Article

Not peer-reviewed version

The Perception of Mistakes Driven by Cognitive Biases: An Experimental Study

Rom Severin *

Posted Date: 20 April 2023

doi: 10.20944/preprints202304.0643.v1

Keywords: cognitive biases; decision-making; experiment; heuristics

 doi:10.20944/preprints202304.0643.v1

Article

The Perception of Mistakes Driven by Cognitive Biases: An Experimental Study

Rom Severin

Eitan Berglas School of Economics, The Gershon H. Gordon Faculty of Social Sciences, Tel Aviv University
* Correspondence: romseverin@mail.tau.ac.il
This study was carried out under the supervision of Prof. Ayala Arad.

Abstract: This thesis presents findings from an online experiment aimed at evaluating individuals' attitudes towards their own cognitive biases that lead to objective mistakes. In a number of incentivized decision problems, a participant in the experiment might make a mistake and choose a dominated lottery (FOSD lottery), which is likely to result from a particular heuristic or cognitive bias. In this case, she is then confronted with her error under one of two treatments: a confrontation with an explanation only regarding the mistake or one with the addition of the supposed mechanism (bias). Following this explanation, her comfort level with her choice is measured. A bias that acted as a mechanism for the observed error could be considered advantageous (i.e., useful given alternative costs) or disadvantageous (i.e., mostly harmful) in other decision-making scenarios encountered during one's lifetime. Consequently, the perception of cognitive biases is subjective and may vary among decision makers, who may emphasize either the positive or negative aspects of such biases. The proposed methodology allows to determine whether these biases are perceived as useful rules of thumb, despite leading to a dominated lottery choice in that particular context, or unfavorable. Results revealed significant differences between the two treatments and across five different decision problems. Interestingly, the results showed that in only 60% of the cases, the participants felt uneasy about their mistakes. In the remaining instances, the participants demonstrated either indifference or a sense of comfort with their choice of a dominated lottery. The net effect of providing information on the cognitive biases that presumably lead to the mistakes, measured as the difference between treatments, revealed a negative perception of the mechanism behind the error in one problem, contentment with the mechanism for another, and no significant effect for the remaining biases.

Keywords: cognitive biases; decision-making; experiment; heuristics

1. Introduction

Classic economic models generally assume that people are rational. However, it has been well established that humans do not behave as entirely rational agents and exhibit pervasive and systematic biases in decision-making. This study aims to investigate the attitudes towards errors that arise from cognitive biases. To accomplish this, an online experiment was conducted to evaluate people's comfort level with respect to their choice of a dominated lottery. The comfort level was assessed based on a self-report. The choice between lotteries was incentivized and the expected payoff loss due to a mistake was similar in all the decision-problems. Nevertheless, the results revealed significant variations in the level of comfort with the mistakes across five different decision-problems. In each of these problems, a different cognitive bias likely led to the choice of the dominated lottery. Thus, a possible explanation for this variation is that the mechanism behind a mistake could influence an individual's level of comfort with the mistake.

The study also employed two distinct treatments: the *no-explanation* treatment involved informing participants that they selected the inferior lottery, while the *explanation* treatment provided participants with additional information about the cognitive bias that may have contributed to their erroneous choice. Comfort level was measured following a dominated lottery choice in each of the two treatments.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

Overall, a majority of the participants (59.6%) felt uncomfortable with their selection of the dominated lottery and a small proportion (12.25%) felt neutral. Notably, a considerable subset of the participants (28.15%) reported feeling comfortable with their choice, despite it being a mistake. In my analysis, I attempt to understand the differences in attitudes between the treatments and between the decision problems (i.e., how the different biases are perceived).

Classic economic theory deems the selection of a dominated lottery as irrational because it violates the FOSD principle, even though it could be argued that this decision is *procedurally rational* and thus deemed satisfactory by participants. The biases that served as mechanisms for mistakes in the experiment might be seen as either useful rules of thumb (heuristics), despite leading to a dominated lottery choice in that particular context, or unfavorable due to faulty reasoning (either in and of itself or due to it leading to an erroneous choice). This flawed reasoning may not be limited to the specific context of the experiment, but rather could extend to more common and critical real-life decision-making scenarios that carry significant material costs.

To discern the participants' attitudes towards the biases themselves, I analyze the net impact of providing information about the cognitive bias that presumably led to the mistake. I do so by examining the difference in comfort levels between the two treatments, for each decision-problem separately. This approach isolates the effect of the supplementary information regarding the mechanism from other factors that might influence level of comfort with the choice, such as the attitude towards the mere violation of FOSD and contextual factors. A negative net impact of the bias information might suggest that the bias was perceived as unfavorable, while a positive net impact could indicate satisfaction with the thought process despite the material outcome.

This methodology may have theoretical and practical implications. Theoretically, it evaluates whether attitudes towards some biases and heuristics align with modern definitions of rationality, which view these phenomena as somewhat normative. This perspective may have practical implications for the design of soft interventions aimed at improving decision-making, by informing the development of more effective "nudges" that exploit these mechanisms. By utilizing the insights gleaned from this experiment, it may be possible to select soft interventions that participants feel more comfortable with, thereby increasing the efficacy of the intervention. Moreover, the study may imply that bias correction procedures are not always desirable, particularly for biases that are perceived as satisfactory by participants.

The net impact of the bias information, i.e., the difference between the two treatments, was found to be significantly positive in one of the problems, negative in another and insignificant in the rest of the problems. This suggests that adding an explanation on the mechanism might not affect the attitude towards a dominated lottery choice, affect it positively or affect it negatively, depending on the underlying bias. These treatment effects are referred to as the "explanation premium" and are discussed in Section 5.1. Section 5.2. presents an analysis of the differences in comfort level across the five problems in each of the two treatments.

The introduction is structured as follows. I begin by defining biases as deviations from rationality and provide an overview of the literature on the five biases implemented in the experiment. I subsequently present a discourse on the normativity of biases, aiming to offer insights into the subjective nature of bias perception among individuals, which not necessarily conforms to an objective standard. Specifically, the discussion highlights the introspective nature of bias perception, particularly in the context of irrational decisions, such as selecting a dominated lottery. It proposes contemporary notions of rationality, which do not rely merely on utility maximization in a given context but rather endorse a more comprehensive approach to maximizing overall well-being or avoiding regret. These notions demonstrate that the way individuals perceive and respond to biases is often shaped by individual perspectives and experiences.

1.1. Rationality and Biases

Economic models generally assume that people are rational. This assumption was the foundation of whole research areas in the social sciences (see, e.g., the *Homo Economicus* theory in economics). Nevertheless, in light of evidence coming from experimental psychology and related

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

areas, it has been well established that humans do not behave as entirely rational agents but instead exhibit pervasive and predictable biases in decision making (Tversky & Kahneman, 1974). These findings shaped scholars' conception of people's statistical intuitions in economics, psychology, and beyond, concluding that *"intuitive judgments of all relevant marginal, conjunctive, and conditional probabilities are not likely to be coherent, that is, to satisfy the constraints of probability theory"* (Tversky & Kahneman, 1983, p. 313). Yet, it seems that humans (and other animals) have always relied on simple strategies or heuristics to solve adaptive problems, ignoring most information and forgoing much computation rather than aiming for as much as possible of both. This work examines people's attitudes towards some of these deviations from rationality: cognitive biases. In the context of this study, irrationality is demonstrated by choosing a first-order stochastically dominated lottery.

A bias is detected when deviation from a norm is observed. Thus, for some given tasks, a cognitive bias is easy to assert if a "correct" response is defined in light of a normative model describing how the situation should be processed. The more systematic such a normative deviation is found to be, the more legitimate the assertion that a bias is indeed present. Blanco (2017) suggests that "*Cognitive bias refers to a systematic (that is, nonrandom and, thus, predictable) deviation from rationality in judgment or decision-making*".

1.2. The Five Biases in the Experiment

Five cognitive biases[1] were carefully chosen as mechanisms to drive participants to choose a dominated lottery. The main reason those five were chosen was practical – the possibility to implement them in a design in which problems consist of a choice between two dice lotteries, where one lottery FOSD the other. The biases in question may result in participants having inaccurate beliefs about the chances of winning the lotteries, leading to their selection of a dominated option based on these subjective probabilities. In addition, while it is intuitive that a higher payoff loss due to a mistake induces more regret and reduces comfort, I aimed at fixing the expected loss due to an error and maintaining a similar decision context across the different problems. This makes it possible to explore the differential attitude towards different cognitive biases. Other reasons included the variability of statistical difficulty that is associated with each cognitive bias problem–ranging from very easy (*magical thinking*) to very hard (*base rate fallacy*)–that may affect the comfort level regarding mistakes. This also contributed to the efficiency of the design. Only observations in which a dominated lottery was chosen were analyzed as those were the instances in which a cognitive bias occurred. Thus, there was a need to challenge participants with varying levels of difficulty to make the most even out of those with a high cognitive ability. This design format, where participants are confronted with their mistakes could be applied to additional biases of course. Suggestions for variations on the design are presented in the discussion section.

Summarized below is the definition of each of the five biases, the mechanisms that account for the bias, the domains where this bias was exhibited, and the axioms found to be violated by each of the cognitive biases used in the experiment.[2] For a more detailed review see Appendix A.

Q1. Representative Conjunctions

In this experiment, the *representativeness heuristic*, a cognitive bias identified by Tversky and Kahneman (1982), is employed as a mechanism for the occurrence of *the conjunction fallacy*. This bias

---

[1] In reference to the five mechanisms that have been selected, it is worth noting that some are commonly referred to as "fallacies". These are defined as errors in reasoning that stem from a cognitive bias. Others are known as "heuristics", which represent a practical solution to problem-solving that is based on a cognitive bias. For the purposes of this discussion, all of these mechanisms shall be referred to as "biases" in a collective manner.

[2] Despite the diversity of axioms described below, in the experiment, all biases led to a violation of first-order stochastic dominance—a choice of a dominated lottery.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

occurs when people do not apply extensional reasoning to a task, but rather assess how similar an event is to an existing mental prototype. It manifests as an inclination to assume that specific conditions are more probable than a general one, thereby violating the conjunction law: $p(A \text{ and } B) \leq p(A)$. The conjunction fallacy has been observed in a variety of contexts, including the well-known Linda problem and its variations (Charness et al., 2010; Zizzo et al., 2000), as well as in other domains (Tversky & Kahneman, 1983), such as in the realm of politics (Kanwisher, 1989).

Q2. Magical Thinking and Greediness in Uncertain Events

Magical thinking refers to the belief that one's actions can influence the outcome of a chance event, even though the likelihood of the event is independent of those actions. This phenomenon is characterized by an expectancy of a personal success probability inappropriately different from what the objective probability would warrant and referred to as *an illusion of control* (Langer et al., 1975). This cognitive bias has been theorized to be the result of a belief in a just world, similar to the concept of karma (Furnham, 2003) or the desire to not tempt fate (Tversky & Kahneman, 1974; Tykocinski, 2008). Research has identified the presence of magical thinking in laboratory experiments (Arad, 2014; Tykocinski, 2008) and has also been incorporated into theoretical behavioral models (Daley & Sadowski, 2017; Passanisi et al., 2017).

Q3. Gambler's Fallacy

The belief that the occurrence of a particular outcome in random events will be balanced by a tendency for the opposite outcome is a cognitive fallacy that is rooted in the false intuition that events are not truly random and independent. Several mechanisms have been proposed to explain this fallacy, including the *representativeness heuristic* (Kahneman & Tversky, 1972), *The law of small numbers,* which posits that probabilistic properties of large samples are applicable to small samples (Tversky & Kahneman, 1971; Rabin, 2002), and the experience of *negative recency* (Ayton et al., 1989; Lopes & Oden, 1987). The fallacy has not only been exhibited in probability learning experiments (cited above), but has also been identified empirically in real gambling (Croson & Sundali, 2005; Suetens et al., 2016), sport (Misirlisoy & Haggard, 2014) and investments (Odean, 1998, Shapira & Venezia, 2001, Chen et al., 2007).

Q4. Probability Matching

Probability matching, a more extreme variant of the representativeness heuristic, concerns choices made by agents faced with competing alternatives. Under probability matching, the likelihood that an agent makes a choice amongst alternatives mirrors the probability associated with the outcome or reward of that choice (Vulkan, 2000) instead of consistently choosing the outcome with the highest probability (maximizing). This behavior represents a special case of *irrational diversification* (Rubinstein 2002). This form of heuristic decision making may also stem from the phenomenon of *attribute substitution*, which occurs when a fast, relatively effortless intuitive assessment is utilized (Kogler & Kuhberger, 2007; West & Stanovich, 2003). Vulkan's (2000) study provides a comprehensive review of experiments, including various versions of the light guessing experiment, that demonstrate the presence of probability matching in laboratory settings. This phenomenon has also been observed using pigeons as participants (Herrnstein, 1958).

Q5. Base Rate Fallacy

The base rate fallacy refers to the tendency to neglect background or "base rate" information in situations where it is highly relevant to the decision-making process. This bias results in inaccurate probability judgments in various contexts. The fallacy arises when base rate information that does not fit into a causal schema is overshadowed by causally relevant data (Tversky & Kahneman 1980, p. 50). Bar-Hillel (1980) proposed a similar mechanism in which base rate information is perceived as irrelevant, while Nisbett et al. (1976) suggested that base rate data is often viewed as remote and abstract. This fallacy has been demonstrated in laboratory experiments, such as the well-known cab

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

problem (Kahneman & Tversky 1972) and the lawyer and engineer item (Kahneman & Tversky, 1973). The base rate fallacy can also be manifested in the interpretation of data, such as in medical tests for rare diseases. An example of this can be found in Appendix A.

1.3. The Normativity of Biases

The normative models that define optimal behavior depend on the definitions of the "rationality" concept. Mainstream economics seeks mathematical models to study how agents should behave to advance their interests. Consequently, the conservative implication of the assumption of rationality pertains to the optimization of choice in accordance with a preference relation, subject to a limited number of conditions regarding the arrangement of these preferences (e.g., monotonicity with respect to money). On the other hand, behavioral economics seeks to study how people *actually* behave (instead of how they should) and exploits the use of surveys and experiments. These experiments have identified suboptimal behavioral – participants that deviate from the classic definitions of rationality.

Simon (1957) introduced the "principle of bounded rationality": "*The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world or even for a reasonable approximation to such objective rationality".*

It has been followed by more recent modifications of rationality that seek to strike a balance between the perspectives of classic and modern economics, acknowledging the limitations of human decision-making while still providing a framework for theoretical models. These new definitions take into account actual human behavior, rather than relying solely on idealized assumptions of rationality. The most salient to this work are the concepts of *satisficing, procedural rationality, rule-rationality, ecological-rationality*, and a definition based on the lack of regret.

