000+ properties across 139 countries. Indigo's focus on hospitality gives guests new ways to experience, connect, and expand their worlds.

* Can you think of a real hotel chain that might have a similar description?

Figure S2B.5. Business-size manipulation in Study 2b: SMALL.

We want to know about your perceptions of an American independent hotel. For privacy reasons, we changed the name to Indigo Hotel.

Below is a sample text from their "About Us" section.

With 42 rooms in a convenient downtown location, Indigo Hotel caters to travelers of different needs. As an independent hotel, Indigo prides itself on its charming rooftop lounge and outdoor dining in the patio. Indigo's focus on hospitality gives guests new ways to experience, connect, and expand their worlds.

* Can you think of a real independant hotel that might have a similar description?

Study 2c

Figure S2C.3. Business-size manipulation in Study 2c: BIG.

We want to know about your perceptions of a large American hotel chain. For privacy reasons, we changed the name to Grand Hotels and Resorts.

Below is a sample text from their "About Us" section.

With more than 8 000 hotels and over 14 000 rooms, Grand Hotels & Resorts is one of the world's largest hotel chains. As an industry leader, Grand Hotels & Resorts has 38 brands and 8 000+ properties across 139 countries. Its focus on hospitality gives guests new ways to expand their worlds.

* Can you think of a real hotel chain that might have a similar description?

Figure S2C.4. Business-size manipulation in Study 2c: SMALL.

We want to know about your perceptions of a small American hotel. For privacy reasons, we changed the name to Little Hotel Parkside.

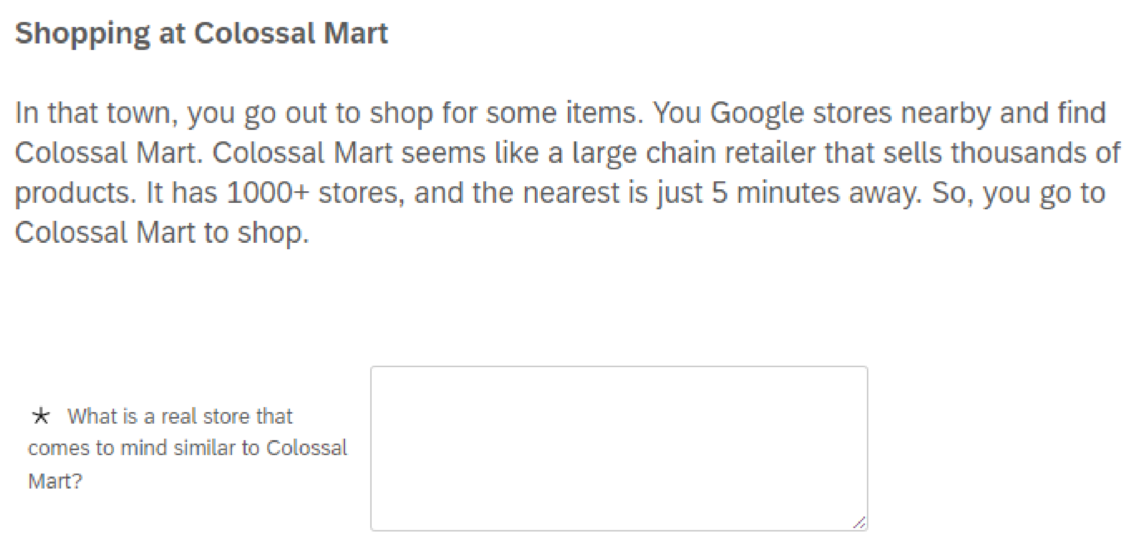
Below is a sample text from their "About Us" section.

With 42 rooms in a convenient location, Little Hotel Parkside caters to travelers of different needs. As a small independent hotel, Little Hotel Parkside prides itself on its quaint rooftop area and outdoor dining in the patio. Its focus on hospitality gives guests new ways to expand their worlds.

* Can you think of a real small hotel that might have a similar description?

Study 3a

Figure S3A.4. Business-size manipulation in Study 3a: BIG.

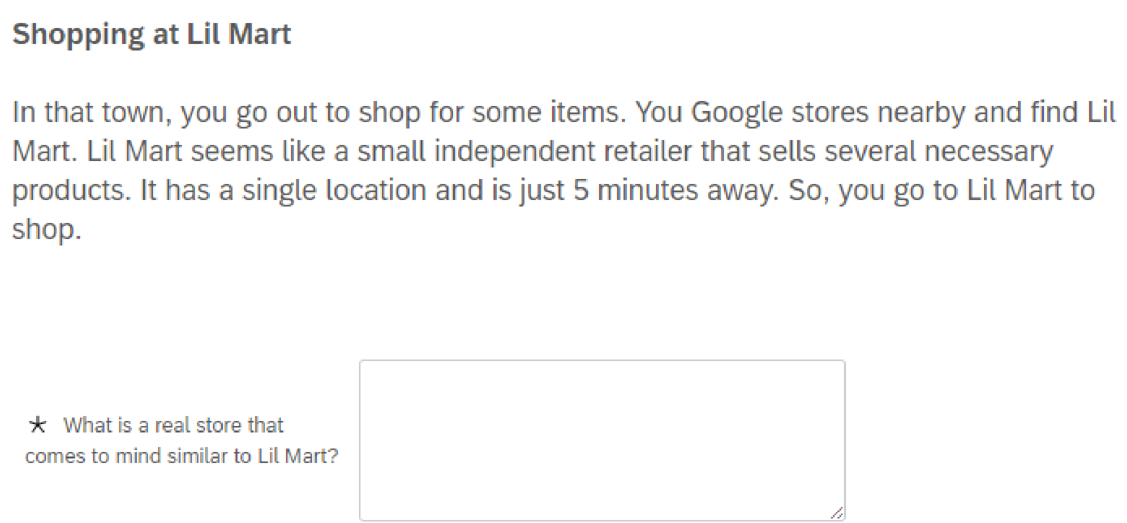


Shopping at Colossal Mart

In that town, you go out to shop for some items. You Google stores nearby and find Colossal Mart. Colossal Mart seems like a large chain retailer that sells thousands of products. It has 1000+ stores, and the nearest is just 5 minutes away. So, you go to Colossal Mart to shop.

* What is a real store that comes to mind similar to Colossal Mart?

Figure S3A.5. Business-size manipulation in Study 3a: SMALL.



Shopping at Lil Mart

In that town, you go out to shop for some items. You Google stores nearby and find Lil Mart. Lil Mart seems like a small independent retailer that sells several necessary products. It has a single location and is just 5 minutes away. So, you go to Lil Mart to shop.

* What is a real store that comes to mind similar to Lil Mart?

Study 3b

Figure S3B.4. Business-size manipulation in Study 3b: BIG.

Shopping at Colossal Mart

In that town, you go out to shop for some items. You Google stores nearby and find Colossal Mart. Colossal Mart seems like a **large chain retailer** that sells thousands of products. It has 1000+ stores, and the nearest is just 5 minutes away. So, you go to Colossal Mart to shop.

What is a real store that comes to mind similar to Colossal Mart?

Figure S3B.4. Business-size manipulation in Study 3b: SMALL.

Shopping at Lil Mart

In that town, you go out to shop for some items. You Google stores nearby and find Lil Mart. Lil Mart seems like a **small independent retailer** that sells several necessary products. It has a single location and is just 5 minutes away. So, you go to Lil Mart to shop.

What is a real store that comes to mind similar to Lil Mart?

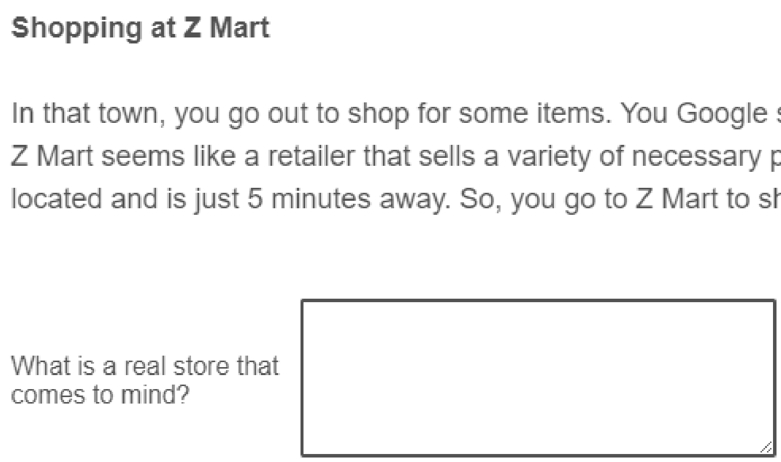
Figure S3B.4. Business-size manipulation in Study 3b: MEDIUM.

Shopping at Mid Mart

In that town, you go out to shop for some items. You Google stores nearby and find Mid Mart. Mid Mart seems like a **medium-sized retailer** that sells a wide range of necessary products. It has several locations, with the nearest one being just 5 minutes away. So, you go to Mid Mart to shop.

What is a real store that comes to mind similar to Mid Mart?

Figure S3B.4. Business-size manipulation in Study 3b: No Size.



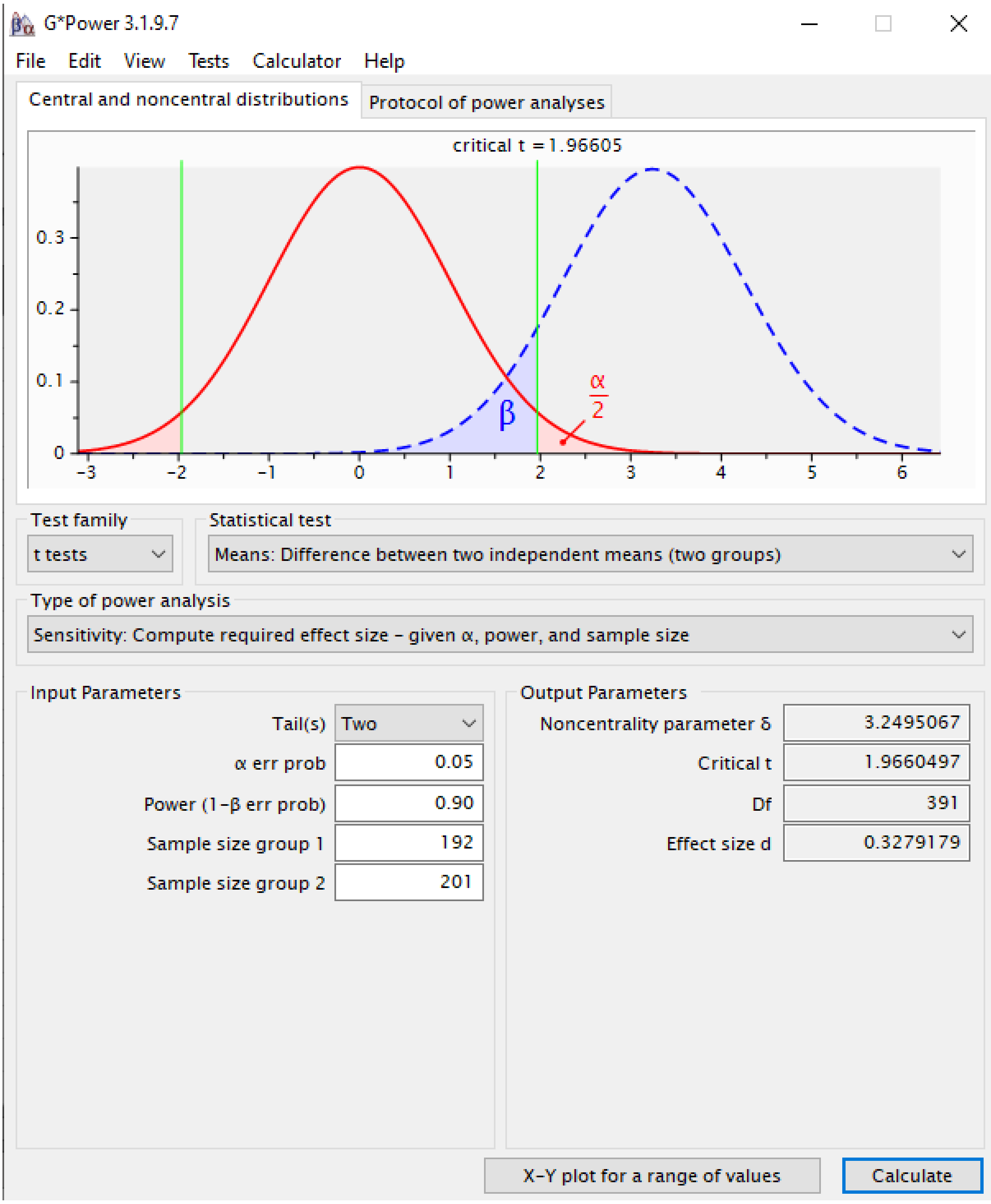
Shopping at Z Mart

In that town, you go out to shop for some items. You Google stores nearby and find Z Mart. Z Mart seems like a retailer that sells a variety of necessary products. It is conveniently located and is just 5 minutes away. So, you go to Z Mart to shop.

