5-2-2024

Essays on Judgment and Decision Making

Chengyao Sun
Washington University in St. Louis, chengyao@wustl.edu

WASHINGTON UNIVERSITY IN ST. LOUIS

Olin Business School

Dissertation Examination Committee:
Cynthia Cryder, Co-Chair
Robyn LeBoeuf, Co-Chair
Fausto Gonzalez
Hannah Perfecto
Scott Rick

Essays on Judgment and Decision Making

by

Chengyao Sun

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy in Business Administration

May 2024
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my deepest gratitude to my amazing advisors, Cynthia Cryder and Robyn LeBoeuf, for giving me the best doctoral journey I could ever imagine. Thank you for creating such an encouraging, inspiring, and rigorous environment for me to grow as a researcher. Thank you for always being so patient, understanding, and knowledgeable when I faced headwinds. It is difficult to overstate how grateful I am to be your doctoral advisee.

I would like to thank Scott Rick for his expertise and unwavering support that are invaluable in shaping critical parts of this dissertation.

I would like to extend my appreciation to Fausto Gonzalez and Hannah Perfecto who helped me navigate challenges in research and beyond.

I offer special thanks to every faculty member on the Consumer Behavior track of the Marketing group for their financial support for my research pursuits.

Finally, I am deeply indebted to my parents. Their unconditional love sustained me through the highs and lows of this journey.

Chengyao Sun

Washington University in St. Louis

March 2024

To my girlfriend, Jinglin

ABSTRACT OF THE DISSERTATION

Essays on Judgment and Decision Making

by

Chengyao Sun

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2022

Professor Cynthia Cryder, Co-Chair

Professor Robyn A. LeBoeuf, Co-Chair

In this dissertation, I describe two programs of research to demonstrate how logically irrelevant factors can influence everyday judgment and decisions. Chapter 1 looks at a common human behavior: prediction. In collaboration with Robyn LeBoeuf, we investigate a particular factor that may bias people's predictions against their own judgment: the absolute likelihood of the most likely outcome. Previous research usually assumes that, when making a prediction from a set of possible outcomes, people select as their prediction the outcome that seems most likely to them. However, we find a disconnect between what people predict and what they believe to be most likely to arise. We find that the disconnect arises because people are sensitive to the absolute likelihood of the most likely outcome, although the absolute likelihood does not determine the best prediction that maximizes accuracy. Specifically, when the most likely outcome has a low (vs. high) absolute likelihood, people less often choose the most likely outcome as their prediction—even though they still believe this outcome is most likely to arise.

Chapter 2 looks at a more specific domain of consumer behaviors: co-branded credit card usage. In collaboration with Cynthia Cryder and Scott Rick, we find that credit card co-branding discourages people from using a co-branded credit card outside the card's featured brand. Co-branded credit cards are typically backed by a payment-processing network such as Mastercard or Visa and can be used anywhere the network is accepted. However, across one descriptive survey and four experiments, we show that consumers are less likely to use a co-branded credit card (compared to its non-co-branded counterpart) for purchases that do not match the featured brand, even when the co-branded card maximizes the cashback reward. We identify two mechanisms. First, the featured brand on a co-branded credit card produces assumptions about the card's reward structure and those assumptions limit consumers' attention to the actual reward structure. Second, the featured brand on a co-branded credit card makes purchases outside of the featured brand feel like a bad "fit", discouraging consumers from using the card outside of the brand. We discuss both consumer and managerial implications.

Chapter 1: Prediction That Conflicts with Judgment

1.1 Introduction

People often make predictions about uncertain events that have several possible outcomes. A traveler may need to forecast whether it will be rainy, cloudy, or sunny on their trip. A voter may want to predict the winner in an election. A parent may speculate about which college their child will attend. A basketball fan may want to predict which team will win the title. One might expect people to predict what they believe to be most likely: that is, if a person thinks Kansas is *most likely* to win March Madness, they would *predict* Kansas to be the winner. Does this statement, albeit intuitive, always reflect behavior? We suggest that it does not, and we document a robust disconnect between prediction and likelihood judgment.

1.1.1 Bias in Prediction and Likelihood Judgment

Extensive research has shown that likelihood judgments and predictions can be biased. Much early research focused on how such judgments and predictions diverge from formal probability models. One well-known finding is that people often rely on shortcuts or heuristics, rather than a formal calculus, to make judgments and predictions (e.g., Kahneman et al., 1982; Tversky & Kahneman, 1974). For example, people may predict by considering the degree to which the key characteristics of the available evidence resemble a possible outcome while ignoring the base rate or probability of that outcome (Kahneman & Tversky, 1973). Judgments and predictions have also been shown to be influenced by a host of factors including, but not limited to, affect (Loewenstein et al., 2001; Slovic et al., 2004, 2007), optimism (Krizan et al., 2009; Massey et al., 2011; Weinstein, 1980), and past history (e.g., Bar-Hillel & Wagenaar, 1991; Gilovich et al., 1985; Jarvik, 1951). For example, people tend to assume that events are more likely to cause harm when those events trigger negative emotions such as dread (Fischhoff et al., 1978; Slovic, 1987), and football partisans tend to give optimistically biased predictions about their favorite teams (Massey et al., 2011; Simmons & Massey, 2012). People's predictions are also biased by their intuitions and their confidence in those intuitions, leading them to predict, for example, that sports teams will beat the point spread when they are more confident in their intuitions that those teams will win, even when that intuitive confidence is objectively uninformative about the likelihood of beating the spread (Simmons & Nelson, 2006).

1.1.2 Prediction versus Likelihood Judgment

Although prior research on judgment and prediction has taken many different directions and explored many potential biasing factors, one thing that unites this research is that people's predictions are thought to follow from, and be consistent with, their likelihood judgments. That is, if a person thinks, for whatever reason, that heads is more likely to come up than tails on an upcoming coin flip, that person will also predict that heads will come up.