Table 1 displays the definitions of different concepts of rationality and how they are related to the experimental design.

In addition to the modern definitions of rationality that aim to capture the ideal standards of human thinking and decision-making, some researchers have attempted to account for cognitive biases, which often cause individuals to deviate from rational behavior. Korteling et al. (2020) define cognitive biases as *"systematic cognitive dispositions or inclinations in human thinking and reasoning that often do not comply with the tenets of logic, probability reasoning, and plausibility. These intuitive and subconscious tendencies are at the basis of human judgment, decision making, and the resulting behavior"*.

Several approaches have tried to account for cognitive biases.

- <u>Limited cognitive resources</u>: An obvious explanation for many reported cognitive biases is the limited processing capacity of the human. As in most complex problems, the optimal, truly rational solution is out of reach. Some argue that we can only aim at "bounded rationality" (Kahneman, 2003)–choosing the best decision after considering only a limited amount of information.
- <u>Influence of motivation and emotion</u>: Emotions are biologically relevant to decision making because they affect behavior (Bechara & Damasio, 2005). For instance, the *loss-aversion bias* (Kahneman & Tversky, 1984) consists of the preference for avoiding losses over acquiring gains of equivalent value, and it could be driven by the asymmetry in the affective value of the two types of outcomes. Other examples concern moral judgment, such as the famous "trolley dilemmas" (Bleske-Rechek et al., 2010).
- <u>Social influence</u>: Social cues could produce certain cognitive biases. One example is *the bandwagon bias*, which describes the tendency of people to conform to the opinions expressed earlier by others and strongly influences collective behaviors, such as voting in elections (Obermaier et al., 2015).
- <u>Heuristics:</u> I believe the most successful attempt to provide a coherent framework to understand cognitive biases is Kahneman and Tversky's research program on heuristics (Kahneman et al., 1974). The rationale of this approach states that making rational choices is not always feasible or desirable. This is because it takes time to collect and weigh all the evidence to solve a problem;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

it needs the investment of lots of cognitive resources that could otherwise be used for other purposes and quite often a rough approximation to the best solution is "good enough". Consequently, the mind uses mental shortcuts and simple rules to arrive at a conclusion in a fast-and-frugal way.

Experimental research has documented several heuristics that could underlie many cognitive biases. Common examples include the representativeness heuristic, the availability heuristic, and the anchoring and adjustment heuristic (Gilovich et al., 2002).

The cognitive biases often discussed in behavioral economics can be classified into two main categories. Some biases are interpreted as *judgment mistakes*. For example, participants calculate subjective probabilities that differ from true ones because of "magical thinking", "framing", or some "anchor". On the other hand, some irrationalities are interpreted as *unique preferences*, ones that violate some normative norm. For example, the compromise effect in which a product gains attractiveness simply by virtue of becoming a middle option in a choice set. This rule is based on consumers' inferences about their personal valuation of alternatives from the portfolio of market offerings and some information about their own relative tastes. In the current experiment, I will focus on five judgmental mistakes in lottery probability assessments. Biases that are based on unique preferences (rather than judgmental mistakes) could potentially be examined under a similar methodology to the one presented in this work, and are addressed in the discussion section.

Table 1. Definitions of rationality.

Concept of rationality	Definition	Relevance to Experiment
Satisficing (Simon, 1947)	People will satisfice when they make a decision that satisfies and suffices for the purpose.	Simon's "satisficing man" operates with rationality which is practically feasible in his environment, as opposed to "rational man" that maximizes utility. I see this definition as the main consequence of bounded rationality. If participants are satisficing, they might feel comfortable exhibiting an erroneous choice when the correct choice requires a tool or method that is out of reach (say, a difficult probabilistic calculation that requires the use of Bayes rule).
Procedural rationality (Simon, 1976)	A behavior is procedurally rational when it is the outcome of appropriate deliberation or some strategy of reasoning, compared to irrational behavior, an outcome of impulsive responses without adequate intervention of thought.	The definition focuses on the rationality of the processes used in arriving at a decision, as opposed to the rationality of the decision itself. Participants might feel comfortable with their decision if they see their thought process as procedurally rational when it aligns with the suggested bias.
Rule-rationality (Aumann, 2019)	Under rule-rationality, people do not maximize over acts (as in ordinary rationality); instead, they adopt rules of behavior that do well in usual, naturally occurring situations. When deciding, they choose an act that accords with the adopted rule. Usually, these	The cognitive biases, heuristics, and fallacies that lead to errors in the experiment may be seen as unconscious rules adopted by the participants. Making participants aware of these rules may make their attitude towards making mistakes more acceptable, since these rules were developed through evolution to optimize behavior in typical situations. However, these

doi:10.20944/preprints202304.0643.v1

	acts maximize utility in the situation at hand, but only sometimes.	rules may result in suboptimal outcomes in atypical or artificial situations like a lottery choice experiment because evolutionary pressures do not apply in these contexts.
Ecological-rationality (Smith, 2003)	According to ecological rationality, behavior arises from cultural and biological evolutionary processes and so is optimal in the environment. It differs from Rule-rationality as it contrasts Constructivist rationality (based on deductive processes of human reason), while rule-rationality contrasts act-rationality.	Similarly to rule-rationality, as people follow the rules without being able to articulate them, the awareness of these rules might affect their perception of them. In addition, the attention this concept received when Vernon L. Smith won the 2002 Prize in Economic Sciences in Memory of Alfred Nobel indicates the growing interest in investigating the positive aspects of what traditional doctrine portrays as irrational behavior.
Subjective regret[3] (Gilboa & Schmeidler, 2001)	A mode of behavior is *irrational* for a decision maker if, when exposed to the analysis of her behavior, the decision maker feels embarrassed or wishes to change her choices, and so forth.	This definition is subjective: a mode of behavior (say, a choice of a dominant lottery) might seem rational for some decision makers and not others. The qualitative nature of this concept also suggests variability in the subjective perception of an erroneous choice. For example, less intelligent decision makers may fail to understand the analysis of their choices and may consequently not exhibit any regret or embarrassment. As a result, they may appear (and internally, see themselves as) more rational than intelligent decision makers who make the same decisions but can understand why these decisions are not coherent.

2. Contribution

Despite the fact that detecting cognitive biases and their robustness has been a key area of focus, the examination of attitudes towards these biases has not been a priority among scholars. I am aware of only one paper, Nielsen and Rehbeck (2022) (which is discussed below), that assessed whether participants are reluctant to violate some behavioral axioms or not. In this study, I focus on the comfort level with regard to a violation of one specific axiom, FOSD of lotteries, and examine how it is influenced by *the reason* for violating the axiom. This evaluation is noteworthy as it sheds light on the perspectives of real participants, rather than relying on the views of researchers, who may have their own preconceptions regarding biases. Additionally, several theories suggest that there may be variations in individual attitudes towards cognitive biases. For example, focusing on either negative or positive accounts and the diffusion between cognitive biases and heuristics suggests that the views toward cognitive biases differ from one subject to another. However, these attitudes may converge towards a common norm, as evidenced by the results in the *explanation* treatment.

As Kahneman and Tversky (1982) point out, judgmental errors (and cognitive biases) might have both positive accounts (focusing on the factors that produced an incorrect choice) and negative accounts (explaining why the correct choice was not made). This suggests that the sense of a mistake differs depending on whether a positive or negative reflection has been made. Another potential reason for variation in the attitude towards biases is the framing of a mechanism that led to some

[3]Lacking a formal notation by the authors, this is my own name for this definition of rationality.

mistake as a heuristic or cognitive bias and the confusion between these two expressions might indicate that the perceptions of these mechanisms are subjective. While using heuristics can lead to cognitive biases, not all cognitive biases result from heuristics. In addition, Besharov (2004) examines second-best considerations in correcting cognitive biases and suggests, through a theoretical model, that correcting any single one might have ambiguous welfare implications in a system of interacting biases.

Some practical contribution is related to the recent proposals for "decisional enhancement", that makes use of "nudges", soft interventions that modify the environment in which choices are made to help individuals make better decisions. These interventions are based on cognitive biases and heuristics. Some work has been done to test the public's attitude toward these soft interventions (Arad & Rubinstein, 2018, Felsen et al., 2013). The findings of this experiment might suggest that different biases in the basis of a nudge intervention are more or less acceptable.

In addition, some papers have suggested bias corrective procedures designed to combat biases in individual and group judgments. For example, they improve elicitation procedures (Cleaves, 1987). This literature might benefit from an assessment of which biases are more tolerable than others as it might suggest participants' willingness to correct them or not. Cognitive bias correction has not only been the realm of scholars but has also been adapted to practical psychology as a form of treatment. Cognitive bias modification (CBM) is an experimental paradigm that uses training to induce maladaptive or adaptive cognitive biases. See (Jones & Sharpe, 2017; Hallion & Ruscio, 2011) for a meta-analysis.

Nielsen and Rehbeck (2022) explored a somewhat similar question to mine in a laboratory experiment. An enormous amount of experimental literature has shown that individuals consistently violate canonical axioms that were the direct mathematical interpretations of the conservative definition of rationality in decision theory. Using an incentivized laboratory experiment, they identify whether decision makers consider it a mistake to violate canonical choice axioms. The findings suggest that most individuals want to follow canonical axioms, shown by their desire to revise their lottery of choice when it conflicts with some axiom they initially wanted to satisfy. This line of research is to be encouraged, in my opinion, as it sheds light on the possible gap between our assumptions of rational agents and their own definitions of "rationality". The following experiment is the beginning of an attempt to check whether some cognitive biases are viewed as normative, similar to Nielsen's findings for classical theory.[4]

Another related paper is Gigerenzer's (2018) "The bias bias in behavioral economics" which challenges the normativity of cognitive biases. Gigerenzer highlights the tendency of scholars to identify biases even when they are absent, for instance, by overlooking differences between small and large sample statistics, or by mistaking people's random errors for systematic errors. The paper also critiques aspects of the *irrationality argument*, which provides reasons for governmental paternalism due to apparent shortcomings in people's rationality, stubbornness, and the substantial welfare-relevant costs that biases may incur. In my opinion, the paper is an attempt by the scholar to present biases as a more normative behavior, and this experiment can support this notion through examining participants' attitudes towards certain biases. As Gigerenzer writes, *"Getting rid of the bias bias is a precondition for psychology to play a positive role in economics".*

3. Experimental Design

3.1. Procedure

An online experiment was conducted to assess participants' attitudes towards mistakes that originate from different cognitive biases. Participants were mainly students from Tel Aviv University and The Academic College of Tel Aviv. They were recruited using e-mail messages and were compensated according to the winnings in one of the lotteries they chose which was randomly

[4] I wouldn't rush to write "mistake" as a synonym for "a choice of a dominated lottery" in my experiment, if it wasn't for the results of that experiment that identified FOSD as a mistake.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

selected at the end of the experiment. There was no fixed participation fee. Participants were paid 12.5 NIS on average. The median time it took participants to complete the experiment was 7.5 minutes. After excluding the 5% of participants with the fastest response time and the ones who tried to participate twice, 397 respondents were left, 84% of them study at public universities (rather than colleges), and 38% are males.

3.2. Design

Participants faced five decision problems (Q1-Q5), in each of which they chose between two lotteries. Participants might make a mistake in each problem and choose a dominated lottery (FOSD lottery). The expected monetary cost of an erroneous choice, that is, the difference in the expected prize, was the same across all problems (5 NIS ~1.4$). See Figure 1 for a flow chart of the general layout and the description of the questions in Appendix B.

In each of the five decision problems, such a mistake is likely to result from a different cognitive bias or heuristic: representativeness heuristic, magical thinking, gambler's fallacy, probability matching, and base rate fallacy. For a review of the biases, see Appendix A.

Following each problem screen, participants were asked to explain why they chose that particular lottery. Then, regardless of whether the participant chose the dominating or the dominated lottery, he was asked whether he was willing to share his choice along with his first name on the following screen. The problem and lottery of choice were displayed at the top of the page on each of these screens, until the next problem. Only if participants chose the dominated lottery (made a mistake), were they then confronted with their mistake (following an explanation of their choice). There were two treatments: a confrontation with an explanation of the potential mechanism and without. Participants were assigned randomly to one of these treatments, which was consistent throughout the questionnaire.

In the *no-explanation* treatment (base treatment), participants were presented with a description explaining that they had made a mistake (a choice of a dominated lottery). In the *explanation* treatment, another paragraph was shown, explaining the mechanism that might have led them to this choice, a cognitive bias or heuristic. This paragraph noted that there is evidence in the literature for that phenomenon, along with its name, a general description, and how it might serve as a mechanism for an erroneous choice in that particular problem.

For example, in the gambler's fallacy problem, a dice with 6 faces–4 blue (B) faces, and two orange (O) faces, was rolled four times and the sequence (from left to right) was: OBBB.

Participants were asked "What color would you choose to bet on for the fifth roll?"

Participants who chose orange (O) were confronted with their mistake. In this screen, in the *no-explanation* treatment, they were only presented with this paragraph:

"In each roll of the dice, the probability of the roll result being B is 4/6 and the probability of the roll result being O is 2/6.
This is true regardless of the number of times we have thrown the dice or the result of previous throws since the dice does not remember the color that it landed on in the past.
In other words, the previous roll of the dice does not affect the next roll.
Since the previous sequence is irrelevant, you should bet on B as the chance of B coming out is higher."

Participants who were assigned to the *explanation* treatment saw the paragraph above as well as this additional paragraph:

"There is research evidence for the "gambler's fallacy", a misconception that if something happens relatively often relative to the expected frequency in a given period, then in the future it will happen less frequently (and vice versa) so that a balance is created. In this case, the gambler's fallacy can lead to the thought that the chance of B drawn again is lower than the chance of O drawn because of the sequence of previous throws."

At the bottom of this page, in both treatments, participants were then presented with the question, *"How do you feel about your choice after reading this passage?"* and were asked to mark a point

 doi:10.20944/preprints202304.0643.v1

on a fixed 7-point scale, ranging from -3 (*"I feel completely uncomfortable with my choice"*) to 3 (*"I feel completely comfortable with my choice"*), where 0 is neutral. This would be referred to as the "comfort level", the main outcome variable of the study.