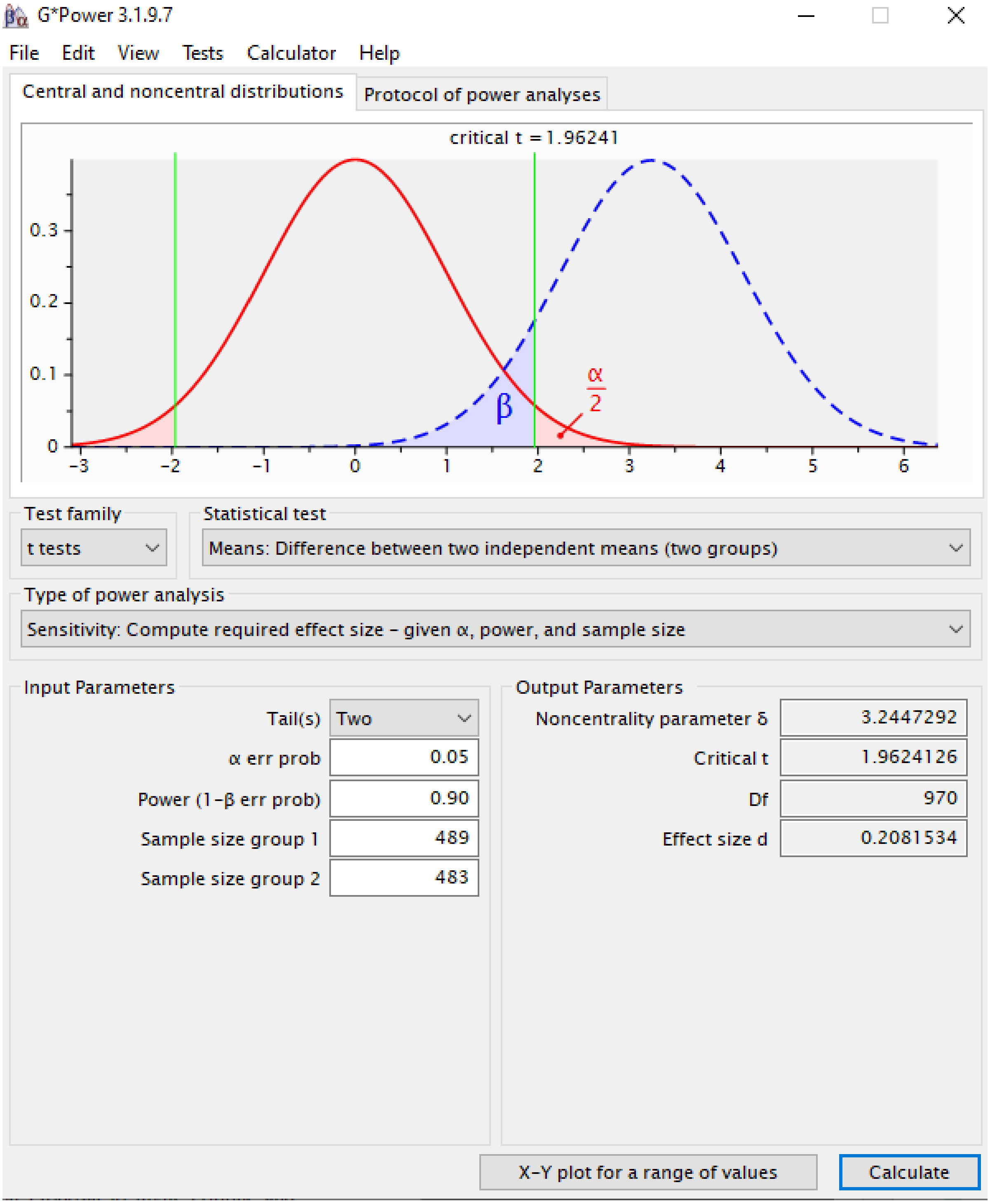
What is a real store that comes to mind?

5. Sensitivity Analyses (power & detectable effect sizes)

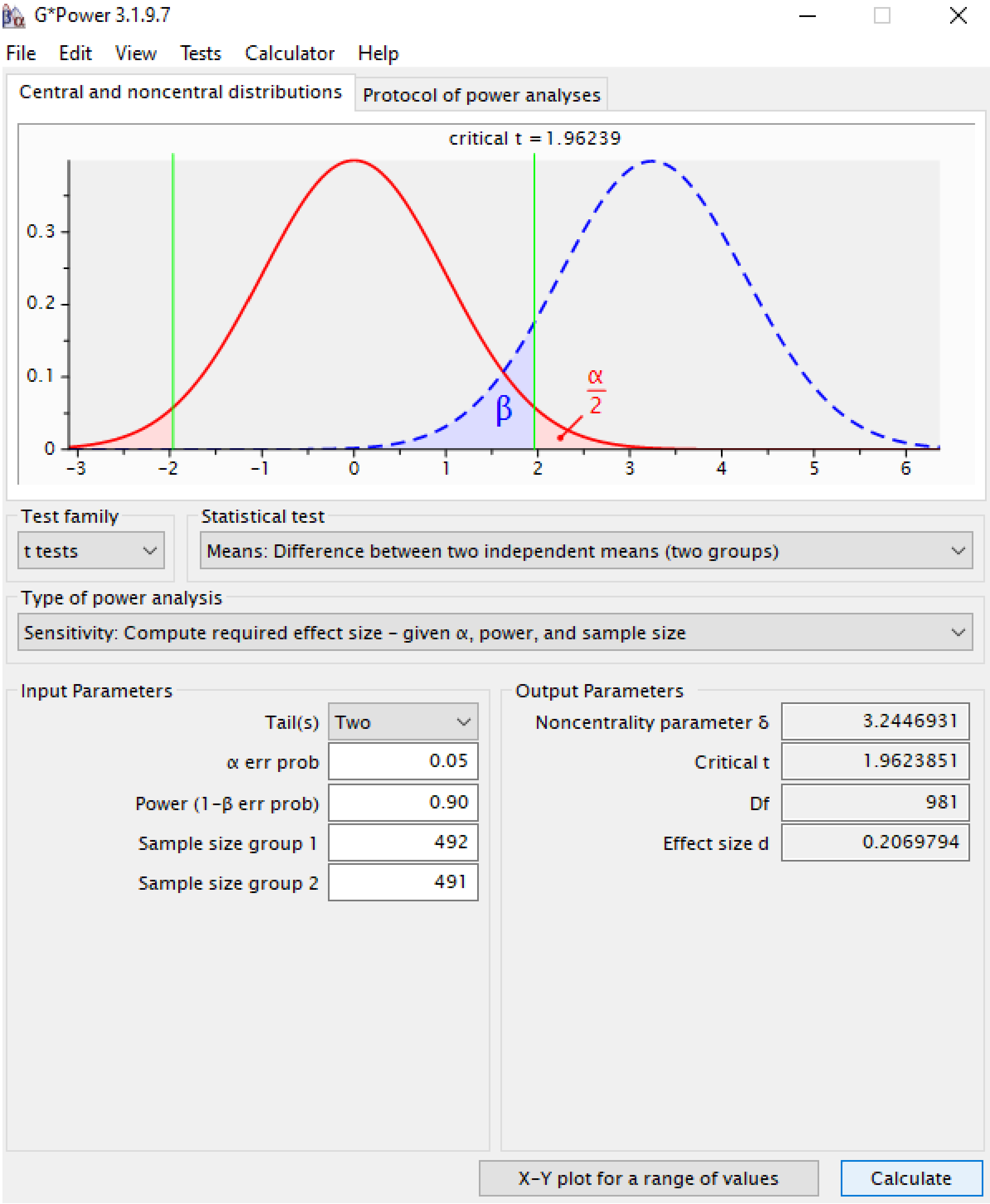
Study 1a



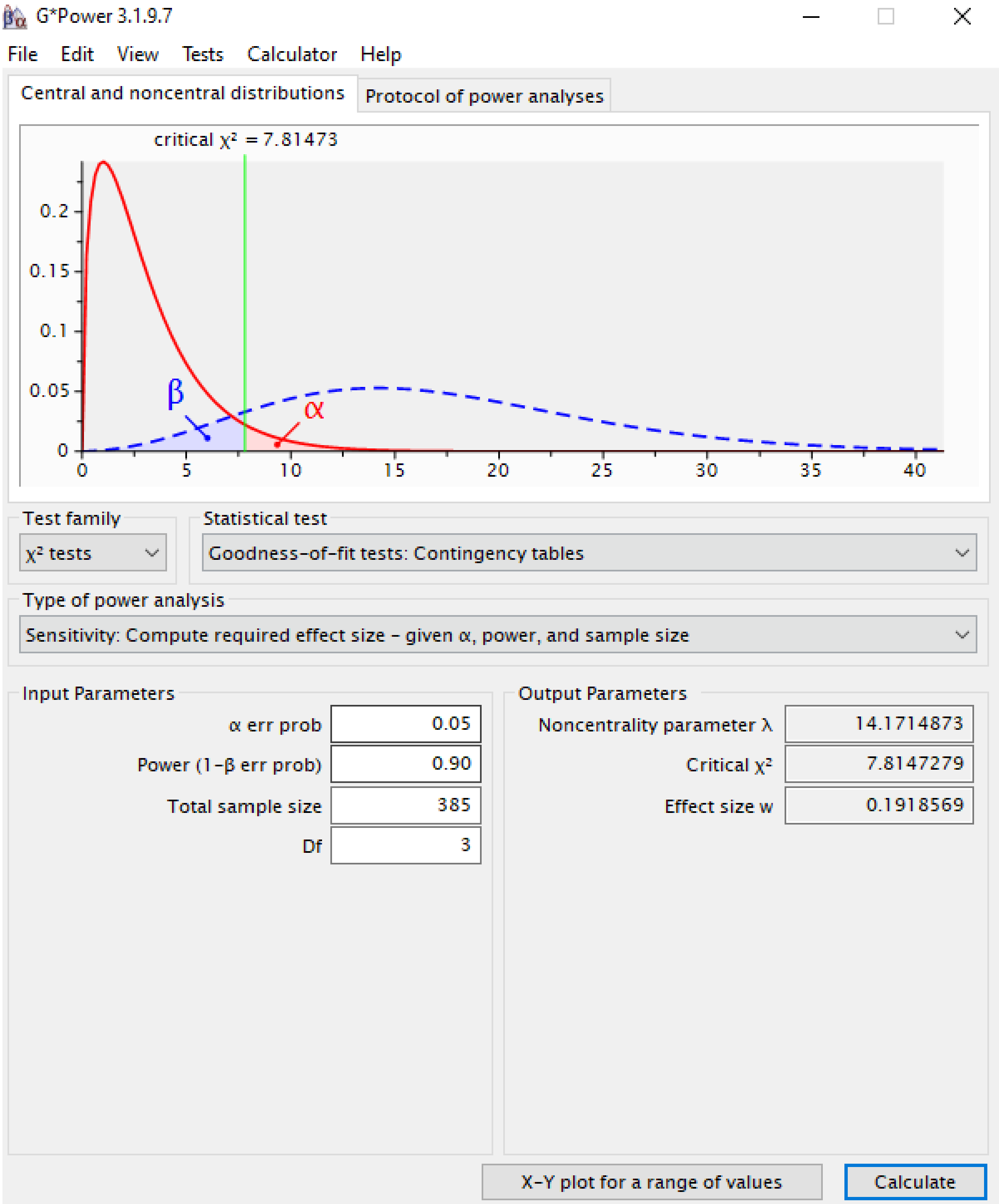
Study 1b


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical t = 1.96241
β
α/2
Test family
t tests
Statistical test
Means: Difference between two independent means (two groups)
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
Tail(s) Two
α err prob 0.05
Power (1–β err prob) 0.90
Sample size group 1 489
Sample size group 2 483
Output Parameters
Noncentrality parameter δ 3.2447292
Critical t 1.9624126
Df 970
Effect size d 0.2081534
X-Y plot for a range of values
Calculate