This assumption is very intuitive and often true. Although little research has directly examined whether predictions and subjective likelihood judgments correspond—probably because such connection usually seems obvious—some work happens to provide evidence that supports this correspondence. For example, Simmons and Massey (2012) show that football fans who optimistically *predicted* their preferred team as the winner indeed *estimated* that their team had a higher likelihood of winning the game than their opponent, even when their preferred team was objectively inferior to the other. In other papers, researchers, quite reasonably, observe predictions as a way of gauging likelihood judgments (and vice versa), indicating a tacit assumption that the two likely often correspond. For example, in demonstrating the representativeness heuristic, Kahneman and Tversky often treated predictions and likelihood judgments as interchangeable measures. In a classic study, they measured people's *predictions* of a graduate student's field of study by asking participants to rank the possible fields in order of their *likelihoods* (Kahneman & Tversky, 1973). In another study, they illustrated a bias in subjective *probabilities* by measuring people's *predictions* about which program a class of students were from (Kahneman & Tversky, 1972). Their view of the relationship between likelihood (or frequency) and prediction was, again quite reasonably, "[i]n category prediction, one predicts the most frequent category" (Kahneman & Tversky, 1973). [1]

However, it is not always the case that people predict that their perceived most likely outcome will arise. Some anomalies have been identified. One such case is the phenomenon of probability matching. When predicting for a class of events that each have the same outcome probabilities, people tend to broadly match their predictions to the probabilities, sometimes disregarding what is most likely for a single event. For example, when predicting a repeated drawing that has a 70% chance of giving red on each draw and a 30% chance of giving black, people may predict red on 70% of the trials and black on 30% even when they are clearly aware that red is always more likely (e.g., Koehler & James, 2009; Neimark & Shuford, 1959; for an extensive review, see Vulkan, 2000). Another case involves the desirability bias. Making an outcome more desirable (e.g., associating it with a monetary payoff) biases people's predictions toward that outcome (e.g., Marks, 1951). However, although people tend to predict in the desired direction,

[1] As described by Kahneman and Tversky (1973), "category prediction" is any prediction that requires people to predict an uncertain event that has several possible nominal outcomes, like predicting the outcome of a coin toss, the result of a roll of the die, or the winner of a tournament. The current paper focuses exclusively on category prediction, so for simplicity, we use the term "prediction."

their likelihood judgments are less affected by desirability (Park et al., 2022; Windschitl et al., 2010; see also Bar-Hillel & Budescu, 1995).

Although not strictly about prediction, research has also identified a disconnect between likelihood judgment and choice that arises when intuitive perceptions conflict with rational analysis. For example, when people try to draw a red bean from a bowl of beans, some people prefer to draw from a bowl of 100 beans that contains 9 red beans over a bowl of 10 beans that contains one red bean. Participants report feeling that the bowl with 9 red beans gives them more ways to win, even though they report knowing that the likelihood of winning is greater in the bowl with one red bean (ratio bias; Denes-Raj & Epstein, 1994). Such a finding suggests that perceptions of what will happen may diverge from pure likelihood assessments.

In this paper, we investigate a different factor that causes prediction and likelihood to diverge: the sense that even the most likely outcome is nevertheless not very likely in an absolute sense. If one's goal is to maximize predictive accuracy, one should focus on the relative likelihoods of the possible outcomes and choose the outcome that is more likely to arise than any other alternative, regardless of whether that outcome has a high or low likelihood itself. However, we suggest that people may take the absolute likelihood of the most likely outcome into account, even when doing so cannot help improve predictions. That is, they may not only consider what is *most likely* but also *what is likely* when predicting.

Why might people focus on absolute likelihood? We suggest two main reasons. First, absolute likelihood is a fundamental element of a probabilistic event that people almost always need to consider, at least initially. When predicting, one first must attend to each outcome's absolute likelihood to deduce the relative likelihoods, although absolute likelihood usually becomes irrelevant once the most likely outcome is determined. However, we suggest that even after people determine the most likely outcome, absolute likelihood may remain salient. Second, in many cases, it may be useful and important to focus on absolute likelihood, especially when people need to consider a particular possibility independently. For example, when booking a flight, people may need to evaluate the absolute likelihood of any accidental disruption to determine whether they need travel insurance. In such cases, absolute likelihood is an important consideration itself. Thus, people may often need to consider absolute likelihood when confronting uncertainty, and they may continue to do so even in situations, such as the ones studied here, where only relative likelihood matters.

Thus, when people predict which of a set of outcomes will arise, we suggest that they may focus both on relative likelihood (which is important for predicting optimally) and also absolute likelihood (which is irrelevant for an optimal prediction). When people perceive the most likely outcome to have a high absolute likelihood, we expect them to predict that their most likely outcome will arise because both the relative likelihood and the absolute likelihood converge towards that outcome.

However, there may be times when people feel that even the most likely outcome's likelihood is objectively low (e.g., the most likely outcome has a 10% chance of arising, and all other possible outcomes have lower but non-zero chances of arising). In such cases, even though the most likely outcome may be clear, no option may seem particularly likely to arise. Instead, because of the low absolute likelihood of the most likely outcome, the final outcome may feel relatively difficult to foresee; that is, people may think "anything can happen" and may feel that there is no longer a clear right or wrong answer. People may thus feel free to predict arbitrarily, by which we mean predicting via an explicitly non-logical method, such as choosing randomly, going with a gut feeling, choosing a desired outcome (e.g., a favorite team, a "fun" outcome, or a lucky number), or simply guessing. Such a process may lead people's predictions to depart from what they perceive to be the most likely outcome.

To make this more concrete, consider again the person who thinks Kansas is most likely to win the championship among all 68 teams in March Madness. If they think Kansas has an 80% chance of winning the title, they can easily predict that Kansas will be the champion: their prediction will be the same as the team they think is subjectively most likely to win.

