İSTANBUL
UNIVERSITY
P R E S S

Check for
updates

The Intention-Behavior Gap in Moral Psychology: Moral Behavior and Economic Games

Gülce Günaydın [1] ✉, Ensar Acem [2] , Ozan Isler [3] & Onurcan Yılmaz [2]

[1] Vrije Universiteit Amsterdam, Faculty of Behavioural and Movement Sciences, Department of Experimental and Applied Psychology, Amsterdam, The Netherlands
[2] Kadir Has University, Faculty of Economics, Administrative and Social Sciences, Department of Psychology, İstanbul, Türkiye
[3] The University of Queensland, School of Economics, Brisbane, Australia

Abstract

Understanding how social factors shape human behavior is a central focus and defining feature of social psychology. While the field's golden age is often associated with behavioral studies, there has been a gradual shift from behavioral measurements to self-report methods, which are less costly in terms of both time and money. However, this shift has revealed a significant limitation: intentions measured through self-report methods often fail to accurately predict actual behavior. This chapter highlights the critical importance of directly measuring behavior in moral psychology and explores the use of behavioral economics methods for assessing moral actions. Key tools such as the Dictator Game (measuring generosity), the Trust Game (measuring trust), the Public Goods and Prisoner's Dilemma Games (measuring cooperation), the Third-Party Punishment Game (measuring punitive behavior), and the Monetary Moral Dilemma Game (assessing utilitarian versus deontological decision-making) are introduced, along with general findings from these games. Additionally, important considerations for the effective application of these economic games are emphasized. Emerging methods for measuring moral behavior, as well as how integrating economic games into these approaches can provide valuable methodological contributions to the field of moral psychology, are discussed.

✉ Corresponding author: Gülce Günaydın g.gunaydin@vu.nl

The Evolution of Measuring Human Behavior and Intentions in Social Psychology

Social psychology studies how social factors influence human behavior. The golden age of the field (circa 1950 to 1970) is closely associated with famous experiments observing human behavior. While behavioral studies continue to provide critical insights into leading research topics in the field, a review of the history of social psychology reveals that behavioral measurements have not been at the forefront of social psychological research for a long time (Yilmaz & Bahçekapili, 2019).

The scientific method of introspection, developed by Wilhelm Wundt, was foundational in early psychological research. Introspection involves an individual systematically describing their own mental processes and is considered the basis of the self-report methods widely used today. However, the use of introspection began to decline with the emergence of the behaviorist movement in the 1920s, which emphasized the direct measurement of behavior through experimental methods.

A famous example of an early behavioral experiment in social psychology is Muzafer Sherif's autokinetic effect study (Sherif, 1935). In the experiment, participants were shown a stable light source in a dark room that, due to the absence of a reference point, appeared to be moving. Participants were then asked to estimate the distance of the light moved. The estimates participants made individually differed significantly from the estimates they made later in a group setting. Similarly, when estimates were reported first in a group setting and then individually, individual estimates continued to be shaped by social information.

A similar phenomenon was observed in Solomon Asch's (1951) conformity experiment. A group of confederates, posing as regular participants, formed a group with actual participants in the same experiment. Participants were asked to identify which of the three lines matched the length of a fourth reference line presented to them. Although the correct answer was obvious, after observing each confederate give a specific incorrect response, nearly a third of the participants chose the same incorrect answer. In contrast, nearly all participants who answered the question individually gave the correct response.

Another famous experiment shedding light on the role of social influence on human behavior is Stanley Milgram's (1963) obedience study. Participants were paired with a confederate and assigned the role of "teacher" in a study ostensibly designed to examine the effects of punishment on learning. The experimenter instructed participants to deliver increasingly higher voltage electric shocks as punishment for any incorrect answers provided by the "learner" (a confederate), who was hidden from view but could be heard. Despite the learner's escalating complaints of pain as the (simulated) shocks intensified, the experimenter's authoritative demeanor led 65% of participants to administer the maximum shock level of 450 volts, though many expressed discomfort and hesitated to continue.

What makes these experiments stand out milestones in the history of social psychology is their ability to vividly and concretely capture the influence of various social factors on human behavior in controlled experimental settings. Considering the difficulty of directly observing psychological phenomena (e.g., self-report methods), these studies underscore the methodological strength of behavioral measures in experimental psychology.

The academic "publish or perish" culture that emerged in the early 1980s changed the priorities of social psychology research. This pressure became one of the primary reasons for a shift back from behavioral measurements to self-report methods in psychology (Baumeister et al., 2007). The ability of self-report methods to allow researchers to gather large amounts of data in a shorter time made them more appealing than the time-intensive and effort-heavy behavioral measurements that often

required field research. Over time, behavioral measurements were increasingly replaced by self-report methods and assessments of intentions to perform behaviors.

However, it is well-established that self-report methods often fall short in accessing the underlying mental processes of behavior (Nisbett & Wilson, 1977). Furthermore, the way questions are posed in such studies can significantly influence participants' responses (Schwarz, 1999). At this juncture, examining the relationship between intention measurements derived from self-reports and actual behaviors becomes crucial. Such scrutiny is essential for ensuring that the study of human behavior—a cornerstone of social psychology—continues to rest on solid and reliable foundations.

Intention-Behavior Gap

The predictive power of intention measurements for actual behavior has been a significant topic of inquiry. Sheeran and Webb (2016) found that only about half of intended behaviors translate into action. A meta-analysis synthesizing past meta-analyses determined that intention measurements explain only 28% of behavioral variance (Sheeran, 2002). This phenomenon, termed the "intention-behavior gap," highlights the limitations of self-reported intentions in accurately predicting behavior. This inconsistency is also evident in the discrepancies between neural activities observed during decisions made in hypothetical versus real-life situations (Camerer & Mobbs, 2017). Similar findings have been reported in consumer behavior studies. For instance, the ethical criteria consumers claim to prioritize in product choices often do not align with their actual purchasing decisions (Auger & Devinney, 2007). Field studies further illustrate this gap. For example, the actual number of reported sexual harassment incidents falls significantly short of the self-reported intentions to report such cases (Goodwin et al., 2020). The intention-behavior gap underscores substantial methodological challenges in many research domains, including social psychology, where self-reported data are widely used. Addressing these challenges is crucial for improving the reliability and validity of findings in studies relying on self-reporting.