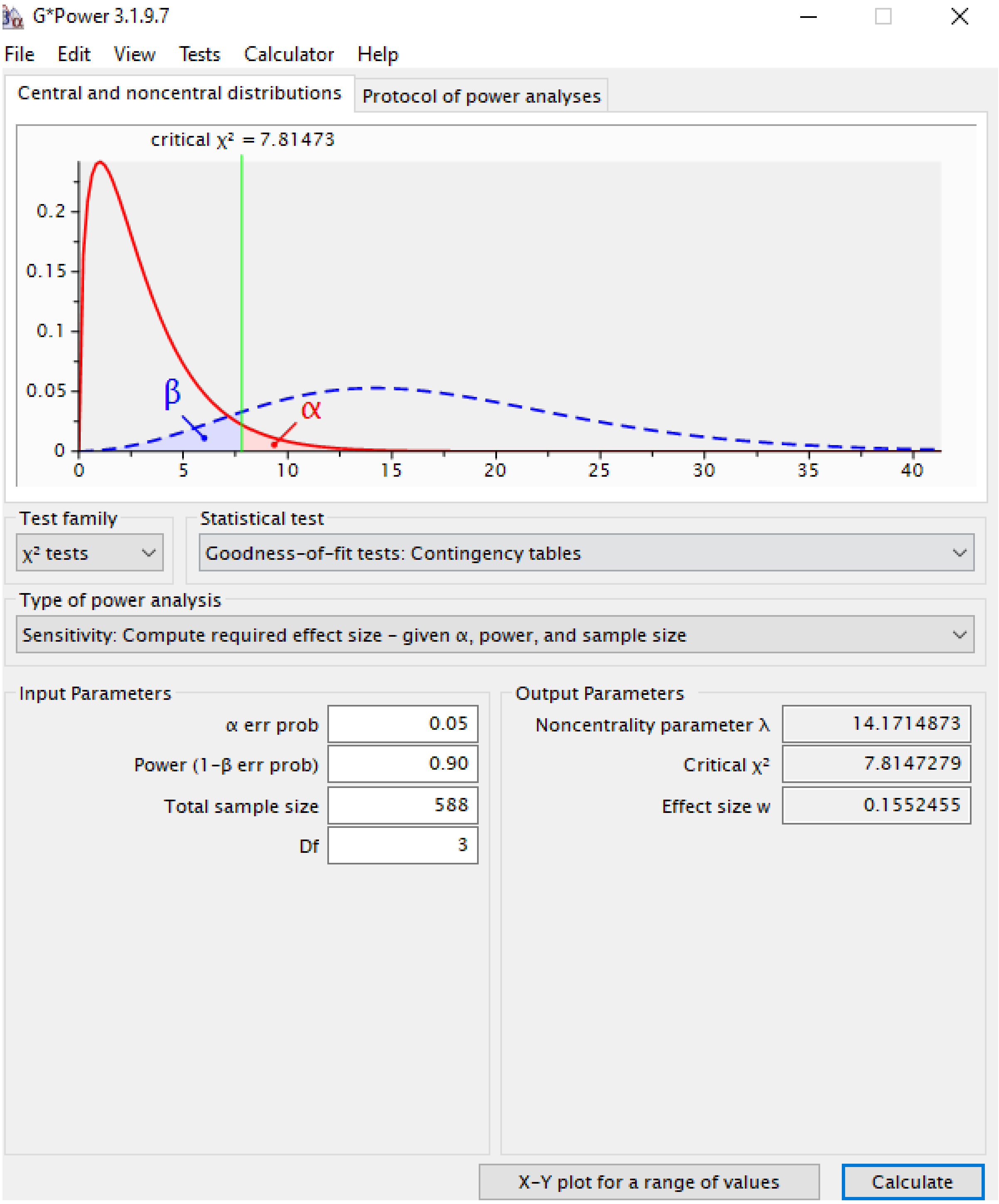
Study 1c


G*Power 3.1.9.7
File
Edit
View
Tests
Calculator
Help
Central and noncentral distributions
Protocol of power analyses
critical t = 1.96239
0.3
0.2
0.1
0
β
α
2
-3
-2
-1
0
1
2
3
4
5
6
Test family
t tests
Statistical test
Means: Difference between two independent means (two groups)
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
Tail(s)
Two
α err prob
0.05
Power (1-β err prob)
0.90
Sample size group 1
492
Sample size group 2
491
Output Parameters
Noncentrality parameter δ
3.2446931
Critical t
1.9623851
Df
981
Effect size d
0.2069794
X-Y plot for a range of values
Calculate

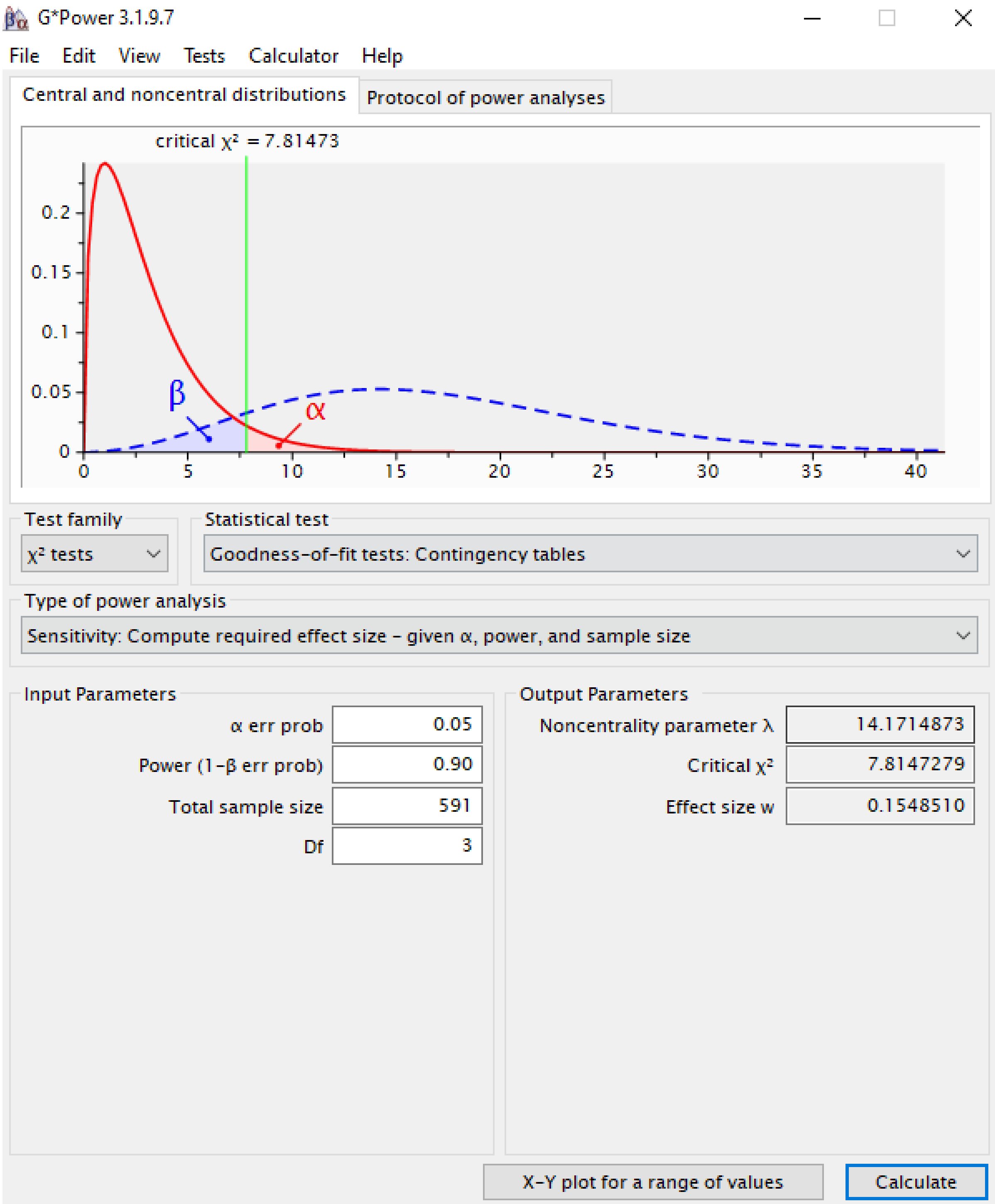
Study 2a


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical χ² = 7.81473
β
α
Test family
χ² tests
Statistical test
Goodness-of-fit tests: Contingency tables
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
α err prob 0.05
Power (1-β err prob) 0.90
Total sample size 385
Df 3
Output Parameters
Noncentrality parameter λ 14.1714873
Critical χ² 7.8147279
Effect size w 0.1918569
X-Y plot for a range of values
Calculate

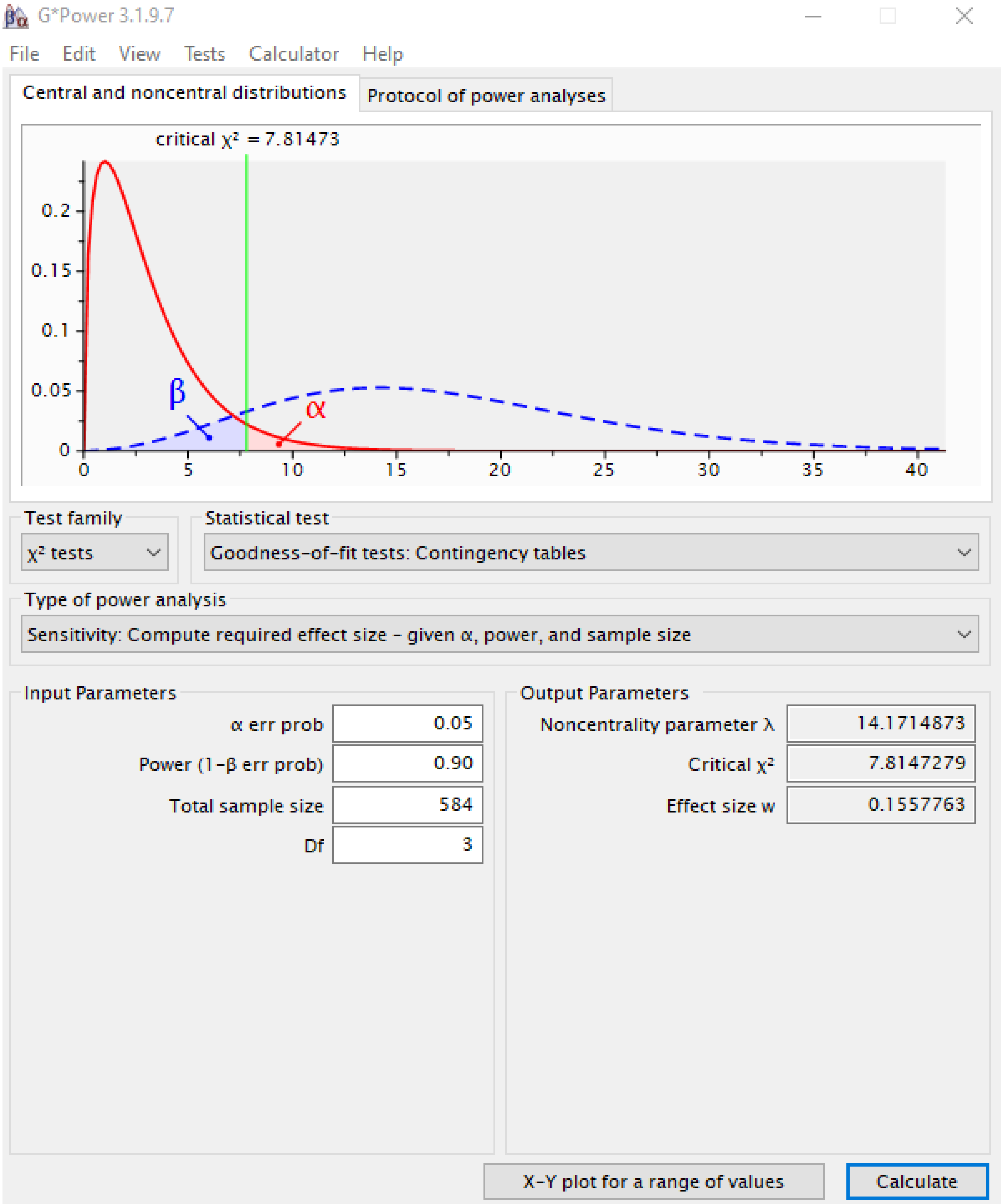
Study 2b


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical χ² = 7.81473
β
α
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25 30 35 40
Test family
χ² tests
Statistical test
Goodness-of-fit tests: Contingency tables
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
α err prob 0.05
Power (1-β err prob) 0.90
Total sample size 588
Df 3
Output Parameters
Noncentrality parameter λ 14.1714873
Critical χ² 7.8147279
Effect size w 0.1552455
X-Y plot for a range of values
Calculate