However, if they think Kansas, albeit the most likely, only has a 10% chance of winning the title, they might conclude that Kansas is not, overall, particularly likely to win. After all, from their point of view, there is a 90% chance that some other team might win the title. The substantial uncertainty posed by the 90% chance of any other team winning could make the actual winner feel hard to foresee. This feeling of low foreseeability may license people to predict arbitrarily, in a manner that is decoupled from their perception of which team is most likely to win: they may rely on a hunch, a guess, or a team they like. Thus, although they still may believe that Kansas is more likely to win than other teams, because that likelihood of winning seems rather low to them, they may predict something other than Kansas.

We thus predict that likelihood judgments will correspond to predictions as long as the most likely outcome also seems likely overall to people, but that predictions will depart from likelihood judgments when the most likely outcome seems overall unlikely (even though the perception of which outcome is most likely remains intact). In other words, we anticipate that, when the absolute likelihood of the most likely outcome is low, there will be a gap between what participants identify as most likely and what they predict. We suggest that this gap is driven (a) by outcomes seeming difficult to foresee when the most likely outcome is unlikely in an absolute sense and (b) by this sense of low foreseeability leading people to predict arbitrarily (e.g., by relying on strategies apart from conventional logic). Because we anticipate that people are sensitive to what is likely (absolute likelihood) in addition to what is most likely (relative likelihood), we also suggest that increasing the focus on relative likelihood or decreasing the focus on the low absolute likelihood will reduce this gap (i.e., will cause predictions to align with likelihood judgments).

This paper thus examines how predictions and likelihood judgments may diverge when the absolute likelihood of the most likely outcome is low. It contributes to research on judgment and decision making by showing that probability judgments and predictions cannot be assumed to be the same. While some important research has indeed shown that predictions and probability judgments may diverge, that research has generally focused on relatively specialized circumstances (e.g., probability matching; Koehler & James, 2009; optimism; Park et al., 2022). Here, we identify a factor that is arguably more pervasive – low absolute likelihood – and show how it can distort predictions.

1.1.3 The Current Research

We begin by showing that people's predictions and their perceived most likely outcomes are largely consistent when the most likely outcome is likely to arise but tend to diverge when the most likely outcome is unlikely to arise (Studies 1, 2a, and 2b). We show that this gap between most likely outcomes and predicted outcomes emerges even when participants are incentivized to make accurate predictions (Studies 1 and 2b) and even when most-likely judgments and predictions are made within moments of each other (Studies 2a and 2b). Studies 3 and 4 provide further evidence from predictions of real-life events: We collect people's predictions for the 2022 NBA championship (Study 3) and 2022 March Madness (Study 4). We find that people are less likely to predict as the winner the team that they think is most likely to win when they think that this team is unlikely to win in an absolute sense (despite being most likely to win in a relative sense). Studies 5a and 5b show that, when the most likely outcome is overall unlikely, people find the final outcome to be less foreseeable, and that they in turn acknowledge choosing randomly and arbitrarily.

Studies 6 through 8 show that emphasizing the relative likelihood of the most likely outcome or reducing the focus on the low absolute likelihood of that outcome reduces the gap between likelihood judgments and predictions. Finally, Study 9 identifies an interesting boundary of the current effects: the gap between likelihood judgments and predictions is much smaller when people advise others than when they predict for themselves, suggesting that people may find it less appropriate to predict arbitrarily when advising others. Study 9 also shows that taking the role of an advisor can debias people when they later predict for themselves.

In all studies, we report all measures, manipulations, conditions, and exclusion criteria. All studies except for Study 2b are preregistered. We report all preregistered analyses in all studies. Occasionally, a preregistered analysis is not central to our main argument, and so we put it in the Appendix and note it in the main text. All preregistration documents, study materials, and data can be found on OSF: https://osf.io/frmz2/?view_only=c5527906c1fe449db06c370e8915eb3f The research was approved by the Institutional Review Board at the authors' institution.

1.2 Study 1: Initial Demonstration

Study 1 examines people's predictions and likelihood judgments in a simple game. The game has an obvious most likely outcome, but we manipulate that outcome's absolute likelihood to be high versus low. We also manipulate whether we ask participants to identify the most likely outcome or to predict which outcome will arise. We predict that participants will easily identify the most likely outcome regardless of whether its absolute likelihood is low or high. However, we predict that absolute likelihood will matter for predictions. When the most likely outcome has a high absolute likelihood, we predict that participants will regularly predict that it will obtain, just like they will regularly recognize it as most likely. However, when the most likely outcome has a low absolute likelihood, we predict that participants will be less likely to choose it as their prediction, even though they will still have no trouble identifying it as most likely.

Prior research suggests that outcomes that are considered "unlikely" usually have a probability around 20-30%, whereas those considered "likely" generally have a probability above 50%, with an average probability around 70% (Budescu & Wallsten, 1995; Clark, 1990; Sirota & Juanchich, 2015; Theil, 2002). Thus, in this and the following studies (except for Studies 3 and 4), we manipulate the most likely outcome to be unlikely or likely by setting its likelihood to approximately 20% or 70%, respectively.

1.2.1 Method

Participants and design. We preregistered (https://aspredicted.org/3MG_7KS) to recruit 600 participants on Amazon Mechanical Turk (MTurk). We recruited 601, but 37 did not pass an attention check (described below) and were excluded (as preregistered), leaving a final sample of 564 (M_{age} = 40.3 years; 52.3% female, 45.9% male, 1.6% non-binary, and 0.2% preferring not to say). Participants were randomly assigned to one cell of a 2 (likelihood of the most likely outcome: high vs. low) by 2 (response: identify the most likely outcome vs. predict the outcome) between-subject design.