After this confrontation, participants were asked if they were willing to share their lottery of choice, stating their name, with other participants in the experiment. No choices were actually shared, but this willingness to share might capture feelings of shame from having made a mistake. Subsequently, the participants moved on to the next problem, with the order of problems randomly assigned for each participant.

After answering five decision problems, one of the chosen lotteries in these problems was randomly selected. A few days later, this lottery was conducted and participants were notified via a text message in which lottery they had participated and how much they had won. Any winnings were then sent using the Israeli "BIT" mobile payments app.

See Figure 1 for a general layout of the experiment and Appendix B for a full English translation of all the screens (which were originally displayed in Hebrew).

3.3. Hypotheses

My main interest was to examine whether the comfort level differs with regard to the two treatments and the different mechanisms that led participants to a mistake. The main two hypotheses, pre-resisted at "ASPREDICTED" on 03/06/2022,[5] are:

Hypothesis 1. Some cognitive biases and heuristics have a positive "explanation premium" while others do not: when participants are only confronted with their mistake of choosing a dominated lottery, their comfort level with their choice is different than when they are confronted with the mistake and the potential heuristic/bias that led to the mistake.

Hypothesis 2. Decision makers' (dis)comfort level with choosing a dominated lottery depends on the decision problem, and in particular on the mechanism behind the mistake in that problem (what led them to make a mistake – different cognitive biases and heuristics in the different problems).

The examination of comfort level was conducted utilizing two outcome variables. The main outcome variable was the subjective comfort level, which was expressed on a numerical scale. The secondary measure was a binary choice indicating willingness to share the lottery of choice along with their first name.

The choice of a point on the comfort level scale was not incentivized and the measure might be noisy, but there is no reason for the comfort level across the five problems and across two treatments to be biased in different ways. Even though comfort level is a subjective measure and participants are not incentivized to share their "true and honest" feelings, the problems themselves are incentivized on the grounds that at the end of the survey, one of the lotteries selected by the participant is chosen and he gets to participate in it. Assuming that participants want to maximize their expected payout, they should strive to avoid choosing a dominated lottery. Participants' willingness to share is "socially incentivized", as they may feel ashamed of choosing a dominated lottery.

[5] Available at https://aspredicted.org/8vk7f.pdf

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

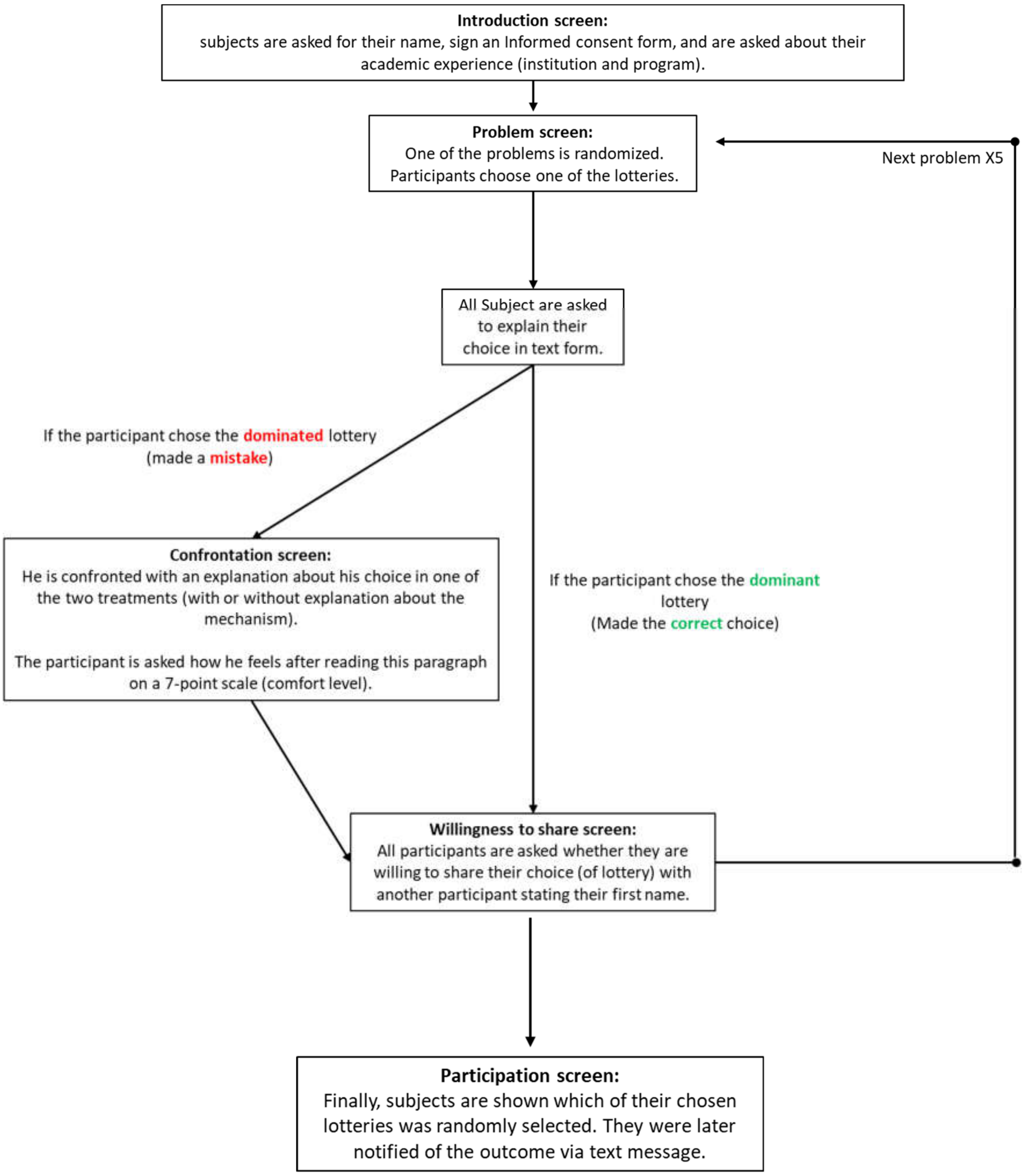


Figure 1. General layout of the experiment.

When comparing the comfort level across the five problems, it was crucial to disentangle extraneous factors that may impact the comfort level from the main factor of interest – the mechanism behind the mistake. Such extraneous factors include the violated axiom, the context of the problem, and the materialistic cost of the mistake. Therefore, the mistake made was designed to be as similar as possible for each problem. First, all mistakes result in a choice of a dominant lottery. FOSD is just one of the many axioms we assume a rational choice maker will satisfy. It was necessary, of course, to have participants violate the same axiom in all problems because, as shown by Nielsen and Rehbeck (2022), there are some axioms that a decision-maker will be more or less willing to violate. Second, for the mistake to be consistent in monetary terms, the choice of a dominated lottery resulted in a difference of 5 NIS in the expected payout of each problem. Third, all problems were framed in the same context as a dice lottery. In this fashion, a mistake always consists of a choice of a dice lottery that pays 5 NIS less on average than the other, so a comparison between these mistakes is reasonable.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

4. Descriptive Statistics

Most of the participants were students from Tel Aviv University, a small proportion were from The Academic College of Tel Aviv-Yafo, and a few from other academic institutions. Overall, 84% of the participants were students at public universities (rather than colleges), and 38% were males. Participants were recruited using e-mail messages and were offered no participation fee, only the lottery winnings.

4.1. Mistakes

Aggregating over all participants and all problems, 545 dominated lotteries were chosen out of a total of 1,985 problems that participants faced (i.e., 27.4% of the choices were mistakes). Figure 2 presents the share of participants who made 0, 1, 2, 3, 4 or 5 mistakes. Most of the participants made at least one mistake, and one mistake was the mode.

Figure 3 shows the share of mistakes made in each problem. The base rate fallacy problem (Q5) was the most difficult for participants according to the error rate. In particular, 47% of the participants selected the dominated lottery from the two options presented. This result was expected, as the base rate fallacy was added to serve as an almost impossible problem that requires advanced calculations to solve correctly.[6] Probability matching was the second hardest, followed by representative conjunctions and finally magical thinking and the gambler's fallacy problems, which had a somewhat similar number of mistakes. The majority of respondents were females (62%), who had 0.24 more mistakes than males, on average. Figure 4 shows the share of mistakes in a particular problem, out of all mistakes made by males (in grey) and females (in blue). Even though the panel is unbalanced, comparing the stacked groups allows us to observe which gender had more mistakes in a particular problem (while the aggregation of the two stacked bars has no direct meaning).

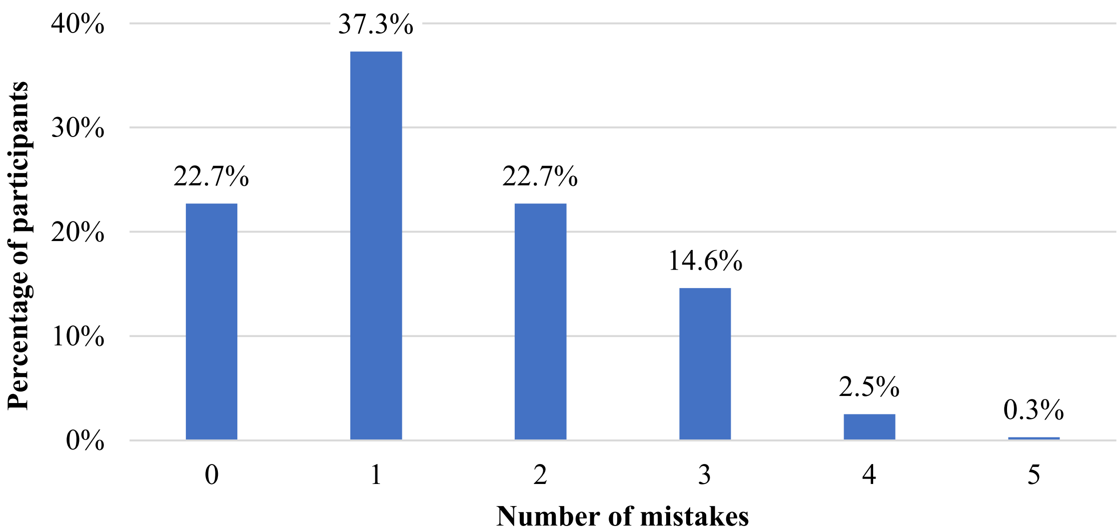


Figure 2. Share of participants with a certain number of mistakes.

[6] Even though 53% of participants selected the correct lottery out of the two options available, their success rate was not much higher than that of flipping a coin.

 doi:10.20944/preprints202304.0643.v1

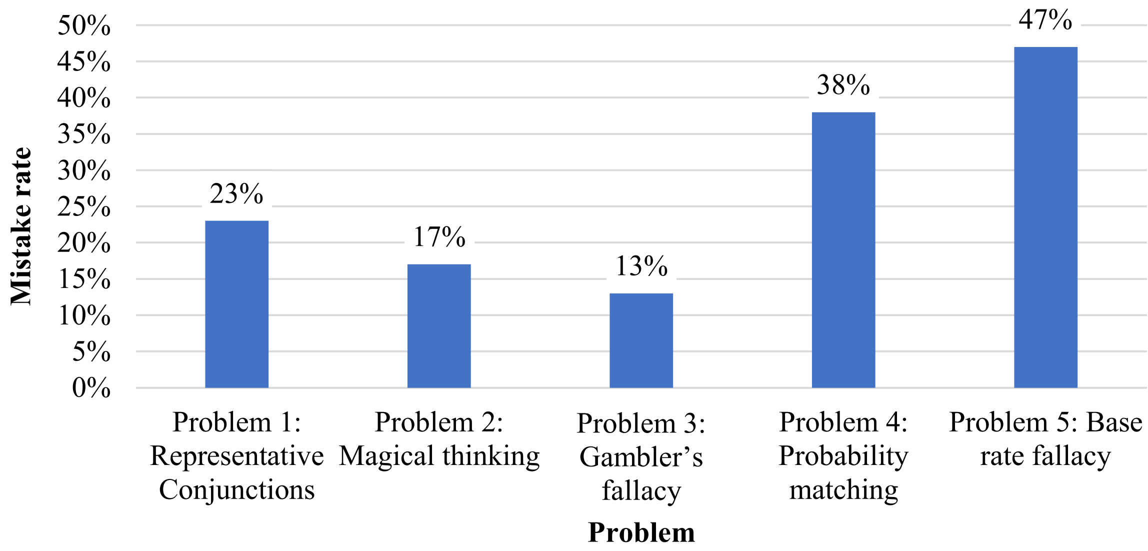


Figure 3. Mistake rate for each problem.

Men and women had a similar share of mistakes across the problems, while problems 1 and 5 were more difficult for males and the others were more difficult for females.

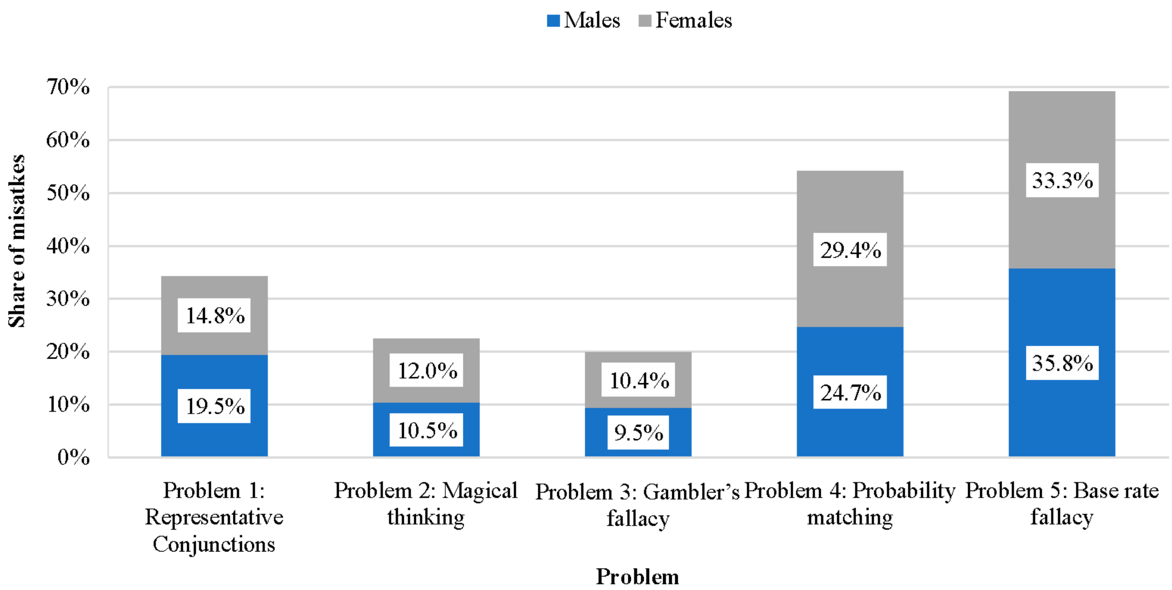


Figure 4. Share of mistakes for each problem out of all mistakes made by gender group.