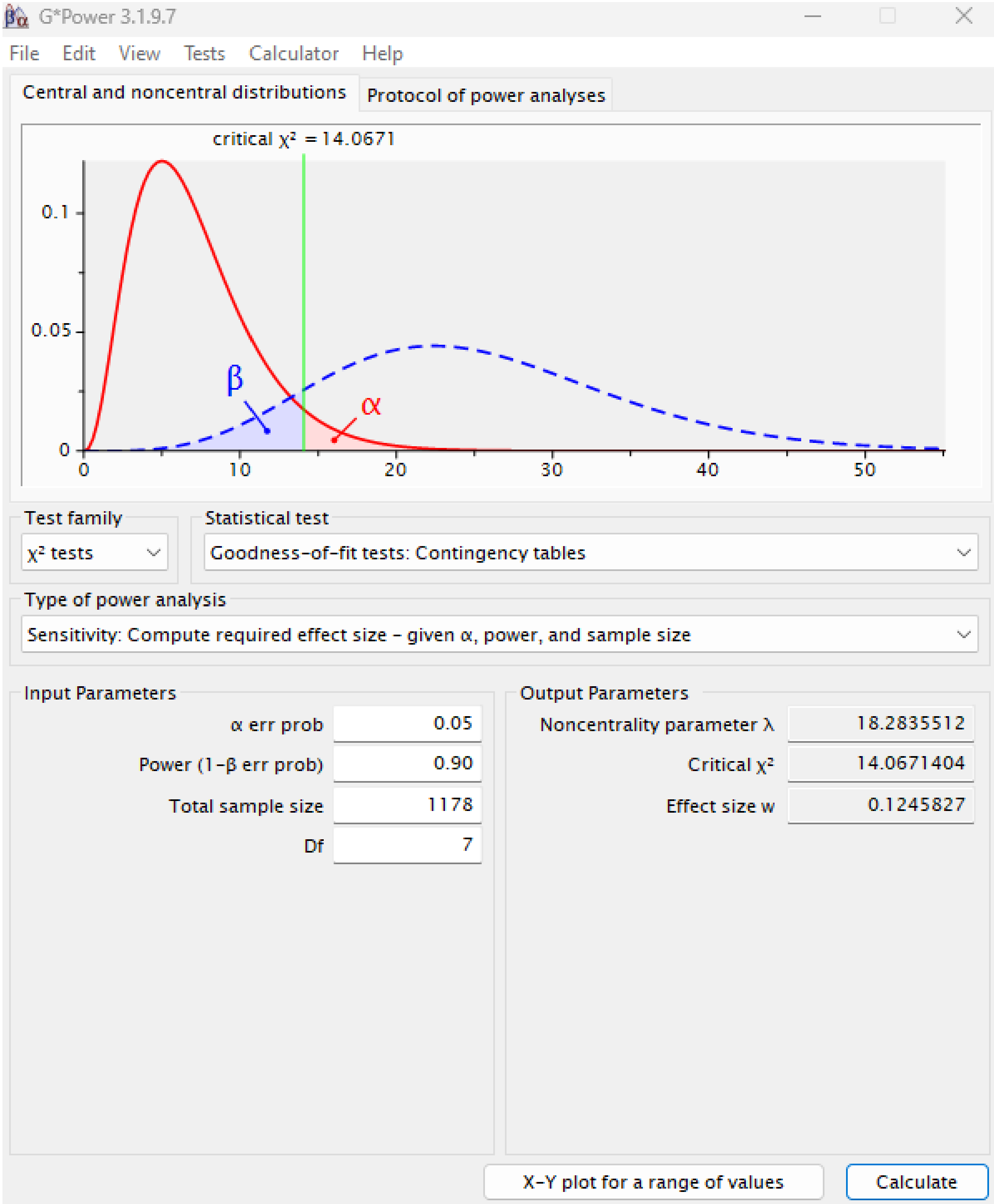
Study 2c


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical χ² = 7.81473
β
α
Test family
χ² tests
Statistical test
Goodness-of-fit tests: Contingency tables
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
α err prob 0.05
Power (1-β err prob) 0.90
Total sample size 591
Df 3
Output Parameters
Noncentrality parameter λ 14.1714873
Critical χ² 7.8147279
Effect size w 0.1548510
X-Y plot for a range of values
Calculate

Study 3a


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical χ² = 7.81473
β
α
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25 30 35 40
Test family
χ² tests
Statistical test
Goodness-of-fit tests: Contingency tables
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
α err prob 0.05
Power (1-β err prob) 0.90
Total sample size 584
Df 3
Output Parameters
Noncentrality parameter λ 14.1714873
Critical χ² 7.8147279
Effect size w 0.1557763
X-Y plot for a range of values
Calculate

Study 3b


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical χ² = 14.0671
0.1
0.05
0
0 10 20 30 40 50
β
α
Test family
χ² tests
Statistical test
Goodness-of-fit tests: Contingency tables
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
α err prob 0.05
Power (1-β err prob) 0.90
Total sample size 1178
Df 7
Output Parameters
Noncentrality parameter λ 18.2835512
Critical χ² 14.0671404
Effect size w 0.1245827
X-Y plot for a range of values
Calculate

5. References

Patil, I. (2021). Visualizations with statistical details: The “ggstatsplot” approach. *Journal of Open Source Software*, *6*(61), 3167. https://doi.org/10.21105/joss.03167

Skowronek, S. E. (2022). DENIAL: A Conceptual Framework to Improve Honesty Nudges. *Current Opinion in Psychology*, 101456. https://doi.org/10.1016/j.copsyc.2022.101456

Supplemental Materials for Article 3

For the manuscript

Intergroup Bias in Dishonesty: Selfish versus Coalitional Lying

Supplemental materials are included to detail all measures, experimental manipulations, and additional analyses that support the current research.

Table of Contents

1. Experiment 1

1.1. List of all measures

Table S1. List of measures – Experiment 1			
Variable	# of items (a)	Question and Items	Response scale/options/box
Age	1	How old are you in years?	Open-ended response box accepting numerical input between 18 and 150.
Gender	1	How do you describe yourself?	Open-ended response box accepting numerical input.
Education	1	What is the highest degree or level of education you have completed?	Multiple choice: Some High School High School Bachelor's Degree Master's Degree Ph.D. or higher Trade School Prefer not to say
Income	1	What is your annual household income?	Multiple choice: Less than \$15,000 \$15,000 - \$24, 999 \$25,000 - \$49,999 \$50,000 - \$100,000 More than \$100, 000 Prefer not to say
Politics	1	Generally speaking, in terms of your political views: Do you usually think of yourself as a Republican, a Democrat, or an Independent?	Multiple choice: Democrat Republican Independent
Vote 2020	1	Which candidate did you vote for in the 2020 Presidential Election?	Multiple choices: Joe Biden Donald Trump Other/Undecided/Prefer not to say
Vote 2024	1	Which candidate do you plan to vote for in the 2024 Presidential Election?	Multiple choices: Joe Biden

			Donald Trump Other/Undecided/Prefer not to say
Die Score	1	What is the number of the die roll?	Multiple choice: 1 2 3 4 5 6
Reported Guess	1	Did you make the correct guess in advance, about whether the die roll would result in an odd number (1, 3, or 5) or an even number (2, 4, or 6)?	Multiple choice: YES, I guessed correctly (You will receive a \$0.75 bonus; your paired ${e://Field/Pair} will not receive any bonus.) NO, I did not guess correctly (Your paired ${e://Field/Pair} will receive a \$0.75 bonus; you will not receive any bonus.)
Ingroup identification (Democrat participants)	14	The statements below assess your reactions to DEMOCRATS. We are not asking about Democratic politicians here. Rather we are asking about people who support the Democratic party and tend to vote for Democratic politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel a bond with Democrats. I feel solidarity with Democrats. I feel committed to Democrats. I am glad to be a Democrat. I think that Democrats have a lot to be proud of. It is pleasant to be a Democrat. Being a Democrat gives me a good feeling. I often think about the fact that I am a Democrat.	1 (Strongly disagree) to 7 (Strongly disagree)

		The fact that I am a Democrat is an important part of my identity. Being a Democrat is an important part of how I see myself. I have a lot in common with the average Democrat. I am similar to the average Democrat. Democrats have a lot in common with each other. Democrats are very similar to each other.	
Ingroup identification (Republican participants)	14	The statements below assess your reactions to REPUBLICANS. We are not asking about Republican politicians here. Rather we are asking about people who support the Republican party and tend to vote for Republican politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel a bond with Republicans. I feel solidarity with Republicans. I feel committed to Democrats. I am glad to be a Republican. I think that Republicans have a lot to be proud of. It is pleasant to be a Republican. Being a Republican gives me a good feeling. I often think about the fact that I am a Republican. The fact that I am a Republican is an important part of my identity. Being a Republican is an important part of how I see myself. I have a lot in common with the average Republican. I am similar to the average Republican. Republicans have a lot in common with each other. Republicans are very similar to each other.	1 (Strongly disagree) to 7 (Strongly disagree)
Sectarianism (Democratic participants)	9	The statements below assess your reactions to REPUBLICANS. We are not asking about Republican politicians here. Rather we are asking	1 (Strongly disagree) to 7 (Strongly disagree)

		about people who support the Republican party and tend to vote for Republican politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel distant from Republicans. I am different from Republicans. No matter how hard I try, I can't see the world the way Republicans do. My feelings toward Republicans are negative. Republicans are immoral. Republicans are evil. Republicans lack integrity.	
Sectarianism (Republican participants)	9	The statements below assess your reactions to DEMOCRATS. We are not asking about Democratic politicians here. Rather we are asking about people who support the Democratic party and tend to vote for Democratic politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel distant from Democrats. I am different from Democrats. No matter how hard I try, I can't see the world the way Democrats do. My feelings toward Democrats are negative. Democrats are immoral. Democrats are evil. Democrats lack integrity.	1 (Strongly disagree) to 7 (Strongly disagree)
Feelings toward the Republican Party	1	On a scale from 0 to 100, how would you rate your feelings towards the Republican Party?	Slider scale from 0 (Very cold or unfavorable) to 100 (Very warm or favorable)
Feelings toward the Democratic Party	1	On a scale from 0 to 100, how would you rate your feelings towards the Democratic Party?	Slider scale from 0 (Very cold or unfavorable) to 100 (Very warm or favorable)

General dishonesty beliefs		What percentage of ${e://Field/Pair}s do you think, completing the same study as you, would falsely state (intentionally) that they guessed correctly in the die-roll game? That is, among all ${e://Field/Pair}s who were not able to correctly guess whether it would be an odd or even number, what percentage of ${e://Field/Pair}s do you think would falsely claim that they guessed correctly? Five participants with the closest guess will receive $1 as additional bonuses within a week of completion of the survey.	Slider scale from 0 to 100.
Intergroup dishonesty beliefs (Pair = Democrat)	1	What percentage of ${e://Field/Pair}s do you think, paired with a Republican Party voter, would falsely state (intentionally) that they guessed correctly? Five participants with the closest guess will receive $1 as additional bonuses within a week of completion of the survey.	Slider scale from 0 to 100.
Intergroup dishonesty beliefs (Pair = Republican)	1	What percentage of ${e://Field/Pair}s do you think, paired with a Democratic Party voter, would falsely state (intentionally) that they guessed correctly? Five participants with the closest guess will receive $1 as additional bonuses within a week of completion of the survey.	Slider scale from 0 to 100.
Feeling no guilt	1	If I were to cheat a ${e://Field/Pair}, I would not feel guilty about it.	1 (Strongly disagree) to 7 (Strongly disagree)
The pair expects to be cheated	1	My paired ${e://Field/Pair} would expect me to cheat them.	1 (Strongly disagree) to 7 (Strongly disagree)
The pair would have cheated	1	My paired ${e://Field/Pair} would have cheated me if I were on the receiving end.	1 (Strongly disagree) to 7 (Strongly disagree)

1.2. Experimental Manipulations (Pairing block)

Pairing

Paired_Participant

Now, you have been **paired with a ${e://Field/Pair}**.

This person has been randomly chosen from a list of ${e://Field/Pair}s.

demo_image

Display this question

If Pair Is Equal to Democratic Party voter

This person voted for **Joe Biden** and the **Democratic Party** in the 2020 presidential election, and plans to do the same in 2024.

repu_image

Display this question

If Pair Is Equal to Republican Party voter

This person voted for **Donald Trump** and the **Republican Party** in the 2020 presidential election, and plans to do the same in 2024.