Procedure. During an online session, participants played a computerized game. Each participant saw a set of nine numbered balls on the screen and could click to randomly draw a ball from the set. As illustrated in Figure 1.1, participants saw one of two sets of balls: In the low-chance set, two balls were labeled “1” and the other seven balls were labeled a unique number from “2” to “8”. In the high-chance set, six balls were labeled “1” and the remaining three balls were labeled “2” to “4”. Although “1” was the most likely number to be drawn from both sets, “1” had a low chance (2/9) of being drawn from the former set and a high chance (6/9) of being drawn from the latter set. Before drawing a ball, participants responded in one of two ways: they either *identified* the most likely outcome (“Which number are you most likely to draw?”) or they *predicted* which number they would draw (“Which number do you predict you will draw?”). Predictors also read that they would win $1.00 if they successfully predicted the number they drew. Then, participants drew a ball and observed the outcome. At the end, they answered an attention check that asked them to identify how many balls were marked “1.”

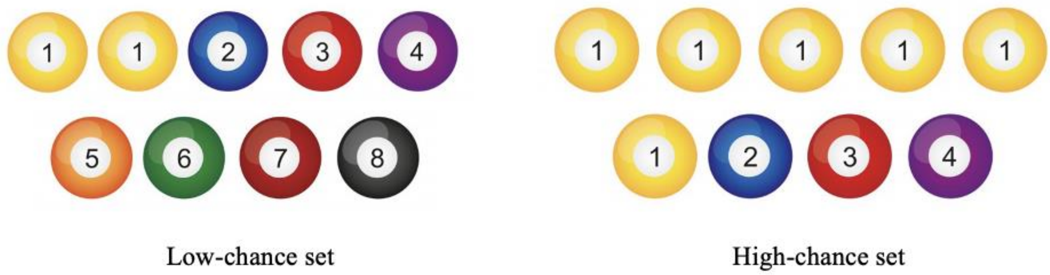

Figure 1.1 Study 1 Stimuli

1.2.2 Results and Discussion

Most participants correctly identified 1 as the most likely number, regardless of whether the likelihood of drawing a 1 was low (92.6%) or high (96.7%, $\chi^2(1) = 1.62$, $p = .203$). However, predictions were sensitive to likelihood, even though participants were incentivized for correct predictions. Participants were less likely to predict 1 as the number that would be drawn when the likelihood of drawing a 1 was low (63.3%) versus high (92.4%, $\chi^2(1) = 31.82$, $p < .001$, see Figure 1.2).

Viewed differently, when there was a high chance of drawing a 1, the percentage identifying 1 as the most likely (96.7%) was not reliably different from the percentage predicting 1 (92.4%, $\chi^2(1) = 1.74$, $p = .188$), suggesting a close correspondence between what participants judged most likely and what they predicted. However, when there was a low chance of drawing a 1, there was a gap between the percentage identifying 1 as the most likely (92.6%) and the percentage predicting 1 (63.3%, $\chi^2(1) = 32.84$, $p < .001$). This most-likely vs. prediction gap reveals that participants tended to choose a different and less likely number as their prediction even when they knew that 1 was most likely to be drawn. Appendix A shows the full distribution of responses in each condition. As seen in Appendix A, there is not a clear regularity in terms of which number people predict when they do not predict 1; Studies 5a and 5b will return to the issue of what governs people's predictions in these cases.

Thus, when people predict from a set of possible outcomes, they do not always predict the most likely outcome, despite recognizing it as most likely. When people can recognize a most likely outcome and it feels likely in an absolute sense, they predict with ease that it will arise. However, when the most likely outcome feels unlikely in an absolute sense, people are less likely to choose that outcome as their prediction—even though they can still easily identify which number is most likely and stand to gain money from making a correct prediction.

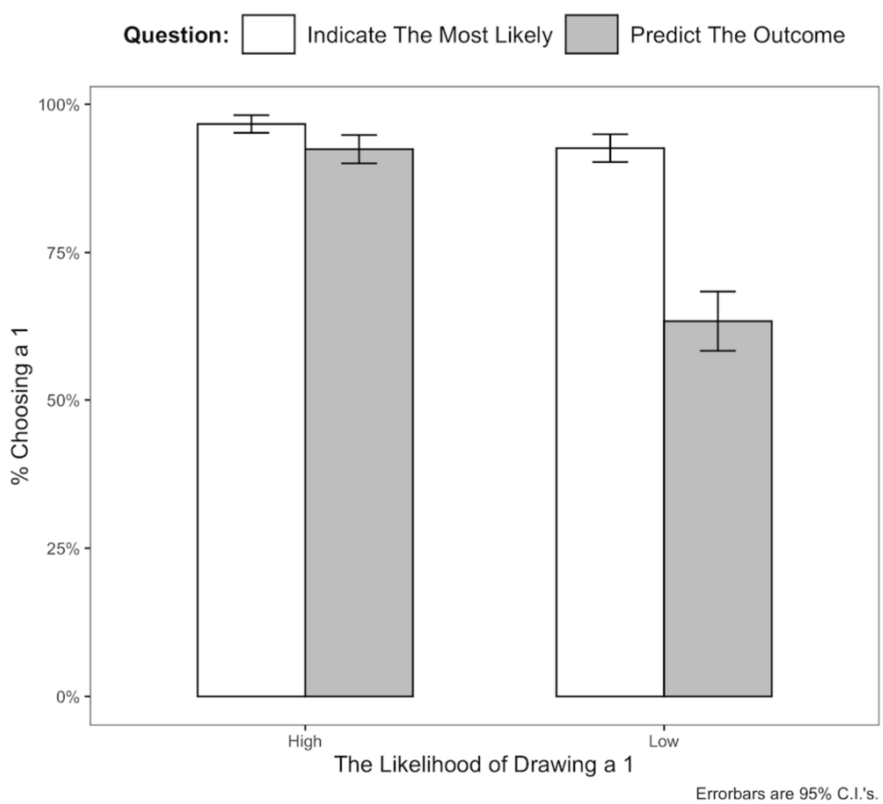

Figure 1.2 Study 1 Results

1.3 Study 2a: Within-Subject Design

In Study 1, we measured people's predictions and likelihood judgments between subjects. Studies 2a and 2b examine these responses within subjects. This design allows us to see if the most-likely vs. prediction gap persists even when participants give their prediction immediately following their most-likely response (and vice-versa); this amounts to a fairly stringent test of whether participants truly mean to predict something different from what they think is most likely. This design also allows us to examine an alternative explanation: It is possible that, in the low likelihood conditions, participants did not assess the likelihoods of the possible outcomes before predicting, but they were able to identify the most likely outcome when they were directly asked about it. That is, rather than predicting contrary to their likelihood judgments, perhaps they simply did not make likelihood judgments at all when predicting. A within-subjects design addresses this explanation by prompting some participants to first identify the most likely outcome before making a prediction.