4.2. Response time

Figure 5 shows the median response time for each of the five problems, clustered by correct choices (in grey) and incorrect (in blue). Participants spent most of their time on the base rate fallacy problem (Q5), which was the hardest. Note that the gambler's fallacy (Q3) had two preceding screens before the actual problem and lotteries were displayed. The first screen described the decision-making scenario, and the second included an animation. Hence, participants had time to contemplate before making a choice. Therefore, the response time indicated here, measured on the third screen of problem 3, does not fully express the overall time spent. Taking into account the total screen time for the three screens, the median total screen time for participants who chose the dominated lottery was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

46.9 seconds, and 55.2 for the others. It is clear that a shorter response time is associated with mistakes. See Section 5.3.2. for an analysis.

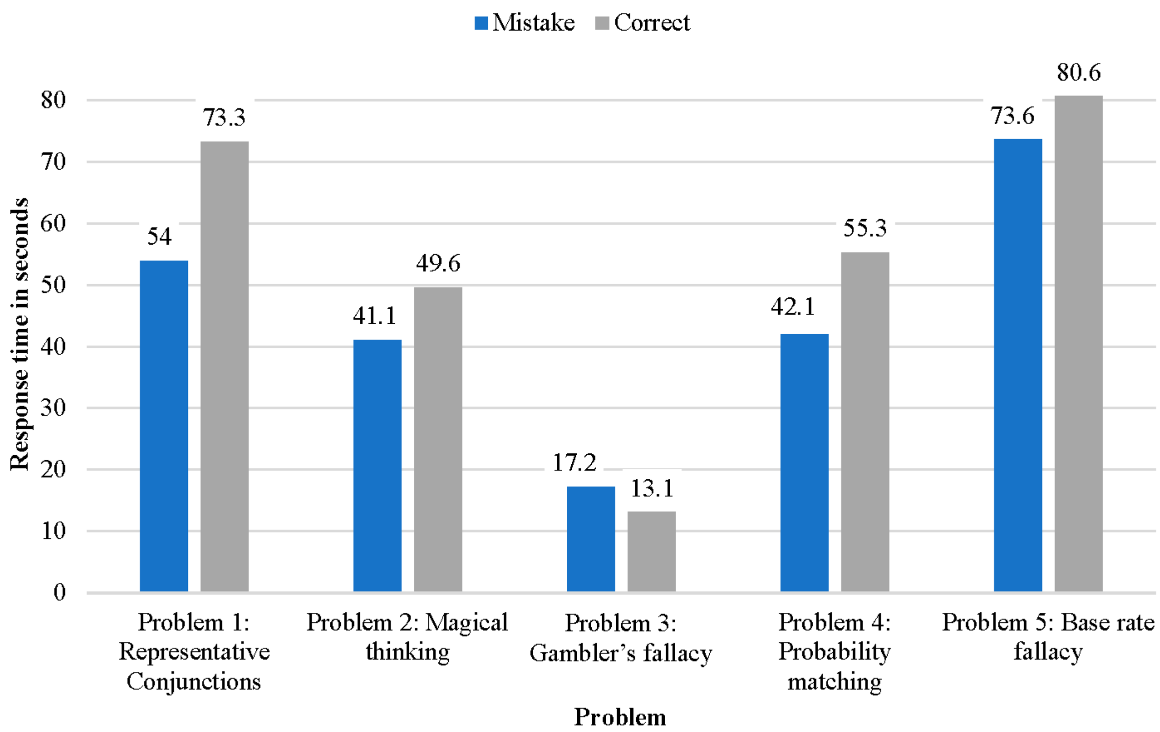


Figure 5. Median response time for a correct or incorrect choice in each problem.

4.3. Explanations

After each choice, whether correct or not, participants were asked to explain their choice in text form. Explanations across all problems consisted of 8.6 words on average.

Sometimes, justifications for dominated lotteries (mistakes) matched the hypothesized mechanism. These would be referred to as *matching explanations*. Table 2 presents some examples of explanations that participants gave to justify their mistakes that suggest the use of the heuristic or cognitive bias, shown in the *explanation* treatment for that problem.

Figure 6 shows the share of explanations classified as matching for each problem. On average, 44% of participants who chose a dominated lottery gave an explanation classified as a matching explanation. Note that there were probably even more cases where the proposed mechanism matched the reason behind the choice, even though it was not expressed in the explanation. Instead, they just wrote "Because it's better", "I guessed" or "Better chances" etc.

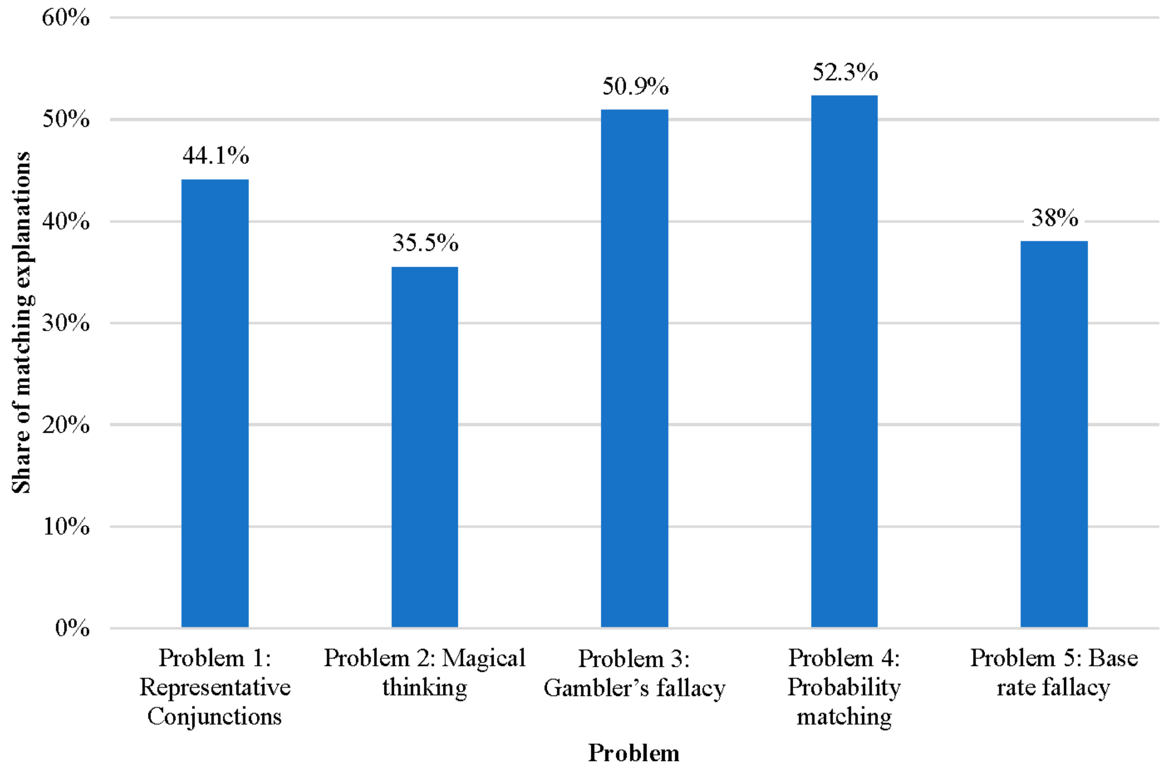


Figure 6. Share of matching explanations in each problem.

Table 2. Examples of matching explanations for each problem.

Problem	Examples of Matching Explanations
Q1. Representatives	Participant 124: *"The dice has more red faces, and my lottery of choice has more red faces"* Participant 365: *"The three last rolls are the same, but, in the lottery I chose, there is a higher chance to get R at the first roll"*
Q2. Magical thinking	Participant 379: *"My lucky number"* Participant 48: *"I settled for the lower prize as I do not want to push it"*
Q3. The gambler's fallacy	Participant 48: *"The chances of O are larger, as the other color showed up a lot"* Participant 113: *"Diversification"*
Q4. Probability matching	Participant 125: *"On one hand, there is a higher chance that each throw will result in black, on the other hand, 2/6 black rolls is also not a very small chance, therefore, it seems likely that out of the six throws, white will come up"* Participant 252: *"The chance for only Bs is low"*
Q5. Base rate fallacy	Participant 354: *"The chances to win 18 NIS are 70% and the chances to win 15 NIS are 75%. I prefer getting 5% more chance instead of 3 NIS of winnings"* Participant 150: *"A greater chance to win something"*

4.4. Comfort Level

The primary explained variable of the experiment was participants' comfort level after being confronted with their mistake in either the *explanation* or *no-explanation* treatment.

Figure 7 displays the cumulative distribution of comfort level, which ranges from -3 to +3, aggregated for the five problems and both treatments. The results demonstrate that a significant proportion of participants feel somewhat comfortable with their dominated lottery. The mean comfort level was -1, indicating a subtle feeling of discomfort that falls on the negative side of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

scale. Additionally, 58.6% of the time, participants felt either neutral or comfortable with their mistake, with 7.8% choosing the highest point on the scale.

Figure 8 presents the mean comfort level for each of the five problems, under the two treatments, along with 5% error bars.

All problems had a negative mean comfort level under the explanation treatment. This means that, on average, participants marked their feelings as less comfortable than "neutral" after reading the paragraph with the explanation.

Notice that for some problems, the comfort level for the explanation treatment was higher than that of the no-explanation treatment: representative conjunctions (Q1), probability matching (Q4), and base rate fallacy (Q5). The other two problems yielded the opposite result. Furthermore, the magical thinking (Q2) and gambler's fallacy (Q3) problems are the only ones with a positive mean comfort level under the no-explanation treatment.

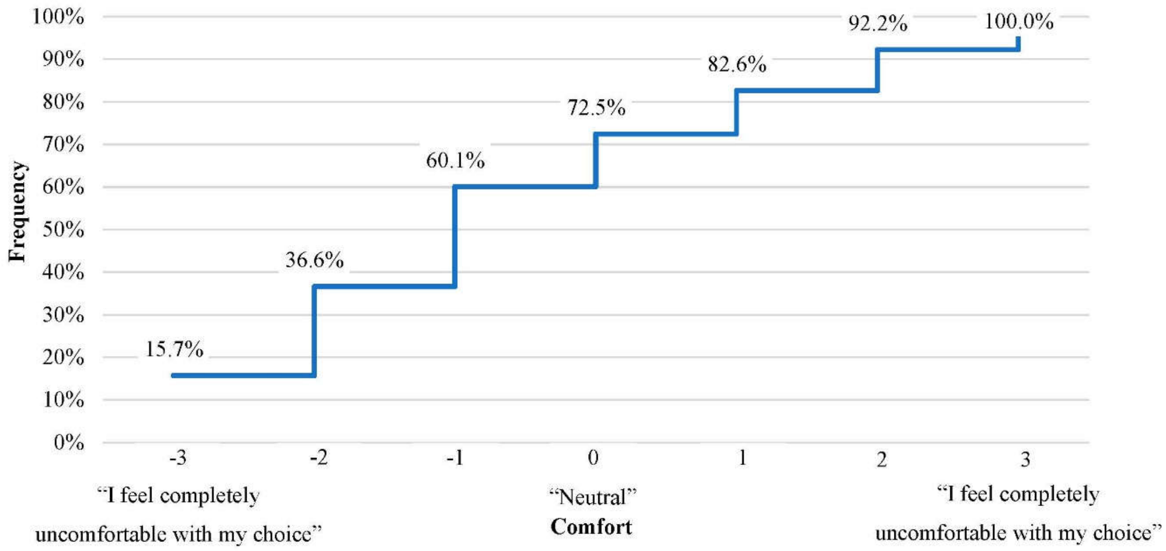


Figure 7. Cumulative precent of comfort levels.

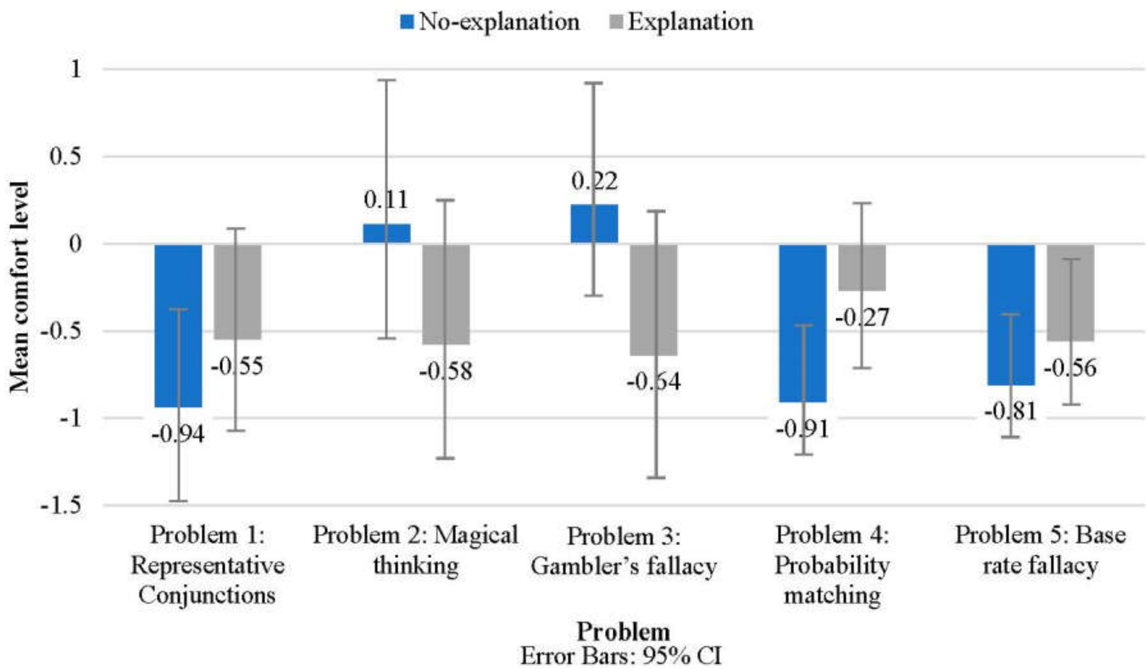


Figure 8. Mean comfort level in each problem and treatment.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

4.5. Willingness to Share

Figure 9 display for each of the five problems, the percentage of participants who agreed to share their lottery of choice, along with their first name, with a different participant. This is shown for each of the three possible scenarios: a correct choice of a dominating lottery (left column), and an incorrect choice of a dominated lottery in either the *explanation* (right column) or the *no-explanation* treatment (middle column).