Confirm_Pair

You will now participate in a die-roll game that will impact both your own and your paired ${e://Field/Pair}'s **earnings**, as it is possible for one of you to win a bonus reward.

Please confirm that you understand before you proceed.

- [ ] I understand that I have been paired with a ${e://Field/Pair} for a die-roll game

Q61

This question lets you record and manage how long a participant spends on this page. This question will not be displayed to the participant.

Import from library | Add new question

1.3. Moderation results of intergroup effects on behavioral dishonesty

Outcome variable: Reported guess of a one-shot die roll being odd or even (1 = Correct, 0 = Incorrect).					
Predictors	*Odds Ratios*	*Odds Ratios*	*Odds Ratios*	*Odds Ratios*	*Odds Ratios*
(Intercept)	0.35 *** (0.27 – 0.45)	0.56 * (0.33 – 0.93)	0.30 *** (0.15 – 0.57)	0.16 *** (0.08 – 0.31)	0.42 ** (0.23 – 0.74)
Condition [Outgroup]	1.00 (0.69 – 1.45)	1.51 (0.71 – 3.19)	3.14 * (1.29 – 7.72)	3.88 ** (1.60 – 9.62)	2.18 (0.98 – 4.86)
Affiliation [Republican]	1.02 (0.72 – 1.46)				
Condition [Outgroup] X Affiliation [Republican]	1.16 (0.69 – 1.94)				
Sectarianism		0.90 (0.80 – 1.01)			
Condition [Outgroup] X Sectarianism		0.92 (0.78 – 1.09)			
Ingroup Identification			1.04 (0.91 – 1.19)		
Condition [Outgroup] X Ingroup Identification			0.79 * (0.65 – 0.95)		
Feelings Ingroup				1.01 * (1.00 – 1.02)	
Condition [Outgroup] X Feelings Ingroup				0.98 ** (0.97 – 0.99)	
Feelings Outgroup					1.00 (0.99 – 1.01)
Condition [Outgroup] X Feelings Outgroup					0.99 (0.98 – 1.00)
Observations	1177	1177	1177	1173	1169
R^2 Tjur	0.001	0.010	0.008	0.008	0.010

** p<0.05 ** p<0.01 *** p<0.001*

1.4. Moderation results of intergroup effects on dishonesty beliefs

Outcome variable: Estimated percentage of dishonest others minus actual percentage of dishonest others.					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	-0.21 (-2.89 – 2.47)	8.27 ** (2.80 – 13.73)	12.16 *** (5.20 – 19.13)	14.76 *** (8.22 – 21.30)	14.78 *** (8.58 – 20.98)
Condition [Outgroup]	16.69 *** (12.78 – 20.60)	-23.43 *** (-31.47 – -15.39)	-13.07 ** (-22.75 – -3.40)	-12.23 * (-21.54 – -2.91)	-18.58 *** (-27.44 – -9.72)
Affiliation [Republican]	2.63 (-1.16 – 6.41)				
Condition [Outgroup] X Affiliation [Republican]	-14.70 *** (-20.21 – -9.19)				
Sectarianism		-1.67 ** (-2.86 – -0.47)			
Condition [Outgroup] X Sectarianism		7.62 *** (5.85 – 9.38)			
Ingroup Identification			-2.34 ** (-3.76 – -0.92)		
Condition [Outgroup] X Ingroup Identification			4.83 *** (2.83 – 6.83)		
Feelings Ingroup				-0.19 *** (-0.28 – -0.10)	
Condition [Outgroup] X Feelings Ingroup				0.30 *** (0.18 – 0.42)	

Feelings Outgroup					-0.18 *** (-0.26 – -0.10)
Condition [Outgroup] X Feelings Outgroup					0.38 *** (0.26 – 0.49)
Observations	1177	1177	1177	1173	1169
R^2 / R^2 adjusted	0.064 / 0.062	0.103 / 0.100	0.053 / 0.051	0.054 / 0.052	0.068 / 0.065

* *p<0.05* ** *p<0.01* *** *p<0.001*

1.5. Moderation results of intergroup effects on intergroup dishonesty beliefs

Outcome variable: Estimated percentage of participants cheating outgroups minus actual percentage of participants cheating outgroups.					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	7.95 *** (5.15 – 10.75)	6.29 * (0.61 – 11.97)	22.90 *** (15.67 – 30.14)	25.30 *** (18.53 – 32.07)	13.17 *** (6.70 – 19.64)
Condition [Outgroup]	8.02 *** (3.94 – 12.10)	-20.48 *** (-28.83 – -12.12)	-18.05 *** (-28.10 – -8.00)	-20.26 *** (-29.91 – -10.61)	-17.25 *** (-26.50 – -8.00)
Affiliation [Republican]	4.59 * (0.64 – 8.54)				
Condition [Outgroup] X Affiliation [Republican]	-12.50 *** (-18.25 – -6.75)				
Sectarianism		0.92 (-0.32 – 2.17)			
Condition [Outgroup] X Sectarianism		5.17 *** (3.34 – 7.00)			
Ingroup Identification			-2.68 *** (-4.15 – -1.21)		
Condition [Outgroup] X Ingroup Identification			4.25 *** (2.17 – 6.32)		
Feelings Ingroup				-0.21 *** (-0.30 – -0.12)	
Condition [Outgroup] X Feelings Ingroup				0.31 *** (0.18 – 0.44)	

Feelings Outgroup					-0.04 (-0.12 – 0.04)
Condition [Outgroup] X Feelings Outgroup					0.26 *** (0.14 – 0.38)
Observations	1177	1177	1177	1173	1169
R^2 / R^2 adjusted	0.017 / 0.015	0.066 / 0.063	0.016 / 0.013	0.022 / 0.019	0.023 / 0.020

* *p<0.05* ** *p<0.01* *** *p<0.001*

1.6. Moderation results of intergroup effects on cheating expectations

Outcome variable: Participants responded to "My paired ${e://Field/Pair} would expect me to cheat them." on a scale from 1 (Strongly disagree) to 7 (Strongly agree).					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	2.56 *** (2.36 – 2.77)	2.70 *** (2.30 – 3.11)	3.84 *** (3.32 – 4.37)	4.02 *** (3.53 – 4.51)	2.88 *** (2.42 – 3.34)
Condition [Outgroup]	1.95 *** (1.65 – 2.25)	-0.68 * (-1.28 – -0.08)	0.05 (-0.68 – 0.77)	0.44 (-0.25 – 1.14)	-0.45 (-1.11 – 0.21)
Affiliation [Republican]	-0.08 (-0.37 – 0.21)				
Condition [Outgroup] X Affiliation [Republican]	-0.46 * (-0.88 – -0.04)				
Sectarianism		-0.04 (-0.13 – 0.05)			
Condition [Outgroup] X Sectarianism		0.56 *** (0.43 – 0.69)			
Ingroup Identification			-0.28 *** (-0.39 – -0.17)		
Condition [Outgroup] X Ingroup Identification			0.36 *** (0.21 – 0.51)		
Feelings Ingroup				-0.02 *** (-0.03 – -0.01)	
Condition [Outgroup] X Feelings Ingroup				0.02 *** (0.01 – 0.03)	
Feelings Outgroup					-0.00 (-0.01 – 0.00)

Condition [Outgroup] X Feelings Outgroup					0.03 *** (0.02 – 0.04)
Observations	1177	1177	1177	1173	1169
R^2 / R^2 adjusted	0.187 / 0.185	0.250 / 0.248	0.198 / 0.196	0.208 / 0.206	0.223 / 0.221

* *p<0.05* ** *p<0.01* *** *p<0.001*

1.7. Moderation results of intergroup effects on lack of anticipated guilt from cheating

Outcome variable: Participants responded to the item "If I were to cheat a ${e://Field/Pair}, I would not feel guilty about it." on a scale from 1 (Strongly disagree) to 7 (Strongly agree).					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	2.52 *** (2.32 – 2.71)	2.08 *** (1.68 – 2.49)	3.27 *** (2.75 – 3.78)	3.39 *** (2.90 – 3.87)	2.34 *** (1.88 – 2.81)
Condition [Outgroup]	0.40 ** (0.11 – 0.69)	-1.19 *** (-1.79 – -0.59)	-0.82 * (-1.54 – -0.10)	-0.64 (-1.33 – 0.05)	-0.61 (-1.27 – 0.05)
Affiliation [Republican]	-0.40 ** (-0.68 – -0.12)				
Condition [Outgroup] X Affiliation [Republican]	-0.44 * (-0.85 – -0.04)				
Sectarianism		0.05 (-0.04 – 0.14)			
Condition [Outgroup] X Sectarianism		0.32 *** (0.19 – 0.45)			
Ingroup Identification			-0.20 *** (-0.31 – -0.10)		
Condition [Outgroup] X Ingroup Identification			0.21 ** (0.06 – 0.36)		
Feelings Ingroup				-0.01 *** (-0.02 – -0.01)	
Condition [Outgroup] X Feelings Ingroup				0.01 * (0.00 – 0.02)	
Feelings Outgroup					-0.00 (-0.01 – 0.01)

Condition [Outgroup] X Feelings Outgroup					0.01 * (0.00 – 0.02)
Observations	1177	1177	1177	1173	1169
R^2 / R^2 adjusted	0.035 / 0.033	0.049 / 0.047	0.014 / 0.012	0.021 / 0.019	0.012 / 0.009

* *p<0.05* ** *p<0.01* *** *p<0.001*

1.8. Moderation results of intergroup effects on perceived reciprocal cheating

Outcome variable: Paired responded to "My paired ${e://Field/Pair} would have cheated me if I were on the receiving end." on a scale from 1 (Strongly disagree) to 7 (Strongly agree).					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	3.15 *** (2.96 – 3.35)	2.94 *** (2.56 – 3.32)	4.69 *** (4.19 – 5.18)	4.76 *** (4.29 – 5.22)	3.14 *** (2.70 – 3.58)
Condition [Outgroup]	1.47 *** (1.19 – 1.75)	-1.33 *** (-1.89 – -0.77)	-1.22 *** (-1.91 – -0.54)	-0.69 * (-1.35 – -0.02)	-0.77 * (-1.39 – -0.14)
Affiliation [Republican]	-0.13 (-0.40 – 0.15)				
Condition [Outgroup] X Affiliation [Republican]	-0.49 * (-0.89 – -0.09)				
Sectarianism		0.03 (-0.05 – 0.12)			
Condition [Outgroup] X Sectarianism		0.59 *** (0.47 – 0.72)			
Ingroup Identification			-0.34 *** (-0.44 – -0.24)		
Condition [Outgroup] X Ingroup Identification			0.52 *** (0.38 – 0.67)		
Feelings Ingroup				-0.02 *** (-0.03 – -0.02)	
Condition [Outgroup] X Feelings Ingroup				0.03 *** (0.02 – 0.04)	
Feelings Outgroup					-0.00 (-0.01 – 0.00)

Condition [Outgroup] X Feelings Outgroup					0.03 *** (0.02 – 0.04)
Observations	1177	1177	1177	1173	1169
R^2 / R^2 adjusted	0.121 / 0.119	0.231 / 0.229	0.149 / 0.147	0.148 / 0.146	0.164 / 0.161

* *p<0.05* ** *p<0.01* *** *p<0.001*

2. Experiment 2

2.1. List of all measures

Table S2. List of measures – Experiment 2			
Variable	# of items (a)	Item	Response scale/options/box
Age	1	How old are you in years?	-100 (No way) to 100 (For sure)
Gender	1	How do you describe yourself?	-100 (Not at all profitable) to 100 (Very profitable)
Education	1	What is the highest degree or level of education you have completed?	Multiple choice: Some High School High School Bachelor's Degree Master's Degree Ph.D. or higher Trade School Prefer not to say
Income	1	What is your annual household income?	Multiple choice: Less than $15,000 $15,000 - $24, 999 $25,000 - $49,999 $50,000 - $100,000 More than $100, 000 Prefer not to say
Politics	1	Generally speaking, in terms of your political views: Do you usually think of yourself as a Republican, a Democrat, or an Independent?	Multiple choice: Democrat Republican Independent
Vote 2020	1	Which candidate did you vote for in the 2020 Presidential Election?	Multiple choice: Joe Biden Donald Trump Other/Undecided/Prefer not to say
Vote 2024	1	Which candidate do you plan to vote for in the 2024 Presidential Election?	Multiple choice: Joe Biden Donald Trump Other/Undecided/Prefer not to say