The intention-behavior gap can be seen as part of the broader challenge of predicting human behavior. Before Milgram's (1963) obedience experiment, a group of students predicted that only 1% of participants would deliver the maximum shock. Contrary to these expectations, the majority of participants reached the maximum voltage, defying both the students' and Milgram's own expectations. Similarly, Asch's (1951) conformity experiment revealed that a significant portion of participants gave incorrect answers to an obviously simple question due to group influence—an outcome that would have been difficult to predict or capture through self-reports. These findings underscore that some of the most compelling discoveries in social psychology have been made possible through the direct measurement of behavior rather than relying solely on intentions or self-reported data. While intention measurements are valuable, especially in areas where direct behavioral measurement is challenging or unethical, they often fall short of providing accurate insights into actual behavior. The replication crisis in psychology has been partially attributed to the over-reliance on intention measurements for practical reasons, with the discrepancy between intention and behavior often overlooked (Camerer et al., 2018).

Intention-based measurements are also a staple in moral psychology research (Yilmaz, 2022), with many studies focusing on participants' hypothetical moral decisions. A well-known example is the Trolley Dilemma (Foot, 1967). In this scenario, a runaway trolley is headed toward five people, but it can be diverted to a track where only one person is present by pulling a lever. Not pulling the lever represents a deontological choice, adhering to the principle that killing is inherently wrong. Conversely, pulling the lever corresponds to a utilitarian decision aimed at maximizing overall welfare. Given the ethical limitations in studying such scenarios behaviorally, decisions are measured only at the level of intention, leaving no room to compare intentions with actual behaviors.

The limited number of studies that have successfully examined both intentions and behaviors found an intention-behavior gap in moral decisions. Bostyn et al. (2018) placed participants in an experimental setting involving deception where five mice were ostensibly at risk of receiving an electric shock. Participants were given the choice to redirect the shock to a single mouse. The results showed that participants were approximately twice as likely to make utilitarian decisions in (what they thought were) real-life situations compared to hypothetical scenarios, choosing to save five mice at the expense of one. Similarly, FeldmanHall et al. (2012) asked participants how much of their money they would spend to reduce the intensity of harmless electric shocks administered to another individual. One group answered hypothetically, while another made decisions in a real scenario involving actual shocks. The study found that participants spent significantly less money in real scenarios compared to their hypothetical responses. A separate group predicting the results of this study incorrectly estimated that real-life decisions would align with hypothetical ones. Interestingly, when the ecological validity of hypothetical scenarios was incrementally increased by providing more contextual details, the decisions participants made gradually approached those observed in real scenarios. The intention-behavior gap can be relevant in not only moral but also social dilemmas involving decisions of cooperation and trust as well as generosity, where numerous factors (e.g., individual, social, cultural, contextual, etc.) could widen the gap between intention and behavior. Social desirability, for instance, is a known influence on self-reported data (Furnham, 1986) and may lead participants to present themselves as more prosocial than they actually are. Therefore, the most insightful information about moral behaviors is often obtained through behavioral measurements, offering a more accurate depiction of human actions in moral contexts.

Experimental economists have a longstanding tradition in measuring individual and social behavior. They regularly employ decision tasks based on game theory incentivized by actual monetary consequences to test the theoretical predictions of game theory's "rationalist" perspective on human behavior. These economic games provide a systematic approach to observe behavior in moral psychological research as well. The next section will provide an overview of the design and common findings of the most relevant economic games for measuring moral or prosocial behaviors including generosity and fairness, trust and trustworthiness, cooperation, norm enforcement, and deontology versus utilitarianism.

Observing Moral Behavior with Economic Games

Generosity and Fairness

Generosity can be defined as personally costly acts of assistance, often performed without expecting direct reciprocity. These behaviors are costly for the individual providing the help, as they require a sacrifice of resources possessed by the helper (Thielmann et al., 2021). Examples of generosity in everyday life include donating blood or contributing to charitable organizations.

Generosity is frequently measured using the Dictator Game (Engel, 2011; Kahneman et al., 1986). The game is played between two participants, one of whom starts with a sum of money and is referred to as the "dictator." The other participant begins the game without any initial funds and plays a passive role. The dictator determines how much of their initial money to give to the passive participant, keeping any remaining amount for themselves. The amount transferred by the dictator is considered a behavioral measure of their generosity.

If we consider humans as rational beings who make optimal decisions solely based on the goal of maximizing self-interest (i.e., homo economicus), it would be expected that the player in the dictator role, motivated by selfishness, would not transfer any money to the passive player. However, research

findings contradict this expectation. According to a meta-analysis, on average, dictators transfer 28.3% of their initial funds to the passive player (Engel, 2011).

The amount transferred is influenced by various factors. For instance, when dictators earn the money through their own efforts, the amount they transfer to their partner decreases (Cherry et al., 2002). Another factor is the amount of information participants have about each other. Dictators are more inclined to donate when their identity is not anonymous, meaning that the passive player knows who the dictators are (Engel, 2011). Similarly, games in which both sides remain entirely anonymous can weaken the dictators' belief that they are interacting with a real partner, which leads to a reduction in the amount donated (Frohlich et al., 2001).

The anonymity of the dictator toward the experimenters also affects their decisions. When dictators are assured that their decisions will remain unknown to the experimenters, the amount of donations decrease. This suggests that decisions are not solely motivated by a desire for fair distribution but also by a desire to avoid being judged negatively for an unfair allocation (Hoffman et al., 1994). When the recipient is perceived to deserve the donation (e.g., in scenarios involving reputable charities), the transferred amount increases (Eckel & Grossman, 1996).

In games involving charitable organizations, donations made in the Dictator Game have been shown to correlate with real-life giving. For example, when participants were asked about their donations to charities in the previous year, those who chose to donate in the game were found to have donated more in real life compared to others (Bekkers, 2007). Another study that sheds light on the validity of the Dictator Game outside laboratory settings was conducted by Stoop (2014). This study compared laboratory and field experiments in which participants encountered a misdelivered envelope with money. The rates at which participants returned the envelope to its rightful recipient were observed to be very similar in both the laboratory and field settings.