The results across all problems indicate that a higher proportion of participants are willing to share their decision when they have made a correct choice, regardless of the explanation provided, lending validity to the notion that the variable in question represents a measure of confidence or shame. For an analysis of willingness to share correct or incorrect choices, see Section 5.3.3. in additional results.

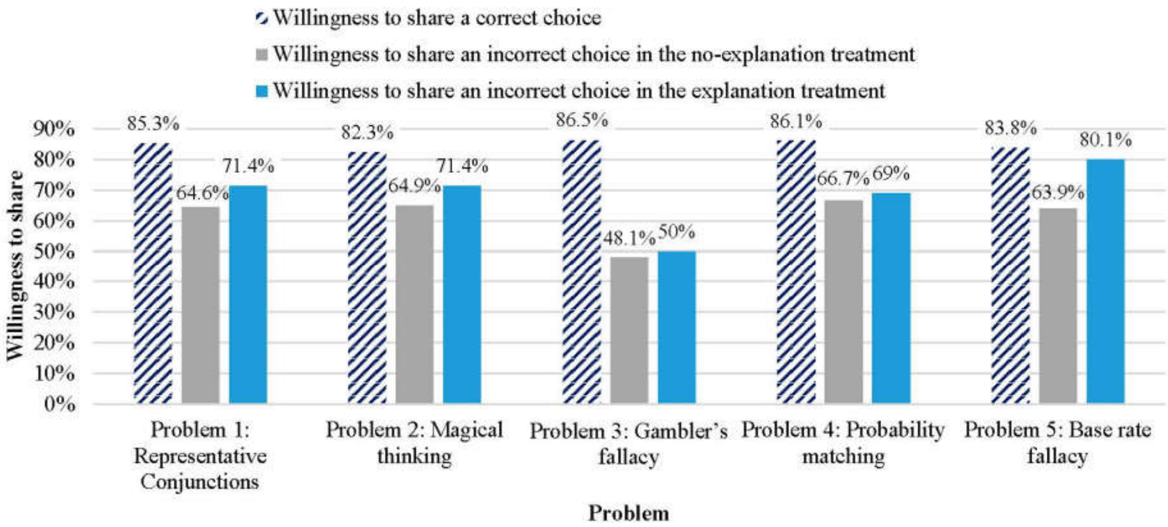


Figure 9. Willingness to share in each problem by treatment.

5. Results

5.1. Differences between Treatments (within Problems)

5.1.1. Explanation Premium for Comfort Level

As seen in Figure 8, there are differences in the comfort level of the two treatments, as hypothesized in Hypothesis 1. These differences will be referred to as the "explanation premium" – an effect on participants' feelings of a mistake that originates from an addition of an explanation of the proposed mechanism.

Two problems were found to have a significant explanation premium, that is, a significant difference in comfort levels of the two treatments. A positive explanation premium was found for the probability matching problem (Q4), that is, adding an explanation about what led to the mistake lowers the discomfort participants feel regarding their mistake. An opposite effect was identified for (Q3), the gambler's fallacy problem, which was found to have a negative explanation premium. That is, adding an explanation about the bias makes participants less comfortable with their mistake. These findings imply that the effect of supplementary information about cognitive biases on participants' attitudes towards their mistakes varies depending on the nature of the underlying bias. In particular, the results indicate that participants perceive probability matching as a beneficial heuristic, while they regard the gambler's fallacy as a flawed thought process.

The comfort levels of participants who chose the dominated lottery in the probability matching problem (Q4) in the *no-explanation* treatment (M = -0.913, SD = 1.667) were significantly lower than those in the *explanation* treatment (M = -0.267, SD = 1.881), according to a *Mann–Whitney U test (U = 2318.5, n1 = 81, n2 = 71, P = 0.037 two-tailed).*

This problem was the second hardest for participants—only 38% chose the dominant lottery. The mere fact that there is researched evidence for some phenomenon that might lead to a mistake might make participants more comfortable with their choice. Another factor is the specific mechanism displayed in the *explanation* treatment – probability matching. It might affect the comfort level, especially when the description indeed matches the actual thinking process that led a participant to a mistake. In other words, if the explanation provided aligns with how they arrived at the incorrect decision, they may be more accepting of the explanation and less likely to feel uncomfortable about the mistake. The classification of matching explanations can aid in investigating this.

In the representative conjunctions problem, Q1, a more subtle case of matching probabilities, the average comfort level in the *explanation* treatment was 0.39 higher compared to the *no-explanation* treatment, though it was not found to have a significant explanation premium.

There was a significant difference in the comfort level of participants who chose the dominated lottery in the gambler's fallacy problem (Q3) under the *no-explanation* treatment (M = 0.222, SD = 1.476) and the *explanation* treatment (M = -0.642, SD = 2.003), according to a *Mann–Whitney U test (U = 265, n1 =27, n2 = 28, P = .053 two-tailed).*

One reason that the explanation makes people feel less comfortable here could be the fallacy's name, which might have a negative association. Gambling, a robust addiction, can be frowned upon by most people, but more so in the context of the questionnaire, which requires proper probabilistic calculations to maximize the payout (as compared to gambling, or guessing). Participants might think that the questioner classified them as "gamblers" whether they actually are ones or not. Indeed, "gambling fallacies" have long been thought to have an etiological role in both the onset of gambling and the development of problem gambling (Delfabbro, 2004; Delfabbro & Winefield, 2000; Gaboury & Ladouceur, 1989).

It is also possible that without the explanation, some participants did not agree entirely that they truly made a mistake. that is, they did not think the previous rolls do not affect the future ones. If this is the case, the average comfort level for the *no-explanation* treatment also reflects the feelings of participants who think they did not make a mistake. It is reasonable to assume, that the share of participants who chose a dominated lottery but did not acknowledge it as a mistake (despite the explanation) is smaller in the *explanation* treatment. This is because giving the fallacy a name and researched evidence might convince participants to agree that they actually made a (common) mistake.

Maybe not surprisingly, the magical thinking problem (Q2) also had a higher average comfort level for the *no-explanation* treatment, although the difference between treatments was not statistically significant. The terms magical thinking and the gambler's fallacy might have a negative association compared to the other biases which are more neutral in context (i.e., probability matching, representative conjunctions, and base rate fallacy).

5.1.2. Explanation Premium for Willingness to Share

Another way to assess in which contexts participants felt more comfortable with making a mistake is to examine how many were willing to share their dominated lottery of choice, along with their first name, with a different participant. This is because sharing a mistake could lead to feelings of shame. The null Hypothesis 1, regarding the secondary variable was that there are no differences in the proportion of participants willing to share their choice of a dominated lottery under the *explanation* and *no-explanation* treatment.

For Problem 2, magical thinking, there is a significant negative explanation premium in "sharing": while 85% agreed to share their lottery in the *explanation* treatment, only 65% agreed to do so in the *no-explanation* treatment χ^2 *(1, N = 63) = 3.304, p = 0.082.*

Magical thinking was the easiest problem; a choice of the dominated lottery could arise because of either lack of attention to the question at hand, misunderstanding the question, or a belief that a choice of the lottery affects the objective probabilities of winning. The last-mentioned fits the magical thinking phenomenon.

One group of participants who chose the dominated lottery did so for the first two reasons- those that do not fit the magical thinking mechanism. For this group, participants who got the *explanation* treatment might feel less comfortable (than without an explanation), as now they understand that they belong to a group that exhibits magical thinking. This suggestion, of belonging to a group that believes unrelated events are causally connected despite the absence of any plausible causal link between them might make participants uncomfortable. For atheists, even the simple suggestion that they "believe…" might be unsatisfying.

The other group, which included participants who actually chose a dominated lottery because of some form of magical thinking, might have been relieved to see the addition of the explanation, first, because of the acknowledgment of the mechanism, as a known phenomenon. Moreover, in sharing their choice, they might have found pride in not being greedy.[7] If this is the case, the group that fits the magical thinking mechanism should have been more willing to share than the other group. Unfortunately, there need to be more observations (mistakes in the magical thinking problem (Q2)) to identify differences between the matching and non-matching explanations division with regard to willingness to share.

Despite this limitation, the results indicate that participants who gave matching explanations expressed a higher comfort level with their chosen dominated lottery than those who provided non-matching explanations. In particular, 13 out of the 36 participants who chose a dominated lottery in Q2 in the *explanation* treatment gave a matching explanation, while the remaining 23 gave a non-matching explanation. The group with matching explanations reported a higher comfort level (mean = 1.69, S.D. = 1.601) compared to the group with non-matching explanations (mean = -0.78, S.D. = 0.422). This difference is statistically significant (*Mann–Whitney U = 54, n1 =23, n2 = 13, P = 0.001 two-tailed).*

In the gambler's fallacy (Q4), the willingness to share is not affected by the addition of an explanation. However, the positive effect of matching explanations on comfort level found in the magical thinking problem (Q2) are also reflected in the increased willingness of participants to share their choice of a dominated lottery in the probability matching problem when the explanation is aligned. In the probability matching problem (Q4), under the *explanation* treatment, about half of the participants gave a matching explanation, and 79% of them agreed to share their choice. In comparison, only 58% of the participants who did not give a matching explanation were willing to share their choice. This difference is statistically significant at the 10% level (χ^2 *(1, N = 71) = 3.772, p = .052).*

5.2. Differences between Problems

5.2.1. Differences in Comfort Levels

Significant differences between comfort level of different problems were identified for the *no-explanation* treatment, but not for the *explanation* treatment. This suggests that the simple fact that participants get some justification for their mistakes makes them feel differently about their mistakes. However, on aggregate—this comfort level diverges into some common comfort level, irrespective of the specific mechanism.

A Kruskal–Wallis test by ranks was adopted to test differences in comfort levels between all pairs of problems under the *no-explanation* treatment. Results suggest that at least for one of the pairs, the comfort levels originated from a different distribution *(H (4) = 15.661, P = 0.004).*

To follow this result, all 11 pairs of problems were assessed, to check for differences in comfort levels. Following are all the pairs that had a significantly different comfort level in the *no-explanation* treatment:

[7] Note that participants didn't really know what sharing really means. They might have assumed that their choice is shared along with an explanation regarding the mechanism if they were in the explanation condition.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

Representative Conjunctions (Q1) Comfort < Gambler's Fallacy (Q3) Comfort

A difference in the comfort level of problem Q1 (mean = -0.94, S.D = 1.873) and Q3 (Mean = 0.22, S.D = 1.476) is significant, according to a *Mann–Whitney U test (U = 389.5, n1 =48, n2 = 27, P = 0.004 two-tailed).*

Probability Matching (Q4) Comfort < Magical Thinking (Q2) Comfort

A difference in the comfort level of problem Q2 (mean = 0.11, S.D = 2.183) and Q4 (Mean = -0.91, S.D = 1.667) is significant, according to a *Mann–Whitney U test (U = 1,099.5, n1 =37, n2 = 81, P = 0.019 two-tailed).*

Base Rate Fallacy (Q5) Comfort < Magical Thinking (Q2)

A difference in the comfort level of problem Q2 (mean = 0.11, S.D = 2.183) and Q5 (Mean = -0.81, S.D = 1.689) is significant, according to a *Mann–Whitney U test (U = 1,312.5, n1 =37, n2 = 93, P = 0.032 two-tailed).*

Probability Matching (Q4) < Gambler's Fallacy (Q3) Comfort

A difference in the comfort level of problem Q3 (Mean = 0.22, S.D = 1.476) and Q4 (Mean = -0.91, S.D = 1.667) is significant, according to a *Mann–Whitney U test (U = 660, n1 =27, n2 = 81, P = 0.032 two-tailed).*

Base Rate Fallacy (Q5) Comfort < Gambler's Fallacy (Q3)

A difference in the comfort level of problem Q3 (Mean = 0.22, S.D = 1.476) and Q5 (Mean = -0.81, S.D = 1.689) is significant, according to a *Mann–Whitney U test (U = 799, n1 =27, n2 = 93, P = 0.003 two-tailed).*

These differences are "problem effects" that do not necessarily derive from the proposed mechanism for the mistake (bias) but from the problems' difficulty, length, expected payout, etc. Although I tried to mitigate problem effects as much as possible by setting a constant 5 NIS difference in expected payouts and phrasing all problems as a choice between two dice lotteries, such differences across the problem cannot be eliminated.

One potential reason that participants felt more comfortable with a dominated lottery choice in the magical thinking problem compared to two other problems is the pre-choice recognition of the axiom violation. It is likely that participants who chose the dominated option in this problem (e.g., their lucky number) were aware of its inferior (objective) expected payout. The gambler's fallacy problem had higher comfort levels compared to the other three problems, potentially due to its animated presentation being more engaging and captivating compared to the other problems that didn't include any animations.

5.2.2. Differences in Willingness to Share

This section presents results regarding differences in willingness to share different problems, in each of the treatments. There are three significant differences in the willingness to share under the *explanation* treatment, and one in the *no-explanation* treatment.

Explanation Treatment

Representative Conjunctions (Q1) Willingness to Share > Gambler's Fallacy (Q3) Willingness to Share

The proportion of participants in the *explanation* treatment who agreed to share their choice of a dominated lottery in Q1 (71%) differed from the proportion in Q3 (50%), χ^2 *(1, N = 70) = 3.304, p = 0.069.*

Magical thinking (Q2) willingness to share > gambler's fallacy (Q3) willingness to share

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

The proportion of participants in the *explanation* treatment who agreed to share their choice of a dominated lottery in Q2 (85%) differed from the proportion in Q3 (50%), χ^2 *(1, N = 54) = 7.269, P = 0.007.*

Probability matching (Q4) willingness to share > gambler's fallacy (Q3) willingness to share

The proportion of participants in the *explanation* treatment, who agreed to share their choice of a dominated lottery in Q4 (69%) differed from the proportion in Q3 (50%), χ^2 *(1, N = 99) = 3.137, P = 0.077.*

No-Explanation Treatment

Probability matching (Q4) willingness to share > gambler's fallacy (Q3) willingness to share

The proportion of participants in the *no = explanation* treatment who agreed to share their choice of a dominated lottery in Q4 (67%) differed from the proportion in Q3 (48%), χ^2 *(1, N = 108) = 3.137, P = 0.086.*

The findings show that individuals are more inclined to reveal their dominated choice in the probability matching problem (Q4) than in the gambler's fallacy problem (Q3) across both treatments. This contradicts the previous comparison of comfort levels between the two problems in the absence of an explanation, where the gambler's fallacy problem (Q5), which was less frequently disclosed, was perceived to be more comfortable.