In Study 2a, we only examine the low-likelihood conditions because that is where we observe the most-likely vs. prediction gap.

1.3.1 Method

Participants. As preregistered (https://aspredicted.org/1VK_BS2), we recruited 300 MTurk participants[2]. None were excluded from the analysis.

Procedure. Participants considered a set of nine (virtual) balls. Two balls were marked "1" and seven were marked "2" to "8". Before randomly drawing a ball from the set, they both identified

[2] We did not collect demographic data in this study. The CloudResearch platform provides overall information about gender (51% female and 49% male) and birth decade (1% from the 1940s, 6% from the 1950s, 12% from the 1960s, 16% from the 1970s, 37% from the 1980s, 25% from the 1990s, 2% from the 2000s, and 1% unknown).

which number they were most likely to draw and predicted which number they would draw. Between these two questions, we told participants, "your answer here can be the same as, or different from, the answer to the previous question." The order of the two tasks was counterbalanced.

1.3.2 Results and Discussion

As preregistered, we first examined participants' responses to their first question for a between-subject comparison. We replicated the most-likely vs. prediction gap. Of those who first identified which number was most likely to be drawn, almost all (93.3%) correctly identified 1 as most likely, but of those who first made a prediction, only 61.3% predicted that 1 would arise ($\chi^2(1) = 42.01$, $p < .001$).

Next, we examined participants' responses to both questions for a within-subject comparison. The most-likely vs. prediction gap persisted within participants. Of the 300 participants, most (88.0%) identified 1 as most likely, but only 59.7% predicted that they would draw a 1 ($\chi^2(1) = 60.87$, $p < .001$). This gap appeared in both question orders. When participants first predicted, 61.3% of them predicted a 1 but subsequently many more of them (82.7%) recognized 1 as most likely ($\chi^2(1) = 15.89$, $p < .001$). Perhaps more notably, when participants first identified the most likely number, almost all of them (93.3%) identified 1, but then only 58.0% predicted 1 immediately thereafter ($\chi^2(1) = 48.95$, $p < .001$). Thus, many people still did not predict the most likely number even when they had just recognized and explicitly stated that number. Predicting after identifying the most likely number did not increase the percentage predicting a 1 ($P_{\text{prediction second}} = 58.0\%$ vs. $P_{\text{prediction first}} = 61.3\%$, $\chi^2(1) = .22$, $p = .638$) even though participants were more likely to identify 1 as most likely before versus after prediction ($P_{\text{prediction second}} = 93.3\%$ vs.

$P_{\text{prediction first}}$ = 82.7%, $\chi^2(1) = 7.10$, $p = .008$). Appendix A shows the full response distribution. Appendix B describes an additional analysis that further examines how many participants were internally (in)consistent.

Study 2a set out to examine whether people's predictions still diverge from the most likely outcome even when the responses are given within moments of each other. We find that, indeed, this divergence still arises in such a setting, suggesting that it does not arise because people have not thought about the most likely outcome before they make a prediction. Study 2b examines this issue further.

1.4 Study 2b: A Replication

Study 2a showed that the most-likely vs. prediction gap persisted even within subjects. Is this because participants really think of the answers to the two questions as different, or is there a more mundane explanation for the discrepancy? For example, perhaps participants did not explicitly connect the two questions, or perhaps they were not motivated to give their best predictions. Study 2b examines these possibilities. All participants in this study first identify the most likely outcome. Then, some simply predict the outcome, while others are reminded of their most-likely responses before making their predictions and are given a monetary incentive for an accurate prediction. These steps allow us to further examine whether the most-likely vs. prediction gap within subjects reflects a real difference between likelihood judgments and predictions, versus a lack of attention or motivation.

1.4.1 Method

Participants and design. We intended to recruit a sample of 300 participants on MTurk. Of the 301 participants who completed the study, 12 failed the attention check, leaving a final sample of 289 (M_{age} = 41.28 years; 48.8% female and 51.2% male). Participants were randomly assigned to one of two conditions: no incentive and no reminder vs. incentive and reminder.

Procedure. As in Study 2a, participants in the no-incentive-and-no-reminder condition considered a set of nine balls. Two balls were marked “1” and seven balls were marked “2” through “8”. Before drawing a ball, they first identified the most likely number, and then they predicted the number that would be drawn. Participants were told that their answer to the second question could be the same as or different from their answer to the first.

The incentive-and-reminder condition followed the same procedure with two additional features: 1) when participants made predictions, they were reminded of the most likely number that they previously identified; 2) participants read that they would win a $1.00 bonus if they correctly predicted the outcome of the drawing. Finally, all participants answered an attention-check question that asked them to recall how many balls were marked “1.”