The Ultimatum Game (Güth et al., 1982) adds a strategic layer to splitting a sum of money between two players. First, the first player proposes a division. Following the first player's proposal, the second player decides whether to accept or reject the offer. If the offer is accepted, the money is distributed as proposed by the first player; if it is rejected, neither player receives any money. The Ultimatum Game is often viewed as a representation of bargaining scenarios encountered in daily or business life. Assuming complete self-regard, the first player is expected to offer the smallest possible amount to the second player, and the second player is expected to accept any offer greater than zero.

However, similar to the Dictator Game, the results of the Ultimatum Game diverge from these predictions. On average, the first player gives 40% of the initial amount to the second player (Camerer, 2003). The amount transferred to partners is substantially higher in the Ultimatum Game than in the Dictator Game. This difference can be attributed to the risk posed by the second player's ability to reject unfair offers, which would result in neither party receiving any money (Suleiman, 1996). The likelihood of an offer being rejected is closely tied to its size: offers of 40% or more are usually accepted, while offers of 20% or less approach zero acceptance rates (Güth & Kocher, 2014).

A critical factor determining the rejection of an offer is not merely the absolute amount offered but its proportion relative to the amount the first player retains for themselves (Bearden, 2001). One of the main motivations for the second player to reject an offer is to punish the first player, aligning with this finding (Bolton & Zwick, 1995). Another factor influencing the decisions of the second player is their expectations regarding the first player's offer. The likelihood of accepting an offer decreases as the difference between the second player's expectations and the actual offer becomes more pronounced, particularly when expectations regarding the offer amount are high (Doğruyol et al., 2021; Vavra et al., 2018). How players are assigned their roles also affects their decisions. In a standard Ultimatum Game,

roles are assigned randomly. However, when players earn the first-player role through individual effort, the offers made to the second player tend to be lower (Hoffman et al., 1994).

Trust and Trustworthiness

Mutual trust is an indispensable component of maintaining interpersonal relationships (Thielmann & Hilbig, 2015). Trust and trustworthiness are frequently measured in the literature using the Trust Game (Berg et al., 1995). This game is played between two participants. In the first stage, the first player is given a sum of money and is asked to decide how much of it to transfer to the second player. The amount chosen by the first player is multiplied by a predetermined factor (typically three) before being transferred to the second player. The second player then decides how much of this money to send back to the first player. The amount transferred by the first player serves as a behavioral measure of the first player's trust in the second player, while the amount returned by the second player is considered a measure of the trustworthiness of the second player.

Assuming complete self-regard and information, the first player is not expected to transfer any money to the second player, since knowing that the second player has complete self-regard as well, the first player would not expect any money to be returned from the second player. However, in the original study where the game was conducted, 94% of the players transferred some amount of money to the other participant (Berg et al., 1995). On average, the first players transferred approximately 50% of their money to the second players. In return, the second players sent back roughly one-third of the tripled amount they had received, meaning the first players effectively recouped nearly the same amount they had initially transferred (Berg et al., 1995; Bolle, 1998). The amounts transferred in the Trust Game exhibit significant variability depending on various factors across studies. For instance, if the second player is also given a certain amount of money at the beginning of the game, this may reduce the first player's transferred amount, possibly by alleviating the perceived pressure to adhere to the principle of equality (Johnson & Mislin, 2011).

Cooperation

Cooperation plays a crucial role in the emergence and sustainability of complex social structures (Bowles & Gintis, 2011). Cooperative behavior refers to decisions made by individuals that prioritize long-term collective welfare over short-term personal gains. Both the causes and solutions to many large-scale contemporary issues, including global warming, involve trade-offs between the personal and the common good (Van Lange et al., 2013). In experimental economics, cooperative behavior is frequently measured using the Public Goods Game (Zelmer, 2003) and its variant, the Prisoner's Dilemma Game (Rapoport & Chammah, 1965). These games are typically based on a social dilemma that requires players to choose between personal interests and the welfare of a group.

In the one-shot version of the game, each participant in the group simultaneously and independently decides how much of their monetary endowment to contribute to a shared project account or common pool. The contributions to the pool are multiplied by a predetermined factor (usually two) before being redistributed equally among all players (usually four), posing a dilemma between individual and total group earnings, often referred to as a "social dilemma".

For example, in a one-shot four-player Public Goods Game with an endowment of 100 tokens for each participant, if three players contribute all of their tokens to the pool while the fourth player contributes nothing. The pool contributions of 300 tokens are doubled to become 600 tokens, divided equally among the four players. As a result, each of the contributing players would receive 150 tokens, while the self-regarding player would end up with more: 250 tokens. The amount a player contributes

to the pool is considered a behavioral indicator of how much they prioritize group welfare over self-interest, or in other words, how cooperative they are.

A completely self-interested individual in the one-shot version of the game would be expected to contribute nothing to the shared pool, as this person would see no dilemma in keeping all tokens. Contrary to this prediction, participants in this game often readily contribute more than half of their initial endowment to the public account (Isler et al., 2021a). The amount contributed is known to depend on several factors (for a meta-analysis, see Zelmer, 2003). For example, communication among players, especially face-to-face interaction, tends to increase cooperation (Balliet, 2010). Similar to the Dictator Game, anonymity between players and toward the experimenter tends to reduce contributions (Ledyard, 1995). In a field study, individuals were more likely to volunteer to reduce their energy consumption for collective benefit when their behavior was observable to others (Yoeli et al., 2013). The standard Public Goods Game is played as a one-shot game among anonymous participants. However, repeated versions of the game also exist. In these repeated games, cooperation tends to decline over time (Ledyard, 1995), unless freeriding can be punished (Fehr & Gächter, 2000).