Die Score	1	What is the number of the die roll?	Multiple Choice: 1 2 3 4 5 6
Reported Guess (Recipient = Democrat)	1	Did you make the correct guess in advance, about whether the die roll would result in an odd number (1, 3, or 5) or an even number (2, 4, or 6)?	Binary choice: YES, I guessed correctly (A Democratic Party voter will receive a $0.75 bonus.) NO, I did not guess correctly (There will be no bonus.)
Reported Guess (Recipient = Republican)	1	Did you make the correct guess in advance, about whether the die roll would result in an odd number (1, 3, or 5) or an even number (2, 4, or 6)?	Binary choice: YES, I guessed correctly (A Republican Party voter will receive a $0.75 bonus.) NO, I did not guess correctly (There will be no bonus.)
Reported Guess (Recipient = Democrat at the cost of a Republican) Democrat participants	1	Did you make the correct guess in advance, about whether the die roll would result in an odd number (1, 3, or 5) or an even number (2, 4, or 6)?	Binary choice: YES, I guessed correctly (A Democratic Party voter will receive a $0.75 bonus.) No, I did not guess correctly (A Republican Party voter will receive a $0.75 bonus.)
Reported Guess (Recipient = Republican at the cost of a Democrat) Republican participants	1	Did you make the correct guess in advance, about whether the die roll would result in an odd number (1, 3, or 5) or an even number (2, 4, or 6)?	Multiple Choice: YES, I guessed correctly (A Republican Party voter will receive a $0.75 bonus.) NO, I did not guess correctly

			(A Democratic Party voter will receive a $0.75 bonus).
Dishonesty beliefs about Democrats		What percentage of Democratic Party voters do you think, completing the same study as you, would falsely state (intentionally) whether they guessed correctly in the die-roll game? That is, among all Democratic Party voters who were not able to correctly guess whether it would be an odd or even number, what percentage of Democratic Party voters do you think would falsely claim that they guessed correctly? Five participants with the closest guess will receive $1 as an additional bonus within a week of completion of the survey.	Slides scale from 0 to 100.
Dishonesty beliefs about Republicans	1	What percentage of Republican Party voters do you think, completing the same study as you, would falsely state (intentionally) whether they guessed correctly in the die-roll game? That is, among all Republican Party voters who were not able to correctly guess whether it would be an odd or even number, what percentage of Republican Party voters do you think would falsely claim that they guessed correctly? Five participants with the closest guess will receive $1 as an additional bonus within a week of completion of the survey.	1 (Strongly disagree) to 7 (Strongly disagree)
Ingroup identification (Democrat participants)	14	The statements below assess your reactions to DEMOCRATS. We are not asking about Democratic politicians here. Rather we are asking about people who support the Democratic party and tend to vote for Democratic politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each.	1 (Strongly disagree) to 7 (Strongly disagree)

		I feel a bond with Democrats. I feel solidarity with Democrats. I feel committed to Democrats. I am glad to be a Democrat. I think that Democrats have a lot to be proud of. It is pleasant to be a Democrat. Being a Democrat gives me a good feeling. I often think about the fact that I am a Democrat. The fact that I am a Democrat is an important part of my identity. Being a Democrat is an important part of how I see myself. I have a lot in common with the average Democrat. I am similar to the average Democrat. Democrats have a lot in common with each other. Democrats are very similar to each other.	
Ingroup identification (Republican participants)	14	The statements below assess your reactions to REPUBLICANS. We are not asking about Republican politicians here. Rather we are asking about people who support the Republican party and tend to vote for Republican politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel a bond with Republicans. I feel solidarity with Republicans. I feel committed to Democrats. I am glad to be a Republican. I think that Republicans have a lot to be proud of. It is pleasant to be a Republican. Being a Republican gives me a good feeling. I often think about the fact that I am a Republican. The fact that I am a Republican is an important part of my identity.	1 (Strongly disagree) to 7 (Strongly disagree)

		Being a Republican is an important part of how I see myself. I have a lot in common with the average Republican. I am similar to the average Republican. Republicans have a lot in common with each other. Republicans are very similar to each other.	
Sectarianism (Democratic participants)	9	The statements below assess your reactions to REPUBLICANS. We are not asking about Republican politicians here. Rather we are asking about people who support the Republican party and tend to vote for Republican politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel distant from Republicans. I am different from Republicans. No matter how hard I try, I can't see the world the way Republicans do. My feelings toward Republicans are negative. Republicans are immoral. Republicans are evil. Republicans lack integrity.	1 (Strongly disagree) to 7 (Strongly disagree)
Sectarianism (Republican participants)	9	The statements below assess your reactions to DEMOCRATS. We are not asking about Democratic politicians here. Rather we are asking about people who support the Democratic party and tend to vote for Democratic politicians. Please carefully read each statement, and indicate to what extent you agree to disagree with each. I feel distant from Democrats. I am different from Democrats.	1 (Strongly disagree) to 7 (Strongly disagree)

		No matter how hard I try, I can't see the world the way Democrats do. My feelings toward Democrats are negative. Democrats are immoral. Democrats are evil. Democrats lack integrity.	
Feelings toward the Republican Party	1	On a scale from 0 to 100, how would you rate your feelings towards the Republican Party?	Slider scale from 0 (Very cold or unfavorable) to 100 (Very warm or favorable)
Feelings toward the Democratic Party	1	On a scale from 0 to 100, how would you rate your feelings towards the Democratic Party?	Slider scale from 0 (Very cold or unfavorable) to 100 (Very warm or favorable)
Intergroup dishonesty beliefs (Pair = Republican)	1	What percentage of ${e://Field/Pair}s do you think, paired with a Democratic Party voter, would falsely state (intentionally) that they guessed correctly? Five participants with the closest guess will receive $1 as additional bonuses within a week of completion of the survey.	Slider scale from 0 to 100.
Feeling no guilt	1	If I were to cheat in this task, I would not feel guilty about it.	1 (Strongly disagree) to 7 (Strongly disagree)
Pair expects to be cheated (1) Recipient = Republican	1	In this task, the Republican Party voter would have expected me to cheat.	1 (Strongly disagree) to 7 (Strongly disagree)
Pair expects to be cheated (2) Recipient = Democrat		In this task, the Democratic Party voter would have expected me to cheat.	1 (Strongly disagree) to 7 (Strongly disagree)
The pair would have cheated (1) Recipient = Republican	1	In this task, the Republican Party voter would have cheated if I was on the receiving end.	1 (Strongly disagree) to 7 (Strongly disagree)
Pair would have cheated (2) Recipient = Democrat	1	In this task, the Democratic Party voter would have cheated if I was on the receiving end.	1 (Strongly disagree) to 7 (Strongly disagree)

2.2. Experimental Manipulations

2.2.1. Outcome Recipient = Democrat

Display this question

If Outcome_Recipient Is Equal to Democrat

Task Description: A die-roll game
In this task, you will guess the outcome of a six-sided die roll.

Specifically, we want you to think about it in advance and **guess in your mind** whether the number that will appear on the die will be an **odd number** (1, 3, or 5) or an **even number** (2, 4, or 6). Do not tell anyone about your guess, just make sure to remember whether you are guessing odd or even.

After the die roll has been completed on the screen and you have observed the result, you should honestly report whether your guess turned out to be correct or not.

Bonus Mechanism
If your guess is **correct**, a **Democratic Party voter** will receive a **$0.75 bonus**. This person voted for **Joe Biden** and the **Democratic Party** in the 2020 presidential election, and plans to do the same in 2024.

If your guess is **incorrect**, there will be **no bonus** for this person.

The outcome of the die-roll game will not affect your own payment in any way. Your identity will remain anonymous.

Important: Do not proceed before you have read the text above
If you have understood the task instructions above, please select "I Understand", and proceed to the next page.

○ I Understand

2.2.2. Outcome Recipient = Republican

Display this question

If Outcome_Recipient Is Equal to Republican

Task Description: A die-roll game
In this task, you will guess the outcome of a six-sided die roll.

Specifically, we want you to think about it in advance and **guess in your mind** whether the number that will appear on the die will be an **odd number** (1, 3, or 5) or an **even number** (2, 4, or 6). Do not tell anyone about your guess, just make sure to remember whether you are guessing odd or even.

After the die roll has been completed on the screen and you have observed the result, you should report honestly whether your guess turned out to be correct or not.

Bonus Mechanism
If your guess is **correct**, a **Republican Party voter** will receive a **$0.75 bonus**. This person voted for **Donald Trump** and the **Republican Party** in the 2020 presidential election, and plans to do the same in 2024.

If your guess is **incorrect**, there will be **no bonus** for this person.

The outcome of the die-roll game will not affect your own payment in any way. Your identity will remain anonymous.

Important: Do not proceed before you have read the text above
If you have understood the task instructions above, please select "I Understand", and proceed to the next page.

○ I Understand

2.2.3. Outcome Recipient = Democrat or Republican

Display this question

If Outcome_Recipient Is Equal to DemOrRep

Task Description: A die-roll game
In this task, you will guess the outcome of a six-sided die roll.

Specifically, we want you to think about it in advance and **guess in your mind** whether the number that will appear on the die will be an **odd number** (1, 3, or 5) or an **even number** (2, 4, or 6). Do not tell anyone about your guess, just make sure to remember whether you are guessing odd or even.

After the die roll has been completed on the screen and you have observed the result, please report whether your guess turned out to be correct or not.

Bonus Mechanism
If your guess is **correct**, a **Democratic Party** voter will receive a **$0.75** bonus. This person voted for **Joe Biden** and the **Democratic Party** in the 2020 presidential election, and plans to do the same in 2024.

If your guess is **incorrect**, a **Republican Party** voter will receive a **$0.75** bonus. This person voted for **Donald Trump** and the **Republican Party** in the 2020 presidential election, and plans to do the same in 2024.

The outcome of the die-roll game will not affect your own payment in any way. Your identity will remain anonymous.

Important: Do not proceed before you have read the text above
If you have understood the task instructions above, please select "I Understand", and proceed to the next page.

○ I Understand

2.3. Moderation results of intergroup effects on behavioral dishonesty.

Outcome variable: Reported guess of a one-shot die roll being odd or even (1 = Correct, 0 = Incorrect).					
Predictors	*Odds Ratios*	*Odds Ratios*	*Odds Ratios*	*Odds Ratios*	*Odds Ratios*
(Intercept)	1.10 (0.87 – 1.39)	2.72 *** (1.70 – 4.41)	1.33 (0.74 – 2.43)	1.30 (0.73 – 2.34)	3.01 *** (1.69 – 5.44)
Condition [Ingroup]	1.39 (1.00 – 1.94)	0.50 (0.24 – 1.01)	0.54 (0.24 – 1.24)	0.52 (0.23 – 1.19)	0.51 (0.22 – 1.18)
Condition [InOverOut]	1.47 * (1.05 – 2.06)	0.28 *** (0.14 – 0.56)	0.85 (0.37 – 1.96)	1.02 (0.45 – 2.32)	0.25 ** (0.11 – 0.57)
Affiliation [Republican]	1.03 (0.74 – 1.43)				
Condition [Ingroup] X Affiliation [Republican]	1.10 (0.69 – 1.76)				
Condition [InOverOut] X Affiliation [Republican]	0.83 (0.52 – 1.33)				
Sectarianism		0.81 *** (0.73 – 0.90)			
Condition [Ingroup] X Sectarianism		1.29 ** (1.10 – 1.51)			
Condition [InOverOut] X Sectarianism		1.45 *** (1.24 – 1.70)			
Ingroup Identification			0.96 (0.84 – 1.09)		
Condition [Ingroup] X Ingroup Identification			1.25 * (1.04 – 1.50)		
Condition [InOverOut] X Ingroup Identification			1.11 (0.92 – 1.33)		
Feelings Ingroup				1.00 (0.99 – 1.01)	

Condition [Ingroup] X Feelings Ingroup				1.01 * (1.00 – 1.03)	
Condition [InOverOut] X Feelings Ingroup				1.00 (0.99 – 1.02)	
Feelings Outgroup					0.99 *** (0.98 – 0.99)
Condition [Ingroup] X Feelings Outgroup					1.01 * (1.00 – 1.03)
Condition [InOverOut] X Feelings Outgroup					1.02 *** (1.01 – 1.03)
Observations	1710	1710	1710	1707	1698
R^2 Tjur	0.007	0.020	0.012	0.012	0.017