1.4.2 Results and Discussion

We replicated Study 2a’s findings in both conditions. In the no-incentive-and-no-reminder condition, almost all participants (97.9%) identified 1 as most likely, but only 56.6% predicted that they would draw a 1 ($\chi^2(1) = 68.29$, $p < .001$). In the incentive-and-reminder condition, also 97.9% of participants recognized 1 as most likely, and only 66.0% predicted that they would draw a 1 ($\chi^2(1) = 47.52$, $p < .001$). The percentage predicting 1 increased slightly but not reliably with the reminder and incentive ($P_{\text{no-incentive-and-no-reminder}}$ = 56.6% vs. $P_{\text{incentive-and-reminder}}$ = 66.0%, $\chi^2(1) = 2.32$, $p = .128$). Thus, the reminder and incentive neither closed the most-likely vs. prediction gap nor reliably improved predictions. Appendix A shows the full response distribution. Appendix B describes an additional analysis that examines how many participants were internally (in)consistent.

The studies thus far suggest that, when a most likely outcome is unlikely overall, even though people can easily identify that outcome, reliably fewer predict that this outcome will be the one to arise. Studies 2a and 2b show that this pattern arises even if people are incentivized to make accurate predictions and even when those predictions are made immediately after participants have identified the most likely outcome.

1.5 Study 3: 2022 NBA Championship

The first three studies provided consistent experimental evidence for the disconnect between predictions and likelihood judgments: when a most likely outcome is unlikely to arise, people may be able to easily identify it as most likely, but they are less likely to choose it as their prediction, seemingly disregarding their beliefs about it being most likely. In Studies 3 and 4, we seek further evidence from predictions of real-life events.

Every year, millions of people follow the postseason tournament of the National Basketball Association (NBA). Before the 2022 postseason started, we collected people's perceived most likely title-winner, their incentivized predictions of the winner, and their belief about whether their most likely team has a high or low absolute likelihood to win the title.

We examine how each person's likelihood assessments of their own most likely team winning relates to their tendency to predict that team to win. One might expect that people will predict their own most likely team as the winner regardless of the subjective absolute likelihood of that team winning: If Jay thinks that the Phoenix Suns are most likely to win the title, he will predict them as the winner. Yet, we suggest that people will be less likely to predict their own most likely team as the winner when they perceive that this team has a low, rather than high, absolute chance of winning. Thus, if Jay picks the Phoenix Suns as his most likely team and thinks their chances are good overall, he will predict them as the winner. However, if Jay thinks Phoenix is the most likely team to win, but if he thinks their overall chances are nevertheless relatively low, he will be more likely to predict another team to win – disregarding his stated belief that the Phoenix is more likely to win than any other team.

1.5.1 Method

Participants. After the 2021-2022 NBA regular season concluded on April 10, 2022, 20 teams advanced to the postseason. The first postseason game was scheduled to start on April 12 at 7:00 pm Eastern. We preregistered to recruit 600 participants from MTurk on April 11 and to stop the recruitment before 7:00 pm Eastern April 12 even if we did not recruit 600 participants (https://aspredicted.org/662_NK8). We recruited 601 participants by the night of April 11 (M_{age} = 39.3 years; 33.3% female, 65.7% male, 0.2% selecting "other," and 0.8% preferring not to say).

Procedure. Participants answered three questions in one of two orders. In one order, they first predicted which of the 20 teams would win the championship. On the next screen, they indicated which team was most likely to win the championship. On the final screen, they indicated whether their most likely winner had a low or high chance of winning the title, by choosing one of the following options:

> I think the [team] is more likely than any other team to win the title, but despite that, I don't think they have a particularly high likelihood of winning the title in an absolute sense.
>
> I think the [team] is more likely than any other team to win the title, and I also think that they have a high likelihood of winning the title in an absolute sense.

In the other order, participants first indicated which team was most likely to win the title. Then, they predicted which team would win the title before finally indicating whether their most likely winner had a low or high absolute chance of winning the title. In both orders, participants were told that they would win $1.00 if their prediction was correct. At the end of the survey, participants indicated whether they followed the NBA and how frequently they watched NBA games on a scale from 0 (never) to 10 (very frequently). Controlling for these measures did not affect any of our results, so we do not discuss them further.

1.5.2 Results and Discussion

On the absolute-likelihood question, 330 participants indicated that their selected most likely winner had a low chance of winning the title in an absolute sense, and 271 indicated that their most likely winner had a high chance.

In keeping with our main prediction, fewer participants predicted their own selected most likely winner as the winner when they indicated that this team had a low absolute chance of winning the title, compared to when they indicated that this team had a high absolute chance ($P_{\text{low likelihood}}$ = 54.2% vs. $P_{\text{high likelihood}}$ = 76.8%, $\chi^2(1) = 31.91$, $p < .001$). This effect emerged both when participants first made a prediction before indicating the most likely winner ($P_{\text{low likelihood}}$ = 50.3% vs. $P_{\text{high likelihood}}$ = 74.8%, $\chi^2(1) = 17.83$, $p < .001$) and when they first indicated the most likely winner before predicting ($P_{\text{low likelihood}}$ = 57.9% vs. $P_{\text{high likelihood}}$ = 78.8%, $\chi^2(1) = 13.77$, $p < .001$).

Study 3 extends our effect to a real-world situation. When people indicate the most likely winner of the NBA championship, their predictions of the winner match that team when people feel that the most likely winner has a high chance of winning. But when people think that their own most likely winner has a lower chance of winning, they are more likely to disregard their beliefs about which team is likely to win, and to predict a different team as the winner.

1.6 Study 4: 2022 March Madness

Study 4 seeks to replicate and extend Study 3's findings with incentivized predictions of another real-life event: March Madness 2022. In Study 4, we elicit absolute likelihood assessments in a different way from Study 3. Rather than asking for a binary choice, we ask people to estimate the percentage likelihood of each eligible team winning the title, including their perceived most likely winner. This method has the virtue of being subtler and also of allowing a continuous analysis of people's sensitivity to (perceived) absolute likelihood.

1.6.1 Method

Participants. The final eight teams in March Madness were decided on March 25, 2022. On March 26 at 6:09pm Eastern, those teams would begin to compete for the final four positions.