The standard version of the Public Goods Game involves a "provision dilemma", where the social outcome is the provision of a non-existent public good. However, the game can also be presented as a "maintenance dilemma", where the social outcome is the maintenance of a preexisting common resource (Gächter et al., 2022; Isler et al., 2021a). Although the decision-outcome relationships in these two versions are mathematically identical, cooperation tends to be lower in the maintenance than in the provision dilemma. In the provision dilemma, cooperation is achieved through giving, while in the maintenance dilemma, cooperation is achieved through refraining from taking. Provision dilemmas involve an investment of resources toward a group project, whereas maintenance dilemmas require restraint in the consumption of resources that otherwise form the group project. Compared to the provision dilemma, the non-cooperative outcomes of the maintenance dilemma may better capture the social psychology that results in emptied store shelves during emergencies (e.g., the COVID-19 pandemic lockdowns). Research by Gächter et al. (2022) and Isler et al. (2021a) shows that these two types of public goods dilemmas can differ psychologically and behaviorally.

Another method for measuring cooperative behavior is the Prisoner's Dilemma Game. The Prisoner's Dilemma Game is a two-person version of the Public Goods Game. In the standard Prisoner's Dilemma (Rapoport & Chammah, 1965), participants are offered two choices: "to cooperate" or "to defect". In another version, cooperation is measured as a continuous variable. In both versions, players start with an equal amount of money and are asked to decide how much of it to transfer to their partner. The transferred amounts are then multiplied by a specific factor (usually two) before being received by the partner, increasing the overall value. Like the Public Goods Game, the Prisoner's Dilemma Game constitutes a social dilemma between individual and group welfare. For both players, transferring more money to their partner increases the pair's total earnings, while transferring less money maximizes individual gains.

As in other games, the results in the Prisoner's Dilemma Game often contradict predictions based on the assumption of complete self-regard. Nearly half of all players choose to cooperate with their partners in the standard Prisoner's Dilemma (Sally, 1995). This rate varies depending on several factors. For instance, communication between players, as in the Public Goods Game, is a significant factor in increasing cooperation in the Prisoner's Dilemma (Balliet, 2010; Sally, 1995; Spadaro et al., 2022). Pre-game discussions between partners enhance mutual trust, leading to greater cooperation (Loomis, 1959). Anonymity also impacts decision-making in the Prisoner's Dilemma, as in other games. For example, players tend to cooperate more when seated in positions where they can see each other (Sally, 1995). Additionally, the framing of the game can influence decisions. When the same game was presented to two groups with different names, such as "'Wall Street'" and "'Community,'" decisions in

the "'Wall Street'" version, which evokes associations with the stock market, tended to be more selfish (Liberman et al., 2004).

Norm Enforcement

The sustainability of human social relationships is closely tied to social norms and their enforcement (Fehr & Gächter, 2000). An essential mechanism in maintaining the continuity of prosocial behavior-regulating norms is the punishment of norm violations (Fehr & Fischbacher, 2004). For instance, failure to cooperate while benefiting from others' cooperation, also called "freeriding" may be punished with social exclusion from future cooperation (Tomasello et al., 2012; Van Lange et al., 2013). Norm enforcing punishment can be administered both by those directly affected by the norm violation, the "second parties", and by personally unaffected witnesses or "third parties."

Granting punishment opportunities (e.g., in response to Prisoner's Dilemma game decisions) to second- or third-parties allows experimental investigation of norm enforcement tendencies. For instance, Fehr and Fischbacher's (2004) study demonstrated that most third parties punished players who violated fairness and cooperation norms, but being a direct reaction to personal harm, second-party punishment tends to be more severe. The presence of a punisher in these games influences decision-making and reduces selfish behaviors (Balliet et al., 2011; Charness et al., 2008; Fehr & Gächter, 2002). The anticipation of punishment encourages norm adherence and fosters cooperative behavior, thereby strengthening social order and collective welfare.

Adding the possibility of retaliation (counter-punishment) into experimental settings allows researchers to investigate both retaliatory and other anti-social behavior. For example, in a Prisoner's Dilemma game, a second- or third-party dissatisfied with a player's selfish allocation may impose a punishment. However, any expectation of retaliation by this player might influence their willingness to punish. Selfish behaviors may prompt punishment, but the risk of retaliation by the punished participant might deter others from enacting punishment. Indeed, studies have shown that when the possibility of retaliation is introduced, punishment behavior decreases by up to 80% compared to games where retaliation is not possible (Balafoutas et al., 2014a). Interestingly, even when retaliation offers no direct benefit to the punished, about 35% of those punished engage in retaliatory punishment that is more severe than the penalty they received.

Punishment and counter-punishment have been studied in real-life settings, finding significant cross-cultural variation in the balance between these pro- and anti-social behaviors (Herrmann et al., 2008). In one such study conducted at a train station, individuals who violated norms by littering received less help from bystanders when they subsequently needed assistance, compared to a control group (Balafoutas et al., 2014b). The same study revealed that most observers refrained from punishing norm violations because they were concerned about the possibility of retaliation. Punishment when there is risk of retaliation or counter-punishment can be conceptualized as moral courage. Research has demonstrated that participants' motivation to punish is not primarily driven by monetary gain but by a desire to penalize non-prosocial behavior and to send a clear message to those who violate prosocial norms (Molnar et al., 2023).

Deontology vs. Utilitarianism

Philosophical approaches underlying moral decision-making play a crucial role in understanding human behavior. Deontology and utilitarianism represent two primary contrasting approaches in normative ethics and can significantly influence how individuals evaluate and make decisions. These approaches have sparked profound philosophical debates about how moral decisions are made, the principles they are based on, and how their outcomes are assessed.

Deontology typically evaluates moral actions based on ethical principles and rules. For instance, Kant's (1797) categorical imperative argues that actions must align with universally acceptable principles and emphasize adherence to these principles regardless of the consequences. However, the utilitarian approach focuses on the outcomes of actions, assessing their moral correctness by considering the benefits and harms they produce. This perspective evaluates whether an action maximizes overall well-being for the greatest number of people. Philosophers such as Bentham (1789) and Mill (1863) appraise actions according to the principle of "the greatest happiness for the greatest number." Allegiance to one versus the other approach can profoundly influence one's moral values, priorities, and decision-making processes.