5.3. Additional Results

5.3.1. Number of Mistakes and Comfort Level

A regression of number of mistakes on comfort level shows that the more mistakes one makes, the more comfortable one feels with choosing a dominated lottery. This trend is observed in the set that includes only the *explanation* treatment ($\hat{\beta} = 0.345, SE = 0.119$), but not in the set that includes exclusively the *no-explanation* treatment. It also holds in the set that includes both treatments, without a dummy for treatment ($\hat{\beta} = 0.189, SE = 0.81$). See Table A1 in Appendix C for the regression table.

5.3.2. Response Time and Mistakes

All five problems[8] have a faster median response time for those who made a mistake. A statistically significant difference in response time was detected for two of these problems, representative conjunctions (Q1) and probability matching (Q4).

A significant difference in response time of the Problem 1 decision screen was observed between participants who made a mistake (M = 54.083 seconds, SD = 48.926) and did not make a mistake (M = 73.388, SD = 4.138); *t(395) = 2.371, P = 0.018.*

This result is supported in a Mann–Whitney nonparametric test:

Mann–Whitney U = 10401, n1 =307, n2 = 90, P = 0.001 two-tailed.

A significant difference in response time of the Problem 4 decision screen was observed between participants who made a mistake (M = 42.143 seconds, SD = 30.915) and did not make a mistake (M = 55.319, SD = 53.546); *t(395) = 2.761, P = .006.*

This result is supported in a Mann–Whitney nonparametric test:

Mann–Whitney U = 14846, n1 =245, n2 = 152, P < 0.001 two-tailed.

5.3.3. Willingness to Share Correct or Incorrect Choices

Participants seem to be more willing to share their lottery of choice, along with their first name, when they do not make a mistake, that is, choose the dominating lottery. Table 3 presents a significant

[8] Taking into account the total screen time for the three screens of the gambler's fallacy (Q3).

difference in the share of participants who were willing to share their correct choice of a dominant lottery in each of the five problems.

The most negligible effect was in the base rate fallacy problem (Q5). In addition, willingness to share is significantly negatively correlated with a mistake in Problems 1-4. The average correlation is 0.285, and correlations are significant at the 0.01 level (2-tailed), but not in Problem 5. This is probably because participants feel more comfortable with a mistake in the base rate fallacy, as it was the most challenging problem according to the error rate.

Table 3. Willingness to share correct and incorrect choices.

Problem	Share of Participants Willing to Share a Correct Choice	Share of Participants Willing to Share an Incorrect Choice	Statistics
Q1. Representative conjunctions	85%	68%	χ^2 (1, N = 397) = 14.158, $p < 0.001$
Q2. Magical thinking	85%	62%	χ^2 (1, N = 397) = 2.974, $p = 0.085$
Q3. Gambler's fallacy	86%	49%	χ^2 (1, N = 397) = 43.838, $p < 0.001$
Q4. Probability matching	86%	68%	χ^2 (1, N = 397) = 19.121, $p < .001$
Q5. Base rate fallacy	84%	77%	χ^2 (1, N = 397) = 2.929, $p = 0.087$

5.4. Results Summary

Out of all the decision-making scenarios, 27.4% of the responses were erroneous choices of dominated lotteries. In these instances, participants were made aware of their mistakes with either *no-explanation* regarding the mechanism or with one, and then their comfort level was measured.

The findings suggest that the effect of adding an explanation on the mechanism might not be consistent and can either affect the attitude positively, negatively, or have no impact, depending on the underlying bias. Specifically, a significant explanation premium, that is, a difference in comfort levels of the *explanation* and the *no-explanation* treatment, was found for two of the problems. In the probability matching problem (Q4), adding an explanation about what led to the mistake lowered the discomfort participants felt regarding their mistakes. This suggests that participants viewed the bias as a potentially useful heuristic or rule of thumb, despite its negative impact in that particular context. An opposite effect was identified for the gambler's fallacy, suggesting that participants perceived the bias as unfavorable.

In the *explanation* treatment, no significant difference between comfort levels of different problems could be identified, while there were prominent differences in the *no-explanation* treatment (attributed to the "problem effects"). This might suggest that when participants get some justification for their mistakes, they each feel differently about their own mistakes. However, on aggregate, this comfort level diverges into some common comfort level, irrespective of the specific mechanism.

6. Discussion

My experiment is pioneering in a line of experiments that examine the normativity of biases. It offers a novel design: an attempt to combine incentives to elicit true behavior- a decision between two lotteries, along with unincentivized scale measures- comfort level, and socially-incentivized measures- willingness to share the mistake.

There are many exciting directions to take this line of thinking, and I outline three below. One key area for further research is the examination of more common and realistic mistakes that go beyond violations of the first-order stochastic dominance (FOSD). By exploring other canonical axioms such as *independence of irrelevant alternatives*, *transitivity*, *independence*, or *consistency*, a broader

 doi:10.20944/preprints202304.0643.v1

range of biases, heuristics and unique preferences can be studied, including those that do not necessarily result in a dominated lottery and are not limited to decision making under uncertainty. The *sunk cost fallacy* for example, is a particularly intriguing area of study, as it often leads to serious real-life losses across various fields (see Roth et al., 2015, for a meta-analytic review). These losses are more tangible than a choice of a dominated lottery in a curated lab experiment. Examining this phenomenon using a similar methodology as in the current study, even without comparisons to other mechanisms, would provide valuable insights. I assume people are aware of the sunk cost fallacy even without formal explanations. It is worth exploring whether such investments are categorized as mistakes by participants, and how such assessments compare to those made for similarly unfavorable investments that didn't originate from sunk costs, but rather from alternative sources.

Secondly, a potential avenue for future research is to explore alternative methodologies for assessing mistakes that do not rely solely on self-evaluation by the individual who committed the error. Utilizing external evaluators to assess participants' criticism of mistakes made by others could enable an examination of whether attitudes towards a specific bias converge towards a certain mean at the aggregate level, as observed in the current study. Moreover, this approach could facilitate an investigation into whether attitudes obtained through self-reflection differ from those obtained through external evaluation. Additionally, this approach could investigate alternative explanatory variables for assessing attitudes towards biases, such as the use of story stimuli. This method involves presenting participants with a description of a person displaying a particular cognitive bias and asking them to evaluate whether the person is suitable for a particular job role. This could provide insight into whether certain biases that impact the output or quality of work are tolerated in different roles. For example, the *compromise effect*, which causes an option to become more attractive when it becomes a middle option in a choice set, may be seen as acceptable for a position such as a minister (where making compromises decisions is a key aspect of the role), but not for a stock trader (where quick, decisive decisions are crucial). On the other hand, the sunk cost fallacy may yield opposite results.

Such alternative designs could offer increased efficiency and cost-effectiveness, as they may require a smaller sample size compared to the current approach. Such alternative designs would be able to assess the attitudes of all participants towards mistakes, while the current study only analyzes attitudes towards a mistake when it occurs, representing a small proportion of the data.

Thirdly, the results of this study may have practical implications for implementing soft interventions (nudges) based on cognitive biases. The acceptance and perception of these interventions may depend on the specific underlying bias, which could be further investigated through comparative studies. Some studies explored public attitudes towards particular soft government intervention and classic paternalistic policies to achieve the same policy goals (Arad & Rubinstein 2018; Hagman et al. 2019; Treger 2023). These could be used as complementary methodologies for future research. Therefore, it may be worthwhile to compare the public's attitudes towards different nudges that are striving to achieve the same policy goals.

In summary, potential future research directions on the normativity of biases, which are complementary, include exploring more realistic mistakes, evaluating mistakes using alternative methods and examining practical implications for implementing soft interventions.

Appendix A. The Five Biases

Five cognitive biases, fallacies, and heuristics were carefully chosen as mechanisms to drive participants to a choice of a dominated lottery. The selection process first included a thorough review of heuristics, biases, and "effects" (such as the compromise and attraction effects).

The main reason those five were chosen was practical – the possibility to implement them as a design of a problem that consists of a choice between two dice lotteries, where one lottery has FOSD over the other.

Other reasons included the variability of statistical difficulty that is attached to each problem – ranging from very easy (magical thinking) to very hard (base rate fallacy). This was important not only because of variability in the dependent variable, comfort level, which might originate from difficulty, but also contributes to the efficiency of the design. Only problems where a dominated lottery was chosen, are analyzed, as those are the instances where a cognitive bias occurs. Thus, there is a need to challenge participants with varying difficulty levels to get the most even out of those with high cognitive ability.

This design format, where participants are confronted with their mistakes, could be applied to more biases of course. Suggestions for additions to the design are presented in the discussion section.

Described below is the definition, the mechanisms that account for the bias, the domains where this bias was exhibited, and the axioms found to be violated by each of the cognitive biases used in the experiment. Despite the diversity of axioms described below, in the experiment, all biases led to a violation of first-order stochastic dominance, a choice of a dominated lottery.

Q1 Representative Conjunctions

The problem in this experiment implements the representativeness heuristic which is the mechanism of the conjunction fallacy.

Definition

Representativeness is an assessment of the degree of correspondence between a sample and a population, an instance and a category, an act and an actor, or, more generally, between an outcome and a model. For example, the model may refer to a person, a coin, or the world economy, and the respective outcomes could be marital status, a sequence of heads and tails, or the current price of bitcoin.

The conjunction fallacy (Famous by the known "Linda problem") occurs when it is assumed that specific conditions are more probable than a single general one (Tversky & Kahneman, 1982).

Mechanisms

Tversky and Kahneman (1982, 1983) proposed an explanation of the conjunction fallacy based on the hypothesis that people do not solve the task by applying extensional reasoning, but rather by using the intuitive heuristic – representativeness. They define representativeness as an assessment of the degree of correspondence between a model and an outcome.

Violations

The conjunction law $p(A \text{ and } B) \leq p(A)$.

Given two events, A and B, the probability of the event A and B is necessarily less than or equal to the likelihood of either of its constituents alone.

Domains

The Linda problem—since its original presentation (Tversky & Kahneman, 1982) there have been many experimental replications of this problem that try to mediate the fallacy rate. Some show how monetary incentives affect the fallacy rate (Zizzo et al., 2000), while Charness et al., (2010) also show that the ability to consult as a group lowers the rate.

Politics – Kanwisher (1989) suggests that the conjunction fallacy might underlie flawed arguments that often recur in debates on U.S. security policy, as strategic priorities have in the past become distorted by overemphasizing the most extreme scenarios at the expense of less flashy but more likely ones.

Other domains – Tversky and Kahneman (1983) show the conjunction rule violations in medical prognosis, decision making under risk, suspicion of criminal acts, and political forecasting.

Q2 Magical Thinking and Greediness in Uncertain Events

Definition

Magical thinking – the belief that one's actions can affect the outcome of some chance event, when in fact its likelihood is independent of those actions, is often referred to as magical thinking.

Mechanisms

Anthropology—magical thinking is predicated on the anthropological observation that humans experience, with their senses and intellect, a perception that all elements in the world are interrelated in a tapestry of natural interactions (Glucklich, 1997).

Tempting fate—the act of tempting fate heightens the accessibility of negative outcomes. This heightened accessibility leads to higher perceived probabilities of those outcomes, through the availability heuristic (Tversky & Kahneman, 1974). Risen & Gilovich (2008) tested whether reminding people of precautions they have taken leads them to see associated risks as less likely and detected an opposite "protection effect".

Belief in a just world–Lerner (1965) suggested that people need to believe that the world is just and form beliefs accordingly (for a literature review, see Furnham, 2003). This is similar to a belief in karma (the law of moral causation in Hinduism and Buddhism). According to both ideas, good behavior will be rewarded and bad behavior will be punished, either by the universe or possibly by a just God.

An illusion of control—an expectancy of inappropriate overconfidence, characterized by a belief in a personal success probability that exceeds the objective likelihood. The illusion of control refers to the belief that one's personal control or Influence can impact random events, while magical thinking encompasses a wider range of beliefs in which chance events are perceived as being under one's personal control. The introduction of skill-related factors, such as competition, choice, familiarity, and involvement, into chance situations can exacerbate these cognitive biases, causing individuals to feel an unwarranted sense of confidence in their outcomes. See Langer et al. (1975) for a meta-analytical review.

Domains

Tempting fate – students believe they are especially likely to be called on to answer a question in class if they have not done the required reading (Risen & Gilovich, 2008). People believe that they are especially likely to experience a mishap while traveling if they have not purchased travel insurance (Tykocinski, 2008).

An incentivized experiment involving a decision-making process similar to the one implemented in this design was presented in Arad (2014). It was found that many participants irrationally forgo the "greedy" option, of choosing the highest-paying lottery among equally likely ones.

The concept of magical thinking has also been implemented in theoretical models:

Strategic games – Daley and Sadowski (2017) formulated a behavioral model with magical thinking where players behave as if their expectations about their opponents' behavior vary with their own choices. The model provides a unified view of documented behavior in a range of often studied games (such as the prisoners' dilemma, the battle of the sexes, hawk-dove, and the stag hunt) and generates predictions across games.

Gambling—Passanisi (2017) tested the fit of an explanatory model of risk that starts from magical thinking and passes through maladaptive decision-making strategies, culminating with pathological gambling.

Q3 Gambler's Fallacy

Definition

The belief that, for random events, runs of a particular outcome (e.g., heads on the toss of a coin) will be balanced by a tendency for the opposite outcome (e.g., tails) .

 doi:10.20944/preprints202304.0643.v1

The recency effect—is the tendency to remember the most recently presented information best. Positive recency refers to "the hot hand fallacy" while negative recency refers to "the gambler's fallacy".

Mechanisms

A misunderstanding of sampling—in situations where a finite population of outcomes is sampled without replacement, expectations with negative recency have some validity because observing a particular outcome indeed lowers the chances of observing that outcome the next time. Accordingly, some authors have suggested that the experience of negative recency in life might be responsible for the gambler's fallacy in experimental tasks where participants are asked to generate or recognize random sequences (Ayton et al., 1989; Lopes & Oden, 1987).