We preregistered to collect 600 responses on MTurk on March 26 and to stop data collection before 6:09pm even if we did not reach 600 responses (https://aspredicted.org/SMZ_T11). We ended up obtaining 602 responses by 3:53pm Eastern on March 26 (M_{age} = 42.5 years; 45.2% female, 53.8% male, 0.3% selecting "other," and 0.5% preferring not to say).

Procedure. Participants answered three questions in one of two orders. In one order, they first indicated which of the final eight teams was most likely to win the title. On the next screen, they estimated each team's likelihood of winning the title (as a percentage; each participant gave eight percentages that were required to sum to 100%). On the third screen, participants predicted which of the eight teams would win the title and were told that they would win $2.00 if their prediction was correct. In the other order, participants first made an incentivized prediction and then indicated the most likely winner before rating each team's likelihood. At the end of the survey, participants indicated whether they followed college basketball and how frequently they watched college basketball games on a scale from 0 (never) to 10 (very frequently). Controlling for these measures did not affect any of our results, so we do not discuss them further.

1.6.2 Results and Discussion

The most likely winner. There was a possibility for people to be inconsistent when identifying the most likely winner. Participants both directly indicated the most likely winner and estimated the percentage likelihood of each team winning the title. This latter estimate could also reveal beliefs about the most likely winner. Most participants (514 of 602) were consistent between their directly indicated most likely winner and their most likely winner as revealed through percentage likelihoods. We focus on these 514 participants in the analyses below. We also report full-sample analyses, all of which yield the same conclusions, in Appendix C.

Main analyses. We first created a dependent variable that equaled 1 if a participant's prediction matched their most likely winner and 0 if it did not. We ran a logistic regression regressing this dependent variable on the reported percentage likelihood of how likely the most likely team was to win, question order (1 = participants first indicated the most likely winner, -1 = participants predicted first), and their interaction.

As predicted, predictions were sensitive to the reported likelihood of the most likely winner: participants were less likely to predict their own most likely team when they reported the absolute likelihood of that team winning to be lower ($\beta = .031$, $SE = .008$, $p < .001$). This effect was not qualified by question order, as the likelihood x order interaction was not significant ($\beta = -.021$, $SE = .017$, $p = .210$). There was an overall main effect of order that is not directly relevant to our predictions: participants were more likely to predict their own most likely team to win when they first indicated the most likely team than when they first predicted ($\beta = 1.00$, $SE = .298$, $p < .001$). Thus, people's predictions are more likely to diverge from their own perceived most likely winner when they perceive that their most likely winner has a lower absolute chance of winning. Appendix C describes a dichotomous analysis that reaches the same conclusion.

Study 4 illustrates the robustness of the prediction-likelihood disconnect in two ways. First, it conceptually replicates the findings of Study 3 with a different real-life prediction. Second, it demonstrates that the effect emerges whether we measure people's perceptions of their team's overall likelihood of winning with a qualitative sense of likelihood (Study 3) or via percentage likelihood (Study 4). Of course, both studies 3 and 4 are correlational studies. We simply collected people's judgments and predictions about real-life events without any experimental manipulations. The results are thus not free of potential confounds. Nevertheless, these two studies provide strong correlational evidence for the discrepancy between likelihood judgment and prediction in ecologically valid settings. These studies, combined with the previous experiments, show compelling evidence that people disregard what they know what to be most likely in specific circumstances, namely, when what is most likely does not appear likely in an absolute sense.

1.7 Study 5a: Free Responses

We have shown consistent evidence that people tend to predict contrary to what they know to be most likely to arise when the most likely outcome is unlikely to happen in an absolute sense. But why do people do this? We have suggested that the most-likely vs. prediction gap may arise because, in such situations, the final outcome seems hard to foresee, which may in turn license people to predict in a less logical and more arbitrary fashion, such as by picking a favorite number, a lucky number, or just making a pure guess. Studies 5a and 5b investigate this potential process, beginning in study 5a by simply asking participants why they predicted something that was different from their professed most likely outcome.

In Study 5a, people give both a most-likely assessment and a prediction. We invite those who give inconsistent responses to the two questions to tell us why their responses diverge. We examine participants' responses to give us some insight into the low absolute likelihood effect.

1.7.1 Method

Participants. As preregistered (https://aspredicted.org/W91_BQB), we recruited 500 participants from Prolific. Eleven did not pass the attention check as described next, leaving a final sample of 489 (M_{age} = 36.3 years; 47.4% female, 49.7% male, and 2.7% selecting "other").

Procedure. Participants saw the low-chance set of balls (Figure 1). Before drawing a ball, they first indicated the most likely number and then predicted which number they would draw. If a participant indicated 1 as most likely but did not predict 1, we then asked them, “What was your reasoning for predicting that you would get [predicted number] instead of a 1?” Participants who gave consistent most-likely and prediction responses or who did not indicate 1 as most likely skipped this open-ended question.

We next asked participants to code their responses in the following way: Each participant reviewed their response and then reported whether it referred to each of the six reasons listed in Table 1.1, by responding “Yes” or “No.” The reasons were presented in a randomized order. Participants could answer “Yes” or “No” to as many reasons as they felt were applicable to them.

Collectively, the six reasons shown in Table 1.1 reflected our hypotheses about why the most-likely vs. prediction gap arises. The first reason in Table 1.1 referred to sense that even the most likely outcome is unlikely to arise. The second and third reasons referred to the perceived low foreseeability of the outcome. The fourth through sixth reasons referred to participants predicting arbitrarily in various ways, such as picking a liked number or a lucky number, or going with a gut feeling.

At the end of the survey, all participants answered an attention check that asked them to recall how many balls were marked “1.”

1.7.2 Results and Discussion

Predictions. We replicated the most-likely vs. prediction gap. Most participants (90.2%) indicated that 1 was most likely to be drawn but only 63.6% of them predicted a 1 ($\chi^2(1) = 95.76$, $p < .001$).