The Trolley Dilemma, as introduced above, is a hypothetical scenario frequently used in studies on moral decision-making (Foot, 1967; Greene, 2007). The differences between deontological and utilitarian approaches are clearly demonstrated in this scenario, which involves sacrificing one person's life to save five others. From a deontological perspective, which upholds a rule prohibiting killing, ending a person's life is inherently immoral, regardless of the outcomes. In contrast, the utilitarian perspective argues that the morally correct choice is the one that seeks to achieve a result where more lives are saved. However, there are significant doubts about whether hypothetical scenarios like the Trolley Dilemma accurately reflect real-life decision-making processes (Sheeran & Webb, 2016). Numerous studies have shown that people's responses in such scenarios are weak predictors of their real-life decisions (e.g., Bostyn et al., 2018). In real life, a more complex array of factors typically influences decisions. Consequently, the use of hypothetical scenarios like the Trolley Dilemma may fall short of fully capturing deontological and utilitarian decision-making approaches. To address these shortcomings, more recent research has focused on developing more realistic measurements and scenarios to better understand decision-making processes.

Bostyn and colleagues (2019) used the Monetary Trolley Dilemma to observe deontological and utilitarian decisions seemingly incentivized with monetary consequences. They found that these behavioral measurements predicted decisions in the hypothetical Trolley Dilemma. Moreover, consistent with findings from previous studies (e.g., Aktaş et al., 2017), psychopathy was found to predict utilitarian decisions in the classic Trolley Dilemma. However, indicating a potential advantage of the Monetary Trolley Dilemma measure, this relationship was not observed in dilemmas involving monetary incentives, which may be related to the fact that the low monetary stakes in the Monetary Moral Dilemma do not involve life-and-death decisions.

Similarly inspired by Bostyn et al. (2019) but without employing experimental deception, Vurgun et al. (2023) further developed the Monetary Moral Dilemma, in particular, avoiding deception by having decisions actual monetary consequences. In Vurgun et al.'s version, a randomly chosen active participant and a passive participant (who won a random draw between six people) are each rewarded a $1 endowment. The active participant is then presented with an option that, if selected, takes away the $1 endowment from the lucky passive participant and instead gives $1 to each of the other five passive participants. This is the utilitarian option, maximizing the group's overall welfare, contrasted with the deontological option of not taking one's earnings. Vurgun et al. (2023) tested whether the Oxford Utilitarianism Scale (Kahane et al., 2018) predicts actual utilitarian behavior, with a time interval of nearly one and a half years between the two measures. The results showed that measuring utilitarian vs. deontological thinking with the Oxford Utilitarianism Scale did not predict utilitarian behavior in the Monetary Moral Dilemma. In contrast, previous measures of cognitive reflection, as well as post-decision measures of utilitarian justifications, were significantly correlated with utilitarian behavior in the task.

Points to Consider When Using Economic Games

The use of economic games in behavioral studies offers several advantages compared to hypothetical and self-report measures. However, various aspects of employing these techniques require careful attention by researchers to maximize their benefits. These include ensuring that participants are not deceived in experimental setups or at least identifying the best ways to minimize deception. Similarly, the choice of additional measures during and after these economic games can provide valuable insights and interactions. Therefore, carefully evaluating which additional measures to incorporate is crucial.

The experimental economic norm against the use of deception has permeated across almost all behavioral sciences, with the exception of psychology. Ortmann and Hertwig (2002) have persuasively argued, drawing on examples from psychology, that the use of deception in research can "damage" participant pools in the long term. They provide evidence that deception fosters distrust among participants over time and that this mistrust can influence their decision-making processes, demonstrating that the use of deception diminishes the credibility of studies conducted in psychology and economics. Thus, they argued that the norm against deception in economics should persist and extend to other fields.

Avoiding deception in the context of economic games implies that if participants are instructed that they will be matched with other players, that these matches will be made randomly, and that the money used in the games represents real money with tangible outcomes, it is essential that the study is conducted, and participants are rewarded exactly as described. Participants can be explicitly assured that there is no experimental deception involved in the study. However, maintaining such transparency and honesty can significantly increase research costs, as real pecuniary incentives need to be used. To avoid using deception while also managing study expenses, certain strategies can be helpful.

Studies involving multiple games can be prohibitively costly. A Random Lottery Incentive system, in which payments are made only for one or a few selected games within the study, can be used to reduce costs. Studies have shown that this type of lottery-based payment does not lead to significant differences in participants' decisions compared to traditional payments for every decision (Cubitt et al., 1998; Starmer & Sugden, 1991). Charness et al. (2016) compared the lottery system of paying for one random task with the conventional monetary incentives of paying for all tasks in a study and found them to be equally effective. In addition, a modified version of the Random Lottery Incentive system, the Conditional Information Lottery, allows incentivization of hypothetical scenarios that would otherwise be possible by using deception (Bardsley, 2000). Accordingly, participants are informed that some of the tasks may be hypothetical but that at least one or more is real and that the monetary rewards would be based on a randomly selected real task. By using either of these two approaches, researchers can reduce data collection costs while avoiding deception.

Exploratory measures taken during economic games can be used to identify moderating factors influencing decision-making processes. In particular, participants' comprehension of the economic game tasks can be assessed see (Goeschl &Lohse, 2018; Isler et al., 2018; Isler et al., 2021b). Differences in the comprehension of social dilemmas in economic games may shape decision-making processes significantly (Isler et al., 2021b; Rand et al., 2014). Similarly, participants' familiarity with these games may affect their decisions. For example, behavioral differences might be observed between naïve participants playing one-shot Public Goods or Prisoner's Dilemma games and experienced participants (Rand et al., 2014). Participants' familiarity with the task can be directly assessed using post-experimental survey questions (e.g., Rand et al., 2014).

Measures of social expectations in the context of the game provide additional insights into strategic games. For example, in the Public Goods Game, a robust strong predictor of cooperation behavior is

the participant's expectation of how much the other party will cooperate. The higher this expectation, the more cooperative the behavior (Isler et al., 2021a; Kudo et al., 2023; Weber et al., 2023).