* *p<0.05* ** *p<0.01* *** *p<0.001*

2.4. Moderation results of experimental condition effects on beliefs about ingroup dishonesty

Outcome variable: Estimated percentage of dishonest ingroups minus actual percentage of dishonest outgroups.					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	35.53 *** (32.85 – 38.22)	30.94 *** (25.62 – 36.27)	33.49 *** (26.63 – 40.35)	35.12 *** (28.44 – 41.81)	39.60 *** (33.14 – 46.06)
Condition [Ingroup]	-14.58 *** (-18.34 – -10.82)	-9.10 * (-17.10 – -1.11)	-9.27 (-18.72 – 0.18)	-8.70 (-18.09 – 0.69)	-13.59 ** (-23.07 – -4.11)
Condition [InOverOut]	-13.74 *** (-17.52 – -9.95)	-6.70 (-14.69 – 1.28)	-5.08 (-14.66 – 4.50)	-3.04 (-12.42 – 6.34)	-14.92 ** (-24.19 – -5.64)
Affiliation [Republican]	-3.87 * (-7.65 – -0.09)				
Condition [Ingroup] X Affiliation [Republican]	0.11 (-5.23 – 5.45)				
Condition [InOverOut] X Affiliation [Republican]	10.05 *** (4.69 – 15.42)				
Sectarianism		0.63 (-0.56 – 1.82)			
Condition [Ingroup] X Sectarianism		-1.25 (-3.01 – 0.50)			
Condition [InOverOut] X Sectarianism		-0.49 (-2.24 – 1.26)			
Ingroup Identification			0.02 (-1.47 – 1.52)		

Condition [Ingroup] X Ingroup Identification			-1.17 (-3.21 – 0.88)		
Condition [InOverOut] X Ingroup Identification			-0.83 (-2.91 – 1.25)		
Feelings Ingroup				-0.02 (-0.11 – 0.07)	
Condition [Ingroup] X Feelings Ingroup				-0.08 (-0.21 – 0.05)	
Condition [InOverOut] X Feelings Ingroup				-0.08 (-0.21 – 0.05)	
Feelings Outgroup					-0.08 (-0.17 – 0.00)
Condition [Ingroup] X Feelings Outgroup					-0.01 (-0.13 – 0.11)
Condition [InOverOut] X Feelings Outgroup					0.08 (-0.04 – 0.20)
Observations	1710	1710	1710	1707	1698
R^2 / R^2 adjusted	0.072 / 0.069	0.064 / 0.061	0.065 / 0.062	0.068 / 0.065	0.067 / 0.064

* *p<0.05* ** *p<0.01* *** *p<0.001*

2.5. Moderation results of experimental condition effects on beliefs about outgroup dishonesty

Outcome variable: Estimated percentage of dishonest outgroups minus actual percentage of dishonest outgroups.					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	45.18 *** (42.32 – 48.04)	30.74 *** (25.20 – 36.28)	38.66 *** (31.33 – 45.99)	42.49 *** (35.31 – 49.66)	39.46 *** (32.61 – 46.30)
Condition [Ingroup]	-19.83 *** (-23.84 – -15.82)	-17.96 *** (-26.27 – -9.65)	-20.48 *** (-30.58 – -10.38)	-20.63 *** (-30.70 – -10.55)	-20.31 *** (-30.36 – -10.27)
Condition [InOverOut]	-4.21 * (-8.25 – -0.17)	-19.57 *** (-27.88 – -11.27)	-12.63 * (-22.87 – -2.39)	-15.04 ** (-25.10 – -4.97)	-25.95 *** (-35.78 – -16.12)
Affiliation [Republican]	-3.69 (-7.72 – 0.34)				
Condition [Ingroup] X Affiliation [Republican]	6.69 * (1.00 – 12.38)				
Condition [InOverOut] X Affiliation [Republican]	-7.94 ** (-13.67 – -2.22)				
Sectarianism		3.00 *** (1.77 – 4.24)			
Condition [Ingroup] X Sectarianism		0.20 (-1.62 – 2.02)			
Condition [InOverOut] X Sectarianism		2.49 ** (0.67 – 4.31)			
Ingroup Identification			1.06 (-0.54 – 2.66)		
Condition [Ingroup] X Ingroup Identification			0.88 (-1.31 – 3.07)		

Condition [InOverOut] X Ingroup Identification			1.02 (-1.21 – 3.24)		
Feelings Ingroup				0.01 (-0.09 – 0.11)	
Condition [Ingroup] X Feelings Ingroup				0.06 (-0.08 – 0.20)	
Condition [InOverOut] X Feelings Ingroup				0.10 (-0.04 – 0.24)	
Feelings Outgroup					0.05 (-0.04 – 0.14)
Condition [Ingroup] X Feelings Outgroup					0.05 (-0.08 – 0.18)
Condition [InOverOut] X Feelings Outgroup					0.24 *** (0.11 – 0.37)
Observations	1710	1710	1710	1707	1698
R^2 / R^2 adjusted	0.089 / 0.086	0.125 / 0.123	0.077 / 0.075	0.073 / 0.070	0.093 / 0.091

* *p<0.05* ** *p<0.01* *** *p<0.001*

2.6. Moderation results of intergroup effects on cheating expectations

Outcome variable: Participants responded to "In this task, the Democratic/ Republican Party voter would have expected me to cheat." on a scale from 1 (Strongly disagree) to 7 (Strongly agree).

Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	4.43 *** (4.24 – 4.62)	2.61 *** (2.25 – 2.97)	3.39 *** (2.91 – 3.86)	3.61 *** (3.15 – 4.07)	2.72 *** (2.27 – 3.16)
Condition [Ingroup]	-2.13 *** (-2.39 – -1.86)	-0.01 (-0.55 – 0.53)	-0.07 (-0.72 – 0.59)	0.09 (-0.55 – 0.73)	0.06 (-0.59 – 0.71)
Condition [InOverOut]	-1.91 *** (-2.17 – -1.64)	-0.13 (-0.67 – 0.41)	-0.26 (-0.93 – 0.40)	-0.09 (-0.73 – 0.56)	-0.12 (-0.76 – 0.52)
Affiliation [Republican]	-0.09 (-0.36 – 0.17)				
Condition [Ingroup] X Affiliation [Republican]	0.21 (-0.17 – 0.58)				
Condition [InOverOut] X Affiliation [Republican]	0.19 (-0.18 – 0.57)				
Sectarianism		0.42 *** (0.34 – 0.50)			
Condition [Ingroup] X Sectarianism		-0.48 *** (-0.60 – -0.36)			
Condition [InOverOut] X Sectarianism		-0.40 *** (-0.52 – -0.28)			
Ingroup Identification			0.23 *** (0.12 – 0.33)		
Condition [Ingroup] X			-0.44 *** (-0.58 – -0.30)		

Ingroup Identification					
Condition [InOverOut] X Ingroup Identification			-0.35 *** (-0.49 – -0.21)		
Feelings Ingroup				0.01 *** (0.00 – 0.02)	
Condition [Ingroup] X Feelings Ingroup				-0.03 *** (-0.04 – -0.02)	
Condition [InOverOut] X Feelings Ingroup				-0.03 *** (-0.03 – -0.02)	
Feelings Outgroup					0.02 *** (0.02 – 0.03)
Condition [Ingroup] X Feelings Outgroup					-0.03 *** (-0.04 – -0.02)
Condition [InOverOut] X Feelings Outgroup					-0.02 *** (-0.03 – -0.01)
Observations	1710	1710	1710	1707	1698
R^2 / R^2 adjusted	0.241 / 0.239	0.286 / 0.284	0.260 / 0.257	0.271 / 0.269	0.267 / 0.265

* *p<0.05* ** *p<0.01* *** *p<0.001*

2.7. Moderation results of intergroup effects on the lack of anticipated guilt from cheating

Participants responded to the statement "If I were to cheat in this task, I would not feel guilty about it." on a scale from 1 (Strongly disagree) to 7 (Strongly agree).					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	2.64 *** (2.42 – 2.85)	1.46 *** (1.03 – 1.88)	2.70 *** (2.15 – 3.25)	2.67 *** (2.14 – 3.21)	1.63 *** (1.11 – 2.15)
Condition [Ingroup]	-0.10 (-0.40 – 0.20)	0.27 (-0.36 – 0.91)	0.08 (-0.68 – 0.84)	0.55 (-0.21 – 1.30)	-0.16 (-0.92 – 0.60)
Condition [InOverOut]	0.22 (-0.09 – 0.52)	0.13 (-0.50 – 0.77)	0.20 (-0.57 – 0.97)	0.65 (-0.10 – 1.40)	-0.08 (-0.83 – 0.66)
Affiliation [Republican]	-0.32 * (-0.62 – -0.02)				
Condition [Ingroup] X Affiliation [Republican]	-0.14 (-0.57 – 0.28)				
Condition [InOverOut] X Affiliation [Republican]	-0.45 * (-0.88 – -0.02)				
Sectarianism		0.24 *** (0.15 – 0.34)			
Condition [Ingroup] X Sectarianism		-0.11 (-0.25 – 0.03)			
Condition [InOverOut] X Sectarianism		-0.04 (-0.18 – 0.10)			
Ingroup Identification			-0.05 (-0.17 – 0.07)		
Condition [Ingroup] X			-0.06 (-0.22 – 0.11)		

Ingroup Identification					
Condition [InOverOut] X Ingroup Identification			-0.05 (-0.21 – 0.12)		
Feelings Ingroup				-0.00 (-0.01 – 0.00)	
Condition [Ingroup] X Feelings Ingroup				-0.01 (-0.02 – 0.00)	
Condition [InOverOut] X Feelings Ingroup				-0.01 (-0.02 – 0.00)	
Feelings Outgroup					0.01 *** (0.00 – 0.02)
Condition [Ingroup] X Feelings Outgroup					-0.00 (-0.01 – 0.01)
Condition [InOverOut] X Feelings Outgroup					0.00 (-0.01 – 0.01)
Observations	1710	1710	1710	1707	1698
R^2 / R^2 adjusted	0.023 / 0.020	0.029 / 0.026	0.006 / 0.003	0.016 / 0.013	0.021 / 0.018

* *p<0.05* ** *p<0.01* *** *p<0.001*

2.8. Moderation results on intergroup effects on perceived reciprocal cheating

Participants responded to "In this task, the Democratic/ Republican Party voter would have cheated if I was on the receiving end." on a scale from 1 (Strongly disagree) to 7 (Strongly agree).					
Predictors	*Estimates*	*Estimates*	*Estimates*	*Estimates*	*Estimates*
(Intercept)	4.33 *** (4.15 – 4.52)	1.80 *** (1.45 – 2.14)	3.17 *** (2.70 – 3.64)	3.41 *** (2.96 – 3.86)	2.25 *** (1.82 – 2.69)
Condition [Ingroup]	-1.76 *** (-2.02 – -1.50)	0.85 ** (0.33 – 1.37)	0.55 (-0.09 – 1.19)	0.64 * (0.01 – 1.28)	0.55 (-0.09 – 1.19)
Condition [InOverOut]	-1.53 *** (-1.80 – -1.27)	0.60 * (0.08 – 1.12)	0.09 (-0.56 – 0.74)	0.19 (-0.44 – 0.83)	0.30 (-0.33 – 0.92)
Affiliation [Republican]	-0.35 ** (-0.61 – -0.09)				
Condition [Ingroup] X Affiliation [Republican]	0.41 * (0.04 – 0.78)				
Condition [InOverOut] X Affiliation [Republican]	0.27 (-0.10 – 0.64)				
Sectarianism		0.56 *** (0.49 – 0.64)			
Condition [Ingroup] X Sectarianism		-0.57 *** (-0.69 – -0.46)			
Condition [InOverOut] X Sectarianism		-0.48 *** (-0.59 – -0.37)			
Ingroup Identification			0.22 *** (0.12 – 0.33)		
Condition [Ingroup] X			-0.48 *** (-0.62 – -0.34)		

Ingroup Identification					
Condition [InOverOut] X Ingroup Identification			-0.34 *** (-0.48 – -0.19)		
Feelings Ingroup				0.01 *** (0.00 – 0.02)	
Condition [Ingroup] X Feelings Ingroup				-0.03 *** (-0.04 – -0.02)	
Condition [InOverOut] X Feelings Ingroup				-0.02 *** (-0.03 – -0.01)	
Feelings Outgroup					0.03 *** (0.02 – 0.03)
Condition [Ingroup] X Feelings Outgroup					-0.03 *** (-0.04 – -0.02)
Condition [InOverOut] X Feelings Outgroup					-0.02 *** (-0.03 – -0.02)
Observations	1710	1710	1710	1707	1698
R^2 / R^2 adjusted	0.165 / 0.163	0.253 / 0.251	0.186 / 0.184	0.196 / 0.194	0.201 / 0.199

* *p<0.05* ** *p<0.01* *** *p<0.001*

Supplemental Materials for Article 5

For the manuscript

A Registered Report on Gender Bias in Interpersonal Dishonesty:
Are Females and Males Cheated Differently?

Supplemental materials are included to detail all measures, experimental manipulations, and additional analyses, that support the current research.