Representativeness heuristic—Kahneman and Tversky (1972) presented a cognitive explanation of the gambler's fallacy in terms of the operation of the representativeness heuristic. They argued that people expect the essential characteristics of a chance process to be represented not only globally in an entire sequence of random outcomes but also locally in each of its parts. Thus, despite their statistical inevitability, long runs of the same outcome lack local representativeness and are thereby not perceived as representative of the expected output of a random device. Therefore, participants will expect runs of the same outcome to be less likely than they are. However, the idea that perceptions of randomness are governed by representativeness has also been used to explain the exact opposite phenomenon—that in a random sequence, people have an incorrect expectation that a run of the same outcome will continue: the "hot hand" fallacy.

The law of small numbers—the standard account and presupposition of the gambler's fallacy are that the gambler's fallacy represents an irrational belief in the law of small numbers (Tversky & Kahneman, 1971; Rabin, 2002): when one is committing the gambler's fallacy, the error one is making lies in believing that the probabilistic properties of large samples also apply to small samples.

Domains

Lab experiments that identify the gamblers fallacy – early probability learning experiments empirically confirmed the reality of this bias in tasks where participants were asked to predict the outcome in a series of random binary alternatives. Some of these experiments are cited in the mechanism section above.

Natural experiments and empirical findings – the gambler's fallacy (and the opposite "hot hand fallacy") are mechanisms to explain many irrational behaviors in different fields:

Investments—the disposition effect can be seen as an exhibition of the gambler's fallacy as investors sell winners too soon and hold losers too long (Odean, 1998; Shapira & Venezia, 2001; Chen et al., 2007).

Real gambling—the fallacy has been identified to be small, but significant in the behavior of actual casino gamblers (Croson & Sundali, 2005). Another example is Suetens et al. (2016) who found that players bet less on lottery numbers drawn in the previous week.

Sports – Misirlisoy and Haggard (2014) show that goalkeepers are prone to the gambler's fallacy: after a series of three kicks in the same direction, goalkeepers are more likely to dive in the opposite direction at the next kick.

Q4 Probability Matching

When faced with the task of choosing between two possible outcomes that differ in their probability of occurrence, some people match their choice probabilities to the corresponding outcome probabilities (matching) rather than always choosing the outcome with the higher probability (maximizing). For instance, in a task in which Outcome A occurs in 70% of the trials and Outcome B in the other 30%, probability matching involves predicting Outcome A in 70% of trials and B in the other 30%. By contrast, maximizing involves predicting the more likely Outcome A in every trial.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

This behavior can be considered a special case of irrational diversification, as it deviates from the rational strategy of always predicting the option with the highest probability (Rubinstein 2002).

Violations

Participants who match probabilities are simply not maximizing their expected utility.

Mechanisms

Attribute substitution—matching arises from a fast, relatively effortless intuitive assessment (Kogler & Kuhberger, 2007; West & Stanovich, 2003) that generates expected outcomes based on relevant probabilities (Tversky & Kahneman, 1971). This is a form of "attribute substitution" (Kahneman & Frederick, 2004) in which the answer to a relatively difficult question (how many Black and White guesses should be made?) is replaced by the answer to an easier one (how many Black and White outcomes are expected?).

Domains

Experiments with pigeons—to many psychologists, the term probability matching is associated with Herrnstein (1958) that used food-deprived pigeons to show early testimonies of probability matching.

The light guessing experiment—participants face two lights, their task being to predict which of the two lights (left or right) will shine at the end of the trial, given a certain probability. Participants are shown to match their frequency of left guesses with P(L). See Vulkan (2000) for a review.

Q5 Base Rate Fallacy

Failure to consider background (base rate) information in situations in which it is actually very relevant is called the base rate fallacy.

Many situations present the decision maker with two kinds of information: base rate (or background) information about how things usually are in such situations and indicator (or diagnostic) information telling how things appear to be in the particular situation. Unless the diagnostic information is excellent, the usual state (base rate) should be an essential guide to judging how they are at the moment. A statistical formula, the Bayes rule, tells precisely how these two kinds of information should be combined. The base-rate fallacy is people's tendency to ignore base rates in favor of individuating information (when such is available), rather than integrating the two.

Violations

Inaccurate probability judgments—the base rate fallacy can lead us to make inaccurate probability judgments in many different aspects of our lives.

Mechanisms

One of the most comprehensive attempts yet to account for the base rate fallacy is to be found in the claim of Tversky and Kahneman (1980, p. 50) that "base rate data that are given a causal interpretation affect judgments, while base rates that do not fit into a causal schema are dominated by causally relevant data". This claim is demonstrated using several inference problems, some based on Bar-Hillel (1980).

Another interpretation of Kahneman and Tversky's results was offered by Nisbett et al. (1976), who suggested that base rate information is ignored in favor of individuating information, since the former is "remote, pallid and abstract", whereas the latter is "vivid, salient, and concrete".

Relevance–Bar-Hillel (1980) suggested that we ignore base rate information because we believe it to be irrelevant to the judgment we are making. Bar-Hillel contends that, before making a judgment, we first categorize the information given to us into different levels of relevance. If something is deemed irrelevant, we discard it and do not factor it into the conclusion we draw. Thus, it is not that we are incapable of integrating different types of information. In fact, if two types of information are assigned equal relevance, we will give them equal consideration.

Domains

The cab problem—this famous problem served as an inspiration for the decision-making scenario implemented in this experiment. Two cab companies operate in a given city, the Blue and the Green (according to the color of the cab they run). Eighty-five percent of the cabs in the city are Blue, and the remaining 15% are Green. A cab was involved in a hit-and-run accident at night. A witness later identified the cab as a Green cab. The court tested the ability of the witness to distinguish between Blue and Green cabs under nighttime visibility conditions. It found that the witness was able to identify each color correctly about 80% of the time but confused it with the other color about 20% of the time. What do you think are the chances that the errant cab was indeed Green, as the witness claimed? (Kahneman & Tversky, 1972).

Lawyer-and-engineer item (Kahneman & Tversky, 1973).

Interpretation of statistical data and causality—I would like to present a recent case study taken from Israeli prime-time television news that serves to Illustrate the base rate fallacy. The report claimed that the efficacy of coronavirus vaccines was called into question as there were more patients hospitalized with COVID-19 among those who received the vaccine compared to those who did not. This conclusion, which implies a causal relationship between the vaccine and the virus, disregarded base rate Information. At the time of the report, over 80% of the population In Israel had already been vaccinated, so it is not surprising that there were more vaccinated patients in the hospital. To establish causality, the proportion of hospitalized patients relative to the total vaccinated and non-vaccinated populations should be compared.

Appendix B. The Experiments' Instructions

This is the general structure for each of the five problems:

- Screen 1: The participant is presented with a choice between two lotteries.
- Screen 2: The participant is asked to explain their choice in text form.
- Screen 3: If the participant chooses the dominated lottery, he is confronted with an explanation about his choice in one of the two treatments. In the *no-explanation* treatment, participants will only be presented with a description regarding the choice of a dominated lottery. In the other treatment, another paragraph will be shown, explaining the mechanism that might have led them to this choice, a cognitive bias or heuristic.

The participant is asked how he feels after reading this paragraph. He chooses a point on a fixed 7-point scale ranging from (-3) "I feel completely uncomfortable with my choice", (0) "neutral", and (3) "I feel completely comfortable with my choice".

Participants who chose the dominating lottery skip from screen 2 to screen 4.

- Screen 4: The participant is asked whether he is willing to share their choice (of lottery) with another participant along with their first name.

Response time is measured on each screen.

The experiment was conducted in Hebrew. Here is a translation of each of the five questions, followed by the confrontation in the two treatments. The full layout of Problem 1 is displayed below. For the other four problems, just the question screen and confrontation screen are shown. Following each question screen is a calculation of the difference in the expected payout between the two lotteries, which is designated to be 5 NIS (or very close). You might find minor grammatical mistakes in the problems. This is because they were translated from Hebrew, and I preferred to portray the original wording over the grammatical correctness when there was a doubt.

Q1 Representative Conjunctions (All Screens)

Screen 1. Question screen

Imagine a dice with 6 faces–2 faces are green (G) and 4 faces are red (R).

The dice will be rolled four times. You can choose to bet that the sequence of four throws will be RGRG or bet that the results in the last three throws will be GRG (no matter what came out in the first throw).

After your selection, the dice will be rolled four times and if the sequence you selected is drawn (when the results are described in the order of throwing from left to right), you will win 51 NIS.

Which sequence will you choose to bet on?

- The last three throws will be GRG.
- The sequence in four throws will be RGRG.

E_1 = The difference in expected payment:

$$E_1(C_1) = C_1 * [14.81\% - 4.94\%]$$

$$E_1(51) = 5.03 \approx 5$$

Screen 2. Participants' explanation

This was the question presented to you: Imagine a dice with 6 faces - 2 faces are green (G) and 4 faces are red (R). The dice will be rolled four times. You can choose to bet that the sequence of four throws will be RGRG or, alternatively, bet that the results in the last three throws will be GRG (no matter what came out in the first throw). After your selection, the dice will be rolled four times and if the sequence you selected is drawn (when the results are described in the order of throwing from left to right), you will win 51 NIS. Which sequence will you choose to bet on?

You chose to bet on [input his choice]

Please explain your choice:

[Text form answer]

Screen 3. Confrontation (only shown to participants who chose RGRG)

Imagine a dice with 6 faces – 2 faces are green (G) and 4 faces are red (R). The dice will be rolled four times. You can choose to bet that the sequence of four throws will be RGRG or, alternatively, bet that the results in the last three throws will be GRG (no matter what came out in the first throw). After your selection, the dice will be rolled four times and if the sequence you selected is drawn (when the results are described in the order of throwing from left to right), you will win 51 NIS. Which sequence will you choose to bet on?

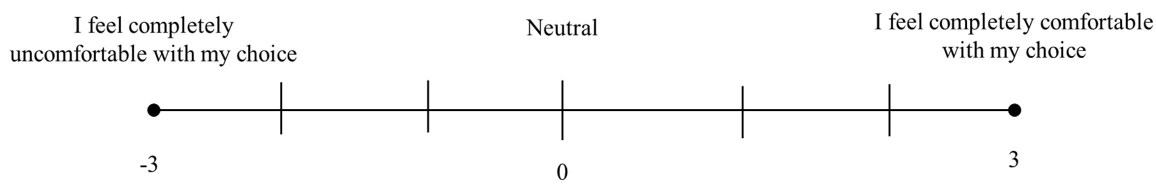
You chose to bet on [input his choice]

Note that there is a higher chance that the GRG sequence will appear last in the four-throw series than that the four-throw sequence will be RGRG. This is because the second sequence is the same as the first sequence, except for the addition of R at the beginning. Therefore, each time the RGRG sequence is realized, the sequence of the last three throws is GRG. Selecting the GRG sequence will result in a bet with a higher chance of winning than choosing the second sequence.

The addition to the *explanation* treatment:

You may have preferred the RGRG sequence over the GRG sequence following the so-called "representational heuristic"—in this example, if a particular sequence of throws seems more "representative", it will be interpreted as more plausible. According to research evidence, when people rely on representativeness, they may misjudge chances, since the fact that something is more representative does not necessarily make it more plausible.

How do you feel about your choice after reading this passage?



Screen 4. Willingness to share (presented to all participants)

This was the question presented to you: Imagine a dice with 6 faces – 2 faces are green (G) and 4 faces are red (R). The dice will be rolled four times. You can choose to bet that the sequence of four throws will be RGRG or, alternatively, bet that the results in the last three throws will be GRG (no matter what came out in the first throw). After your selection, the dice will be rolled four times and if the sequence you selected is drawn (when the results are described in the order of throwing from left to right), you will win 51 NIS. Which sequence will you choose to bet on?

You chose to bet on [input his choice]

Would you be willing for us to share your lottery of choice, stating your first name, [input first name] with another participant in the experiment?

- Yes
- No

Q2 Magical Thinking and Greediness in Uncertain Events

Question screen:

Imagine a standard dice with 6 faces with the numbers 1-6. You can choose to bet on any of the numbers 1-6.

Once you select a number, the dice will be rolled once. You will win a prize if the result of the dice roll is the number you choose.

You can bet on the number 6 and win 60 NIS if it comes out or bet on one of the other numbers: 1 or 2 or 3 or 4 or, 5 and win 30 NIS if the particular number you choose comes out.

Which number will you choose to bet on?

$E_2 =$ The difference in expected payment

$$E_2(C_{21}, C_{22}) = \frac{C_{21}}{6} - \frac{C_{22}}{6}$$

$$E_2(60,30) = 5$$

Confrontation screen:

The chance that the dice will stop on the number 6 is 1/6 regardless of the number you choose. Therefore, choosing number 6 is actually a choice to participate in a lottery with a 1/6 chance of winning 60 NIS. Choosing any number other than 6 is a selection of a lottery with the same chance of winning, 1/6, although the prize will be lower – only 30 NIS.

The addition to the *explanation* treatment:

You may have chosen not to bet on the number 6 because you intuitively believe that your chances of winning the prize are lower if you choose 6, perhaps because you are afraid of being "punished" for greed by a higher power.

There is research evidence for a type of "magical thinking" – according to which people believe that their actions can affect a random result when in practice this result does not depend on their choices. In particular, such thinking can lead to the belief that a greedy choice will lead to lower chances of winning.

In this example, magical thinking can lead to a gut feeling that the chance of winning will be lower if the number 6 (with the higher prize) is chosen since it is a greedy choice.

How do you feel about your choice after reading this passage?

Q3 Gambler's Fallacy

Question screen (1 out of 3)

Imagine a dice with 6 faces—4 faces are blue (B) and 2 faces are orange (O).

The dice will now be rolled four times.

After the rolls, you can choose which color to bet on prior to the fifth throw.

If the color you bet on comes out in the fifth roll, you will receive 15 NIS.

Click the arrow to continue the virtual throws.

Question screen (2 out of 3)

*An animation imitates a series of four rolls, from left to right.

Question screen (3 out of 3)

A dice with 6 faces—4 blue (B) faces, and two orange (O) faces, was rolled four times and the sequence (left to right) was: OBBB.

If the color you bet on comes out on the next roll, you will receive 15 NIS.

What color would you choose to bet on for the fifth roll?

- Orange (O)
- Blue (B)

Confrontation screen

In each roll of the dice, the probability of the roll result being B is 4/6 and the probability of the roll result being O is 2/6.

This is true regardless of the number of times we have thrown the dice and the result of previous throws since the dice does not remember the color that it landed on it in the past.