Similarly, the internal conflict participants feel at the moment of decision-making could be measured as a source of potential insight into decision-making processes. For instance, in social dilemmas, the moral conflict experienced by an individual might lead to more deliberative decision-making, as opposed to automatic, intuition-based responses (Pennycook et al., 2015). The decision conflict or difficulty participants experience while making a decision can be measured with a single question, as demonstrated by Kocher et al. (2008) and Isler et al. (2021b).

Finally, the sense of closeness participants feel toward the other participant(s) can also significantly influence social behavior. This can be measured with a single-item Inclusion of Other in the Self scale, as in Gächter et al. (2015) or in Baader et al. (2024). These tasks ask participants to position two circles —one representing themselves and the other(s) representing the participant(s)—at varying degrees of proximity or overlap.

Surely, this is not an exhaustive list of relevant additional measures in the study of moral behavioral decisions. Incorporating these and additional measures can help identify the boundary conditions of decision-making processes in social and moral dilemmas, paving the way for future research.

Future Studies and Conclusion

Technological advancements provide new methods for measuring moral behaviors. For example, virtual reality (VR) technology has been used to present moral dilemmas, traditionally explored through hypothetical scenarios, in much more realistic contexts (e.g., Patil et al., 2014). This allows researchers to study moral decision-making processes under conditions that are ecologically closer to real-life scenarios. This capability is particularly valuable for presenting realistic contexts of moral dilemmas that are otherwise impractical to replicate for ethical reasons. Presenting the frequently used Trolley Dilemma to participants through VR has been shown to be more emotionally stimulating (Patil et al., 2014) and to result in more utilitarian decisions (Niforatos et al., 2020; Patil et al., 2014). This finding aligns with Bostyn et al.'s (2018) finding that real behaviors tend to be more utilitarian compared to responses measured using hypothetical scenarios. VR technology enables researchers to maintain experimental control while more effectively integrating contextual variables into experimental designs compared to traditional laboratory or field experiments (Innocenti, 2017). This method can also help evaluate the consistency between behavioral measurements taken in the laboratory and the field. For example, Fiedler et al. (2011) conducted a Trust Game in VR to examine the impact of social distance and virtual communication on participants' economic decisions and found these effects to be more pronounced than the traditional laboratory settings. Overall, VR technology can address one of social psychology's persistent challenges: measuring actual behavior in controlled settings.

Understanding the factors that predict human behavior and determining how and when these factors influence behavior are among the core objectives of social psychology. To achieve this goal, researchers should aim to observe actual behavior and not just intentions, much like during the golden age of social psychology, and prioritize the precision and accuracy of behavioral measurements. Such a shift will ensure that the field's scientific outputs continue to contribute meaningfully to our understanding of human behavior.

List of Basic Texts and Further Readings

Bowles, S., & Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science*, *29*(7), 1084-1093. https://doi.org/10.1177/0956797617752640

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106

Isler, O., Gächter, S., Maule, A. J., & Starmer, C. (2021a). Contextualised strong reciprocity explains selfless cooperation despite selfish intuitions and weak social heuristics. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-93412-4

Isler, O., Yilmaz, O., & Maule, A. J. (2021b). Religion, parochialism and intuitive cooperation. *Nature Human Behaviour, 5*(4), 512-521. https://doi.org/10.1038/s41562-020-01014-3

Rand, D. G. (2016). Cooperation, fast and slow: Meta-Analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science*, *27*(9), 1192-1206. https://doi.org/10.1177/0956797616654455

Saribay, S. A., & Yilmaz, O. (2024). *Reflection and intuition in a crisis-ridden world: Thinking hard or hardly thinking?*. Taylor & Francis.

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 https://doi.org/10.1111/spc3.12265

Acknowledgement

Author Details

Gülce Günaydın (PhD Student)

¹ Vrije Universiteit Amsterdam, Faculty of Behavioural and Movement Sciences, Department of Experimental and Applied Psychology, Amsterdam, The Netherlands

0000-0001-7748-3491   ✉ g.gunaydin@vu.nl

Ensar Acem (PhD Student)

² Kadir Has University, Faculty of Economics, Administrative and Social Sciences, Department of Psychology, İstanbul, Türkiye

0000-0002-7033-2911   ✉ ensar.acem@khas.edu.tr

Ozan Isler (Doctor)

³ The University of Queensland, School of Economics, Brisbane, Australia

0000-0002-4638-2230   ✉ o.isler@uq.edu.au

Onurcan Yılmaz (Associate Professor)

² Kadir Has University, Faculty of Economics, Administrative and Social Sciences, Department of Psychology, İstanbul, Türkiye

0000-0002-6094-7162   ✉ onurcan.yilmaz@khas.edu.tr

References

Aktaş, B., Yilmaz, O., & Bahçekapili, H. G. (2017). Moral pluralism on the trolley tracks: Different normative principles are used for different reasons in justifying moral judgments. *Judgment and Decision Making*, *12*(3), 297-307. https://doi.org/10.1017/S1930297500005891~

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men; research in human relations* (pp. 177-190). Carnegie Press.

Auger, P., & Devinney, T. M. (2007). Do what consumers say matter? The misalignment of preferences with unconstrained ethical intentions. *Journal of Business Ethics*, *76*(4), 361-383. https://doi.org/10.1007/s10551-006-9287-y

Baader, M., Starmer, C., Tufano, F., & Gächter, S. (2024). Introducing IOS11 as an extended interactive version of the "Inclusion of Other in the Self" scale to estimate relationship closeness. *Scientific Reports, 14*(1), 8901.