Table of Contents

1. List of measures

Table S1. List of measures

Variable	# of items (a)	Item	Response scale/options/box.
Age	1	How old are you in years?	Open-ended response box accepting numerical input between 18 and 150.
Income	1	What is your annual household income? Please insert number only.	Open-ended response box accepting numerical input.
Education	1	What is the highest degree or level of education you have completed?	Multiple choice: Some High School High School Bachelor's Degree Master's Degree Ph.D. or higher Trade School Prefer not to say
Dice Rolls Reporting	1	You are now ready to start the task. Please roll a six-sided die privately 10 times and note down the scores below, using numbers from 1 to 6. You can use a physical die or a digital version (simply search "Roll a die" on a DIFFERENT TAB in your web browser). As you input, you can see your potential additional earnings at the bottom of the page. Your paired ${e://Field/Victim} will receive £60 minus your earnings.	Open-ended response box accepting numerical input between 1 and 6 beside each roll. 1st Roll 2nd Roll 3rd Roll 4th Roll 5th Roll 6th Roll 7th Roll 8th Roll 9th Roll 10th Roll
Confirming earnings	1	Please move the sliders below to confirm the total additional earnings you and your paired ${e://Field/Victim} may receive.	Two slider scales from 0 to 60 were set to a constant sum of 60 altogether.

Morality of action	1	I believe my actions in the previous tasks are acceptable.	Slider scale anchored at 0 (Completely disagree) to 10 (Completely agree).
Inferences about the paired participant	1	I believe the other male/ female/ Prolific participant would view my actions as fair.	Slider scale anchored at 0 (Completely disagree) to 10 (Completely agree).
Beliefs about cheating different sexes	3	It is equally unacceptable to cheat men or women for money. It is more acceptable to cheat men than women for money. It is more acceptable to cheat women than men for money.	Slider scales anchored at 0 (Completely disagree) to 10 (Completely agree).
Biological sex	1	What is your biological sex?	Multiple choice: Male Female Intersex Not listed [please specific below]
Gender identity	1	What is gender identity?	Multiple choice: Male Female Nonbinary/ fluid Queer/ Genderqueer Not listed [please specify below]
Sexual orientation	1	What is your sexual orientation?	Multiple choice: Heterosexual Gay or Lesbian Bisexual Queer Pansexual Asexual Not listed [please specify below]
Politics	1	In general, how would you describe your political views?	Slide scale from 0 [Left-wing (Liberal)] to 10 [Right-wing (Conservative)].
Dishonesty beliefs		What would be the average additional earnings of participants completing the exact same study as you? Use the slider below to state your estimate.	Slider scale from 0 to 60.

		Note: If all participants report their 10 die rolls truthfully, statistically speaking, the average earnings should be £35. Ten participants with the closest guess will receive £1 as additional bonuses within a week of completion of the study.	
Cheating expectation	1	My paired ${e://Field/Victim} would have cheated if they were in my shoes.	Slider scales anchored at 0 (Completely disagree) to 10 (Completely agree).
No anticipated guilt	1	If I were to cheat my paired ${e://Field/Victim}, I would not feel guilty.	Slider scales anchored at 0 (Completely disagree) to 10 (Completely agree).
Cheating reciprocity	1	My paired ${e://Field/Victim} would have expected me to cheat them.	Slider scales anchored at 0 (Completely disagree) to 10 (Completely agree).

2. Manipulations

Male

Display this question

If Victim Is Equal to male participant

Your **paired ${e://Field/Victim}** will be randomly chosen from a pool of Prolific participants.

+ Add page break

Female

Display this question

If Victim Is Equal to female participant

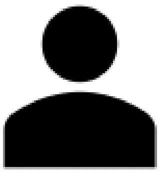
Your **paired ${e://Field/Victim}** will be randomly chosen from a pool of Prolific participants.

Prolific

Display this question

If Victim Is Equal to participant

Your **paired ${e://Field/Victim}** will be randomly chosen from a pool of Prolific participants.



Confirm_Earnings

Your decisions in the task can impact both **your** own and your paired **${e://Field/Victim}'s** earnings. All personal identities will remain completely anonymous.

Please confirm that you understand.

- [ ] I understand that my decisions can impact **my own** and the **paired ${e://Field/Victim}'s** earnings

3. Task instructions

Dice_Instructions

Task: Roll a Die 10 Times

What to do: Roll a six-sided die 10 times. Please do this privately. You can use a physical die or a digital version (simply search "Roll a die" on a DIFFERENT TAB in your web browser).

Then, simply note the 10 numbers from the die rolls, and do not tell anyone about them.

How it affects your earnings:

- **Base Payment**: You already have a guaranteed payment of £0.60 for participating.
- **Additional Earnings**: By completing this task, you will be entered into a lottery. From this lottery, **40 participants** will be randomly selected to receive additional earnings. If you are chosen in the lottery, you will receive an **additional earning** corresponding to the **sum of numbers** of the 10 die rolls. For example, if the sum of your die rolls is 35, you may receive £35 in additional earnings if chosen in the lottery.
- **Remaining Money**: If your additional earning is less than £60, the remaining amount, that is **£60 minus your earnings,** will be given to your **paired ${e://Field/Victim}**.
- **Payment Timeline**: All additional earnings will be paid within two weeks of completion of the study. All personal identities will remain completely anonymous.

Before You Begin: Please ensure you have read and understood these instructions.

Once you are ready, click "I Understand" to start the task.

○ I Understand

+ Add page break

Q39

This question lets you record and manage how long a participant spends on this page. This question will not be displayed to the participant.

4. Exploratory Analyses

4.1. Main effect of target sex on total earnings across datasets

Table S2. Results from main effects (Model 2) of target sex across different datasets. Outcome variable: Total earnings from 10 rounds of self-reported die rolls.

	Pre-registered		**Gender = M/F**		**Heterosexual**		**All complete responses**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	36.40 (36.04 – 36.75)	**<0.001**	36.44 (36.08 – 36.80)	**<0.001**	36.49 (36.09 – 36.88)	**<0.001**	36.32 (35.97 – 36.67)	**<0.001**
Target [Male]	0.47 (-0.03 – 0.97)	0.064	0.41 (-0.09 – 0.92)	0.107	0.25 (-0.31 – 0.81)	0.380	0.53 (0.04 – 1.02)	**0.036**
Target [Neutral]	0.51 (0.01 – 1.01)	**0.045**	0.50 (-0.00 – 1.00)	0.052	0.48 (-0.07 – 1.04)	0.089	0.58 (0.09 – 1.07)	**0.021**
Random effects								
σ^2	34.21		34.17		34.66		35.57	
τ_{00}	0.00 $_{\text{Country}}$		0.00 $_{\text{Country}}$		0.00 $_{\text{Country}}$		0.00 $_{\text{Country}}$	
N	10 $_{\text{Country}}$		10 $_{\text{Country}}$		10 $_{\text{Country}}$		10 $_{\text{Country}}$	
Observations	3166		3110		2570		3380	
Marginal R^2 / Conditional R^2	0.002 / NA		0.001 / NA		0.001 / NA		0.002 / NA	

4.2. Main and interaction effect of target and decision-maker sex across datasets

Table S3. Results from main and interaction effects (Model 2) of target sex and decision-maker sex across different datasets. Outcome variable: Total earnings from 10 rounds of self-reported die rolls.

	Pre-registered		**Gender = M/F**		**Heterosexual**		**All complete responses**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	36.77 (36.27 – 37.27)	**<0.001**	36.75 (36.25 – 37.25)	**<0.001**	36.88 (36.35 – 37.41)	**<0.001**	36.65 (36.16 – 37.15)	**<0.001**
Decision-maker [Female]	-0.74 (-1.44 – -0.03)	**0.040**	-0.63 (-1.34 – 0.08)	0.081	-0.86 (-1.65 – -0.08)	**0.031**	-0.67 (-1.37 – 0.02)	0.058
Target [Male]	-0.28 (-0.99 – 0.44)	0.449	-0.29 (-1.00 – 0.43)	0.429	-0.58 (-1.34 – 0.17)	0.131	-0.18 (-0.88 – 0.52)	0.617
Target [Neutral]	0.35 (-0.35 – 1.06)	0.321	0.38 (-0.32 – 1.09)	0.282	0.22 (-0.53 – 0.96)	0.570	0.39 (-0.30 – 1.08)	0.270
Decision-maker [Female] × Target [Male]	1.47 (0.47 – 2.46)	**0.004**	1.39 (0.39 – 2.40)	**0.007**	1.81 (0.69 – 2.93)	**0.002**	1.39 (0.41 – 2.38)	**0.006**
Decision-maker [Female] × Target [Neutral]	0.30 (-0.70 – 1.29)	0.561	0.21 (-0.79 – 1.22)	0.676	0.58 (-0.54 – 1.69)	0.311	0.40 (-0.59 – 1.38)	0.428
Random Effects								
σ^2	34.13		34.10		34.56		35.49	
τ_{00}	0.00 $_{\text{Country}}$		0.00 $_{\text{Country}}$		0.00 $_{\text{Country}}$		0.00 $_{\text{Country}}$	
N	10 $_{\text{Country}}$		10 $_{\text{Country}}$		10 $_{\text{Country}}$		10 $_{\text{Country}}$	
Observations	3166		3110		2570		3378	
Marginal R^2 / Conditional R^2	0.005 / NA		0.004 / NA		0.005 / NA		0.004 / NA	

4.3. Main effect of target sex on exploratory measures

Table S4. Results from main effects of target sex across three exploratory measures as outcomes.

	No guilt cheating pair		The pair would have cheated		The pair expects to be cheated	
Predictors	*std. Beta*	*p*	*std. Beta*	*p*	*std. Beta*	*p*
(Intercept)	-0.07 (-0.15 – 0.00)	**<0.001**	-0.13 (-0.20 – -0.06)	**<0.001**	-0.08 (-0.19 – 0.02)	**<0.001**
Target [Male]	0.10 (0.02 – 0.19)	**0.017**	0.24 (0.16 – 0.33)	**<0.001**	0.14 (0.06 – 0.23)	**0.001**
Target [Neutral]	0.12 (0.03 – 0.20)	**0.007**	0.15 (0.07 – 0.24)	**<0.001**	0.14 (0.05 – 0.22)	**0.002**
Random Effects						
σ^2	8.13		8.32		8.93	
τ_{00}	0.04 Country		0.03 Country		0.16 Country	
ICC	0.01		0.00		0.02	
N	10 Country		10 Country		10 Country	
Observations	3166		3166		3166	
Marginal R^2 / Conditional R^2	0.003 / 0.008		0.010 / 0.014		0.004 / 0.022	

4.3. Main and interaction effect of target and decision-maker sex on exploratory measures

Table S5. Results from main and interaction effects of target and decision-maker sex across three exploratory measures as outcomes.

	No guilt cheating pair		**The pair would have cheated**		**The pair expects to be cheated**	
Predictors	*std. Beta*	*p*	*std. Beta*	*p*	*std. Beta*	*p*
(Intercept)	-0.00 (-0.10 – 0.10)	**<0.001**	-0.05 (-0.15 – 0.04)	**<0.001**	0.01 (-0.11 – 0.13)	**<0.001**
Decision-maker [Female]	-0.14 (-0.26 – -0.02)	**0.020**	-0.15 (-0.27 – -0.03)	**0.014**	-0.19 (-0.31 – -0.07)	**0.001**
Target [Male]	0.11 (-0.01 – 0.23)	0.073	0.23 (0.11 – 0.35)	**<0.001**	0.19 (0.07 – 0.31)	**0.002**
Target [Neutral]	0.16 (0.04 – 0.28)	**0.007**	0.20 (0.09 – 0.32)	**0.001**	0.16 (0.04 – 0.28)	**0.009**
Decision-maker [Female] × Target [Male]	-0.01 (-0.18 – 0.16)	0.904	0.04 (-0.13 – 0.20)	0.685	-0.09 (-0.25 – 0.08)	0.317
Decision-maker [Female] × Target [Neutral]	-0.10 (-0.27 – 0.07)	0.251	-0.11 (-0.28 – 0.06)	0.202	-0.05 (-0.22 – 0.12)	0.544
Random Effects						
σ^2	8.07		8.25		8.80	
τ_{00}	0.04 $_{\text{Country}}$		0.03 $_{\text{Country}}$		0.16 $_{\text{Country}}$	
ICC	0.01		0.00		0.02	
N	10 $_{\text{Country}}$		10 $_{\text{Country}}$		10 $_{\text{Country}}$	
Observations	3166		3166		3166	

Marginal R^2 / Conditional R^2	0.011 / 0.017	0.019 / 0.022	0.019 / 0.036