In other words, the previous roll of the dice does not affect the next roll.

Since the previous sequence is irrelevant, you should bet on B as the chance of B coming out is higher.

The addition to the *explanation* treatment

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

There is research evidence for the "gambler's fallacy", a misconception that if something happens relatively often relative to the expected frequency in a given period, then in the future it will happen less frequently (and vice versa) so that a balance is created. In this case, the gambler's fallacy can lead to the thought that the chance of B being drawn again is lower than the chance of O being drawn because of the sequence of previous throws.

How do you feel about your choice after reading this passage?

$E_3 =$ The difference in expected payment

$$E_3(C_3) = C_3 * \left[\frac{4}{6} - \frac{2}{6}\right] = C_3 * \frac{2}{6}$$

$$E_3(15) = 5$$

Q4 Probability Matching

Question screen

Imagine a dice with 6 faces—4 black (B) faces and 2 white (W) Faces.

The dice will be rolled six times and the sequence of blacks and whites will be recorded from left to right.

You can choose to bet on one of these two sequences: BBBBBB or BWBWBB.

You will win 76 NIS if the sequence you selected is drawn.

On what sequence would you like to bet?

- BBBBBB
- BWBWBB

$E_4 =$ The difference in expected payment

$$E_4(C_4) = C_4 * (8.78\% - 2.19\%) = C_4 * 6.58\%$$

$$E_4(76) = 5$$

Confrontation screen

The BBBBBB sequence is more likely than the BWBWBB sequence since the dice have more black than white faces. So, with each and every throw, the chance of a B is greater.

The probability of the sequence being BBBBBB is 8.7% while the chance of the sequence being BWBWBB is only 2.2%.

The addition to the *explanation* treatment

There is research evidence for "Probability matching", a case where people diversify by choosing from a variety of options in proportions similar to the probabilities of success.

How do you feel about your choice after reading this passage?

Q5 Base Rate Fallacy

Question screen

A box contains 20 dice, in two different sizes: 16 large dice and 4 small dice.

It is difficult to determine whether a dice is large or small without accurate measurement.

Suppose we have tested your identification ability and found that 70% of the time, you correctly identify the size of the dice and 30% of the time you are wrong.

Now suppose you pulled one dice out of the box and it seems large to you. You can choose to bet that the dice is large and win 18 NIS if it is indeed large.

Alternatively, you can participate in a lottery, and win 15 NIS with a 75% chance.

What will you choose?

- Participate in a lottery and win 15 NIS with a 75% chance.
- Bet that the dice is large and win 18 NIS if it turns out that it is indeed large.

$E_5 =$ The difference in expected payment

$$E_5(C_{51}, C_{52}) = 0.903 * C_{51} - 0.75C_{52}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

$$E_5(18{,}15) = 5.008$$

Confrontation screen

The chance that the dice you pulled out is indeed large is 90% (and not 70%) because you have to consider not only your chance of estimating the size of the dice but also the number of large and small dice in the box.

One has to weigh the fact that only in a few cases are the dice small, so if you did not have an estimate of the size of the dice from looking at it, it was much more likely to be large.

This means that when it looks big to you, the chances of this being correct are very high. On the other hand, even when it looks small to you, it is very likely that it is not small.

Therefore, a lottery in which you win NIS 15 with a 75% chance will pay a lower prize (15 NIS vs. 18 NIS) with a lower chance (75% vs. 90%)

To calculate the chance that the dice is indeed large we need to concentrate on the number of times that the dice actually pulled out will be classified as large. In 6% of the cases, a small dice will be mistakenly classified as large (4/20*0.3). In 56% of the cases, a large dice will be classified as large (16/20*0.7) and therefore in total, in 62% of the cases, a dice will be classified as large. Divide the number of times the dice is classified as large when it is really large out of the total number of times the dice is classified as large and we get the chance that the dice is indeed large:

56%/62% = 90.3%

The addition to the explanation treatment

There is evidence in the literature for the "base rate fallacy" according to which people ignore information about the relative probability of results.

The cube is really big in only 70% of the cases after ignoring the calculation of the chance that the cube is indeed big out of the total number of cases where you can think that the cube is big.

How do you feel about your choice after reading this passage?

Appendix C. Regression Tables

Table A1. A regression of comfort level on number of mistakes**.**

	Full sample (Without dummy variables)	*Explanation* treatment	*No-explanation* treatment
Constant	−0.991 ***	−1.237 ***	−0.777 ***
	(0.197)	(0.282)	(0.276)
Number of mistakes	0.189 **	0.345 ***	0.059
	(0.081)	(0.119)	(0.111)
R-squared	0.01	0.031	0.01
Adjusted R-squared	0.008	0.028	0
No. of observations	545	259	286

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively.

References

1. Arad, A. Avoiding greedy behavior in situations of uncertainty: The role of magical thinking. *J. Behav. Exp. Econ.* **2014**, *53*, 17–23.
2. Arad, A.; Rubinstein, A. The people's perspective on libertarian-paternalistic policies. *J. Law Econ.* **2018**, *61*, 311–333.
3. Aumann. A synthesis of behavioral and mainstream economics. *Nat. Hum. Behav.* **2019**, *3*, 666–670.
4. Ayton, P.; Hunt, A.J.; Wright, G. Psychological conceptions of randomness. *J. Behav. Decis. Mak.* **1989**, *2*, 221–238.
5. Bar-Hillel, M. The base-rate fallacy in probability judgments. *Acta Psychol.* **1980**, *44*, 211–233.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2023 doi:10.20944/preprints202304.0643.v1

6. Bechara, A.; Damasio, A.R. The somatic marker hypothesis: A neural theory of economic decision. *Games Econ. Behav.* **2005**, *52*, 336–372.
7. Besharov, G. Second-best vonsiderations in correcting cognitive biases. *South. Econ. J.* **2004**, *71*, 12–20.
8. Blanco, F. Cognitive bias. *Encycl. Anim. Cogn. Behav.* **2017**, 1.
9. Bleske-Rechek, A.; Nelson, L.A.; Baker, J.P.; Remiker, M.W.; Brandt, S.J. Evolution and the trolley problem: People save five over one unless the one is young, genetically related, or a romantic partner. *J. Soc. Evol. Cult. Psychol.* **2010**, *4*, 115.
10. Charness, G.; Karni, E.; Levin, D. On the conjunction fallacy in probability judgment: New experimental evidence regarding Linda. *Games Econ. Behav.* **2010**, *68*, 551–556.
11. Chen, G.; Kim, K.A.; Nofsinger, J.R.; Rui, O.M. Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *J. Behav. Decis. Mak.* **2007**, *20*, 425–451.
12. Cleaves, D.A. Cognitive biases and corrective techniques: Proposals for improving elicitation procedures for knowledge-based systems. *Int. J. Man-Mach. Stud.* **1987**, *27*, 155–166.
13. Croson, R.; Sundali, J. The gambler's fallacy and the hot hand: Empirical data from casinos. *J. Risk Uncertain.* **2005**, *30*, 195–209.
14. Daley, B.; Sadowski, P. Magical thinking: A representation result. *Theor. Econ.* **2017**, *12*, 909–956.
15. Delfabbro, P. The stubborn logic of regular gamblers: Obstacles and dilemmas in cognitive gambling research. *J. Gambl. Stud.* **2004**, *20*, 1–21.
16. Delfabbro, P.H.; Winefeld, A.H. Predictors of irrational thinking in regular slot machine gamblers. *J. Psychol.* **2000**, *134*, 117–128.
17. Felsen, G.; Castelo, N.; Reiner, P.B. Decisional enhancement and autonomy: Public attitudes towards overt and covert nudges. *Judgm. Decis. Mak.* **2013**, 8.
18. Furnham, A. Belief in a just world: Research progress over the past decade. *Personal. Individ. Differ.* **2003**, *34*, 795–817.
19. Gaboury, A.; Ladouceur, R. Erroneous perceptions and gambling. *J. Soc. Behav. Personal.* **1989**, *4*, 411.
20. Gigerenzer, G. The bias bias in behavioral economics. *Rev. Behav. Econ.* **2018**, *5*, 303–336.
21. Gilboa, I.; Schmeidler, D. *A Theory of Case-Based Decisions*; Cambridge University Press: 2001.
22. Gilovich, T.; Griffin, D.; Kahneman, D. *Heuristics and Biases: The Psychology of Intuitive Judgment*; Cambridge University Press: New York, 2002.
23. Glucklich, A. *The End of Magic*; Oxford University Press: 1997.
24. Hagmann, D.; Ho, E.H.; Loewenstein, G. Nudging out support for a carbon tax. *Nat. Clim. Change* **2019**, *9*, 484–489.
25. Hallion, L.S.; Ruscio, A.M. A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychol. Bull.* **2001**, *137*, 940.
26. Herrnstein, R.J. SECTION OF PSYCHOLOGY: Some factors influencing behavior in a two-response situation. *Trans. N. Y. Acad. Sci.* **1958**, *21*, 35–45.
27. Jones, E.B.; Sharpe, L. Cognitive bias modification: A review of meta-analyses. *J. Affect. Disord.* **2017**, *223*, 175–183.
28. Kahneman, D. A perspective on judgment and choice: Mapping bounded rationality. *Am. Psychol.* **2003**, *58*, 697.
29. Kahneman, D.; Frederick, S. Attribute substitution in intuitive judgment. *Model. A Man Essays Mem. Herbert A Simon* **2004**, 411–432.
30. Kahneman, D.; Tversky, A. Subjective probability: A judgment of representativeness. *Cogn. Psychol.* **1972**, *3*, 430–454.
31. Kahneman, D.; Tversky, A. On the psychology of prediction. *Psychol. Rev.* **1973**, *80*, 237.
32. Kahneman, D.; Tversky, A. On the study of statistical intuitions. *Cognition* **1982**, *11*, 123–141.
33. Kahneman, D.; Tversky, A. Choices, values, and frames. *Am. Psychol.* **1984**, *39*, 341–350.
34. Kanwisher, N. Cognitive heuristics and American security policy. *J. Confl. Resolut.* **1989**, *33*, 652–675.
35. Kogler, C.; Kühberger, A. Dual process theories: A key for understanding the diversification bias? *J. Risk Uncertain.* **2007**, *34*, 145–154.
36. Korteling, J.E.; Toet, A. Cognitive biases. In *Encyclopedia of Behavioral Neuroscience*; 2020.
37. Langer, E.J. The illusion of control. *J. Personal. Soc. Psychol.* **1975**, *32*, 311.
38. Lerner, M.J. Evaluation of performance as a function of performer's reward and attractiveness. *J. Personal. Soc. Psychol.* **1965**, *1*, 355.
39. Lopes, L.L.; Oden, G.C. Distinguishing between random and nonrandom events. *J. Exp. Psychol. Learn. Mem. Cogn.* **1987**, *13*, 392.
40. Misirlisoy, E.; Haggard, P. Asymmetric predictability and cognitive competition in football penalty shootouts. *Curr. Biol.* **2014**, *24*, 1918–1922.
41. Nielsen, K.; Rehbeck, J. When choices are mistakes. *Am. Econ. Rev.* **2022**, *112*, 2237–2268.

42. Nisbett, R.E.; Borgida, E.; Crandall, R.; Reed, H. Popular induction: Information is not necessarily informative. In *Cognition and Social Behavior*; Carroll, J.S., Payne, J.W., Eds.; Lawrence Erlbaum: 1976.
43. Obermaier, M.; Koch, T.; Baden, C. Everybody follows the crowd? *J. Media Psychol*. *J. Media Psychol. Theor. Methods Appl.* **2015**, *29*, 69–80.
44. Odean, T. Are investors reluctant to realize their losses? *J. Financ.* **1998**, *53*, 1775–1798.
45. Passanisi, A.; Craparo, G.; Pace, U. Magical thinking and decision-making strategies among late adolescent regular gamblers: A mediation model. *J. Adolesc.* **2017**, *59*, 51–58.
46. Rabin, M. Inference by believers in the law of small numbers. *Q. J. Econ.* **2002**, *117*, 775–816.
47. Risen, J.L.; Gilovich, T. Why people are reluctant to tempt fate. *J. Personal. Soc. Psychol.* **2008**, *95*, 293.
48. Roth, S.; Robbert, T.; Straus, L. On the sunk-cost effect in economic decision-making: A meta-analytic review. *Bus. Res.* **2015**, *8*, 99–138.
49. Rubinstein, A. Irrational diversification in multiple decision problems. *Eur. Econ. Rev.* **2002**, *46*, 1369–1378.
50. Shapira, Z.; Venezia, I. Patterns of behavior of professionally managed and independent investors. *J. Bank. Financ.* **2001**, *25*, 1573–1587.
51. Simon, H.A. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*; Macmillan: 1947.
52. Simon, H.A. *Models of Man; Social and Rational*; Wiley: 1957.
53. Simon, H.A. From substantive to procedural rationality. In T.J. Kastelein, S.K.; 1976.
54. Smith, V.L. Constructivist and ecological rationality in economics. *Am. Econ. Rev.* **2003**, *93*, 465–508.
55. Suetens, S.; Galbo-Jørgensen, C.B.; Tyran, J.R. Predicting lotto numbers: A natural experiment on the gambler's fallacy and the hot-hand fallacy. *J. Eur. Econ. Assoc.* **2016**, *14*, 584–607.
56. Treger, C. When do people accept government paternalism? Theory and experimental evidence. *Regul. Gov.* **2023**, *17*, 195–214.
57. Tversky, A.; Kahneman, D. Belief in the law of small numbers. *Psychol. Bull.* **1971**, *76*, 105.
58. Tversky, A.; Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science* **1974**, *185*, 1124–1131.
59. Tversky, A.; Kahneman, D. *The Framing of Decisions and the Rationality of Choice*; Stanford University, California, Department of Psychology: 1980.
60. Tversky, A.; Kahneman, D. (Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychol. Rev.* **1983**, *90*, 293.
61. Tykocinski, O.E. Insurance, risk, and magical thinking. *Personal. Soc. Psychol. Bull.* **2008**, *34*, 1346–1356.
62. Vulkan, N. An economist's perspective on probability matching. *J. Econ. Surv.* **2000**, *14*, 101–118.
63. West, R.F.; Stanovich, K.E. Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Mem. Cogn.* **2003**, *31*, 243–251.
64. Zizzo, D.J.; Stolarz-Fantino, S.; Wen, J.; Fantino, E. A violation of the monotonicity axiom: Experimental evidence on the conjunction fallacy. *J. Econ. Behav. Organ.* **2000**, *41*, 263–276.