Balafoutas, L., Grechenig, K., & Nikiforakis, N. (2014a). Third-party punishment and counter-punishment in one-shot interactions. *Economics Letters*, *122*(2), 308-310. https://doi.org/10.1016/j.econlet.2013.11.028

Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2014b). Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences*, *111*(45), 15924-15927. https://doi.org/10.1073/pnas.1413170111

Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution, 54*(1), 39-57. https://doi.org/10.1177/0022002709352443~

Balliet, D., Mulder, L. B., & van Lange, P. A. M. (2011). Reward, punishment, and cooperation: a meta-analysis. *Psychological Bulletin, 137*(4), 594-615. https://doi.org/10.1037/A0023489

Bardsley, N. (2000). Control without deception: Individual behaviour in free-riding experiments revisited. *Experimental Economics, 3*(3), 215-240.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2010). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? In C. R. Agnew, D. E. Carlston, W. G. Graziano & J. R. Kelly (Ed.), *Then a miracle occurs: Focusing on behavior in social psychological theory and research* (pp. 12-27). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195377798.003.0002

Bearden, J. N. (2001). Ultimatum bargaining experiments: The state of the art. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.626183~

Bekkers, R. (2007). Measuring altruistic behavior in surveys: The all-or-nothing dictator game. *Survey Research Methods, 1*(3), 139-144. https://doi.org/10.18148/SRM/2007.V1I3.54~

Bentham, J. (1789). *An introduction to the principles of morals and legislation.* Clarendon Press, 1907.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122-142. https://doi.org/10.1006/GAME.1995.1027~

Bolle, F. (1998). Rewarding trust: An experimental study. *Theory and Decision, 45*(1), 83-98. https://doi.org/10.1023/A:1004930101743

Bolton, G. E., & Zwick, R. (1995). Anonymity versus punishment in ultimatum bargaining. *Games and Economic Behavior*, *10*(1), 95-121. https://doi.org/10.1006/GAME.1995.1026~

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science, 29*(7), 1084-1093. https://doi.org/10.1177/0956797617752640

Bostyn, D. H., Sevenhant, S., & Roets, A. (2019). Beyond physical harm: How preference for consequentialism and primary psychopathy relate to decisions on a monetary trolley dilemma. *Thinking & Reasoning, 25*(2), 192-206. https://doi.org/10.1080/13546783.2018.1497536

Bowles, S., & Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press. https://doi.org/10.23943/princeton/9780691151250.001.0001

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction.* Russell Sage Foundation.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour, 2*(9), Art. 9. https://doi.org/10.1038/s41562-018-0399-z

Camerer, C., & Mobbs, D. (2017). Differences in behavior and brain activity during hypothetical and real choices. *Trends in Cognitive Sciences, 21*(1), 46-56. https://doi.org/10.1016/j.tics.2016.11.001

Charness, G., Cobo-Reyes, R., & Jiménez, N. (2008). An investment game with third-party intervention. *Journal of Economic Behavior & Organization, 68*(1), 18-28. https://doi.org/10.1016/J.JEBO.2008.02.006

Charness, G., Gneezy, U., & Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization, 131*(Part A), 141-150. https://doi.org/10.1016/j.jebo.2016.08.010~

Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *American Economic Review, 92*(4), 1218-1221. https://doi.org/10.1257/00028280260344740

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive System. *Experimental Economics, 1*(2), 115-131. https://doi.org/10.1007/bf01669298

Doğruyol, B., Yilmaz, O., & Bahçekapili, H. G. (2021). Evolutionary by-product. In: Vonk J., Shackelford T. (eds) *Encyclopedia of Animal Cognition and Behavior*. Springer, Cham. https://doi.org/10.1007/978-3-319-47829-6_1800-1

Eckel, C. C., & Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior, 16*(2), 181-191. https://doi.org/10.1006/GAME.1996.0081

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

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63-87. https://doi.org/10.1016/S1090-5138(04)00005-4

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

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*(4), 980-994.

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T.(2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition, 123*(3), 434-441. https://doi.org/10.1016/J.COGNITION.2012.02.001

Fiedler, M., Haruvy, E., & Li, S. X. (2011). Social distance in a virtual world experiment. *Games and Economic Behavior, 72*(2), 400-426. https://doi.org/10.1016/j.geb.2010.09.004

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review, 5,* 5-15.

Frohlich, N., Oppenheimer, J., & Bernard Moore, J. (2001). Some doubts about measuring self-interest using dictator experiments: the costs of anonymity. *Journal of Economic Behavior & Organization, 46*(3), 271-290. https://doi.org/10.1016/S0167-2681(01)00178-0

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences, 7*(3), 385-400. https://doi.org/10.1016/0191-8869(86)90014-0

Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the closeness of relationships: A comprehensive evaluation of the 'inclusion of the other in the self' scale. *PLoS ONE, 10*(6), e0129478. https://doi.org/10.1371/journal.pone.0129478

Goeschl, T., & Lohse, J. (2018). Cooperation in public good games. calculated or confused? *European Economic Review, 107,* 185-203. https://doi.org/10.1016/j.euroecorev.2018.05.007

Goodwin, R., Graham, J., & Diekmann, K. A. (2020). Good intentions aren't good enough: Moral courage in opposing sexual harassment. *Journal of Experimental Social Psychology, 86,* Article 103894. https://doi.org/10.1016/j.jesp.2019.103894~

Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences, 11*(8), 322-323. https://doi.org/10.1016/j.tics.2007.06.004~

Güth, W., & Kocher, M. G. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization, 108,* 396-409. https://doi.org/10.1016/J.JEBO.2014.06.006~

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization, 3*(4), 367-388. https://doi.org/10.1016/0167-2681(82)90011-7~

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science, 319*(5868), 1362-1367. doi:10.1126/science.1153808

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

Innocenti, A. (2017). Virtual reality experiments in economics. *Journal of Behavioral and Experimental Economics*, *69*, 71-77. https://doi.org/10.1016/j.socec.2017.06.001

Isler, O., Maule, J., & Starmer, C. (2018). Is intuition really cooperative? Improved tests support the social heuristics hypothesis. *PLoS ONE, 13*(1), Article e0190560. https://doi.org/10.1371/journal.pone.0190560~

Isler, O., Gächter, S., Maule, A. J., & Starmer, C. (2021a). Contextualised strong reciprocity explains selfless cooperation despite selfish intuitions and weak social heuristics. *Scientific Reports, 11*(1), 13868. https://doi.org/10.1038/s41598-021-93412-4~

Isler, O., Yilmaz, O., & Maule, A. J. (2021b). Religion, parochialism and intuitive cooperation. *Nature Human Behaviour, 5*(4), 512-521. https://doi.org/10.1038/s41562-020-01014-3~

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

Kahane, G., Everett, J. A., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131-164.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *The American Economic Review, 76*(4), 728-741. http://www.jstor.org/stable/1806070

Kant, I. (1797). On a supposed right to lie from philanthropy. In M. J. Gregor (Ed.), *Practical philosophy* (pp. 611-615). Cambridge University Press.

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., & Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters, 101*(3), 175-178. https://doi.org/10.1016/j.econlet.2008.07.015~

Kudo, Y., Kato, T., Spadaro, G., & Balliet, D. (2023). Machine learning approaches for meta-analytic estimates of important predictors in behavioral science studies: An analysis of cooperation in social dilemmas. https://doi.org/10.31234/osf.io/pybxr

Ledyard, J. O. (1995). Public goods: A survey of experimental research. *The Handbook of Experimental Economics*, 111-194. https://doi.org/10.1515/9780691213255-004~

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

Loomis, J. L. (1959). Communication, the development of trust, and cooperative behavior. *Human Relations, 12*, 305-315. https://doi.org/10.1177/001872675901200402

Mieth, L., Buchner, A., & Bell, R. (2021). Moral labels increase cooperation and costly punishment in a Prisoner's Dilemma game with punishment option. *Scientific Reports*, *11*(1), 1-13. https://doi.org/10.1038/s41598-021-89675-6

Mill, J. S. (1863). *Utilitarianism*. Parker, Son and Bourn.

Milgram, S. (1963). Behavioral Study of obedience. *Journal of Abnormal and Social Psychology*, *67*(4), 371-378. https://doi.org/10.1037/H0040525

Molnar, A., Chaudhry, S. J., & Loewenstein, G. (2023). "'It's not about the money. It's about sending a message!'" Avengers want offenders to understand the reason for revenge. *Organizational Behavior and Human Decision Processes, 174,* Article 104207. https://doi.org/10.1016/j.obhdp.2022.104207~

Navarrete, C. D., McDonald, M. M., Mott, M. L., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional "'trolley problem.'" *Emotion*, *12*(2), 364-370. https://doi.org/10.1037/a0025561

Niforatos, E., Palma, A., Gluszny, R., Vourvopoulos, A., & Liarokapis, F. (2020). Would you do it?: Enacting moral dilemmas in virtual reality for understanding ethical decision-making. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-12. https://doi.org/10.1145/3313831.3376788

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231-259. https://doi.org/10.1037/0033-295X.84.3.231

Ortmann, A., & Hertwig, R. (2002). The costs of deception: Evidence from psychology. *Experimental Economics, 5*(2), 111-131. https://doi.org/10.1023/A:1020365204768

Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social Neuroscience, 9*(1), 94-107. https://doi.org/10.1080/17470919.2013.870091

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34-72. https://doi.org/10.1016/j.cogpsych.2015.05.001~

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications, 5*(1), 3677. https://doi.org/10.1038/ncomms4677

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation.* University of Michigan Press. https://doi.org/10.3998/mpub.20269

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7*(1), 58-92. https://doi.org/10.1177/1043463195007001004~

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*(2), 93-105. https://doi.org/10.1037/0003-066X.54.2.93

Sheeran, P. (2002). Intention—behavior relations: A conceptual and empirical review. *European Review of Social Psychology, 12*(1), 1-36. https://doi.org/10.1080/14792772143000003

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 https://doi.org/10.1111/spc3.12265

Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology (Columbia University), 187*, 60.

Spadaro, G., Graf, C., Jin, S., Arai, S., Inoue, Y., Lieberman, E., Rinderu, M. I., Yuan, M., Van Lissa, C. J., & Balliet, D. (2022). Cross-cultural variation in cooperation: A meta-analysis. *Journal of Personality and Social Psychology, 123*(5), 1024-1088. https://doi.org/10.1037/pspi0000389

Stoop, J. (2014). From the lab to the field: Envelopes, dictators and manners. *Experimental Economics, 17*(2), 304-313. https://doi.org/10.1007/s10683-013-9368-6

Suleiman, R. (1996). Expectations and fairness in a modified Ultimatum game. *Journal of Economic Psychology, 17*(5), 531-554. https://doi.org/10.1016/S0167-4870(96)00029-3

Thielmann, I., & Hilbig, B. E. (2015). Trust: An integrative review from a person-situation perspective. *Review of General Psychology, 19*(3), 249-277. https://doi.org/10.1037/GPR0000046

Thielmann, I., Böhm, R., Ott, M., & Hilbig, B. E. (2021). Economic games: An introduction and guide for research. *Collabra: Psychology, 7*(1). https://doi.org/10.1525/collabra.19004

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence Hypothesis. *Current Anthropology, 53*(6), 673-692. https://doi.org/10.1086/668207

Van Lange, P. A. M., Joireman, J., Parks, C. D., & van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes, 120*(2), 125-141. https://doi.org/10.1016/J.OBHDP.2012.11.003

Vavra, P., Chang, L. J., & Sanfey, A. G. (2018). Expectations in the ultimatum game: Distinct effects of mean and variance of expected offers. *Frontiers in Psychology, 9,* 992. https://doi.org/10.3389/fpsyg.2018.00992

Vurgun, R., Acem, E., Şeker, F., Yilmaz, O., Isler, O., & Dogruyol, B. (2023). Oxford Utilitarianism Scale's predictive power on the monetary trolley dilemma. https://osf.io/skjmh~

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~

Yilmaz, O., & Bahçekapılı, H. G. (2019). Sosyal psikolojinin altın çağını yeniden düşünmek. *Psikoloji Çalışmaları, 39*(1), 195-207. https://doi.org/10.26650/SP2018-0016~

Yilmaz, O. (Ed.). (2022). *Ahlakın yeni soyağacı: Psikolojik ve evrimsel bir bakış.* Koç University Press.

Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences of the United States of America, 110*(Suppl. 2), 10424-10429. https://doi.org/10.1073/pnas.1301210110

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics, 6*(3), 299-310. https://doi.org/10.1023/A:1026277420119