Free responses. One hundred forty-five participants indicated 1 as most likely but then predicted another number. They explained their thought processes in the subsequent free response question. Their explanations were generally consistent with our proposed process. They often acknowledged that their predictions were influenced by the low likelihood of drawing a 1. For example, one said, "Although there are two 1s, the likelihood of getting it is still pretty low in comparison to the others, as it is still a 2/9 chance that you will get a 1. So, despite being a tad hopeful, chances are, I won't get a 1." Others further communicated the difficulty of foreseeing the outcome and mentioned predicted arbitrarily as a result. One explained, "I just picked a random ball that wasn't one of the two 1's. Because you never know..." As expected, participants used various arbitrary (non-logical) strategies. Some chose a number they liked, as one said, "It is my lucky number." Others went with their gut feelings and frankly said so: "I just went with a gut feeling." Many of the rest simply chose a random number, as one described, "I closed my eyes, shook my cursor and it let choose."

To get a more systematic sense of these responses, we can examine how participants coded their responses according to the six reasons shown in Table 1.1. The six reasons represent different aspects of our proposed process, and we were interested in whether participants' verbal responses reflected any or all aspects of that process. Over half of these participants (51.7%) agreed that their response reflected, "The likelihood of drawing a 1 was small overall and/ or 1 was overall unlikely to be drawn." This suggests that the absolute likelihood of the most likely outcome indeed influenced people's predictions. Even more participants resonated with the sense of low foreseeability: 80.7% agreed that their response reflected, "The drawing is random and anything could happen," and 68.3% agreed that their response reflected, "The outcome is hard to predict or know in advance."

Our account suggests that this low foreseeability could promote arbitrary predictions of various kinds. Indeed, 40.7% agreed that their response reflected, "I picked a number I just liked for some reason, such as my lucky number, my birthday, my favorite number, and so on," and 66.2% agreed that their response reflected, "I guessed or picked a number at random." Finally, 71.7% agreed that their response reflected, "I went with my gut feeling." (These percentages sum to more than 100%, reflecting that these categories are not mutually exclusive.) Collectively, the six reasons covered all responses: no one answered "No" to all of them.

In Study 5a, we see that people's explanations for why their judgments and predictions diverged fit with several aspects of our proposed process. Although participants' verbal explanations may not always accurately reflect the forces that drive their behavior (Nisbett & Wilson, 1977), the fact that their verbal reports converge with our hypotheses is a sign that our proposed account may capture some of the reasons that predictions diverge from most likely assessments. Study 5b tests our proposed process in a more structured way.

Table 1.1 Study 5a Self-Coded Results.

Reason	% Yes
The likelihood of drawing a 1 was small overall and/ or 1 was overall unlikely to be drawn.	51.7%
The drawing is random and anything could happen.	80.7%
The outcome is hard to predict or know in advance.	68.3%
I picked a number I just liked for some reason, such as my lucky number, my birthday, my favorite number, and so on.	40.7%
I guessed or picked a number at random.	66.2%
I went with my gut feeling.	71.7%

Note: The "% Yes" column records the percentage of the total participants who indicated that their response referred to the corresponding reason.

1.8 Study 5b: Mediation

As noted, when making predictions about the outcome of an uncertain event, we propose that people attend to the absolute likelihood of the most likely outcome, and not just its relative likelihood. We further suggest that when this absolute likelihood is small, people consider the outcome to be rather difficult to foresee, because even this most likely outcome seems unlikely to happen. We suggest that this sense of low foreseeability promotes arbitrary predictions that are not based on reasoning about relative likelihood but rather are based on things like gut feelings, personal preferences, or even simply guesses. Study 5a gave some suggestion that participants might go through this process. Study 5b measures perceptions of foreseeability and measures how participants claim to make their predictions to test this process more formally.

1.8.1 Method

Participants and design. We preregistered to recruit 300 participants (https://aspredicted.org/7DH_BB2), and 301 completed the study. As pre-registered, we excluded participants who failed the attention check (n = 15), leaving 286 participants for our analyses (M_{age} = 41.1 years; 42.7% female, 54.5% male, 1.0% non-binary, and 1.4% preferring not to say). Participants were randomly assigned to either the low-chance or high-chance condition.

Procedure. Participants saw either the low-chance set or the high-chance set shown in Figure 1. Before drawing a ball, they first indicated which number was most likely to be drawn and then predicted which number would be drawn.

After prediction, participants rated two sets of items. The first set contained three statements, order randomized, that asked participants what they based their predictions on: "My prediction was based on subjective or personal factors, such as a gut feeling, a lucky number, a pure guess, or something similar;" "my prediction was based on the objective probabilities of drawing different numbers;" and "my prediction was based on logic and reasoning." For each item, participants responded on a scale ranging from 0 (disagree strongly) to 10 (agree strongly).

The second set contained another three statements, order randomized, that measured how foreseeable the outcome felt: "On a scale from 0 (very uncertain) to 10 (very certain), how certain versus uncertain do you feel about which number will be drawn?;" "on a scale from 0 (very difficult to predict) to 10 (very easy to predict), how easy do you think it is to predict which number will be drawn?;" "on a scale from 0 (very unforeseeable) to 10 (very foreseeable), how foreseeable is the number that will be drawn?"

Then, participants clicked a button to draw a ball. At the end, they answered an attention check that asked them to recall how many balls were marked "1."

1.8.2 Results and Discussion

First, we replicated the discrepancy between predictions and likelihood judgments. Most participants correctly identified 1 as most likely regardless of whether they were in the high-chance or low-chance condition ($P_{high\text{-}chance}$ = 96.6% vs. $P_{low\text{-}chance}$ = 93.4%, $\chi^2(1) = .97$, $p = .325$). Predictions, however, reliably differed between conditions. Fewer participants predicted a "1" in the low-chance condition than in the high-chance condition ($P_{high\text{-}chance}$ = 91.9% vs. $P_{low\text{-}chance}$ = 59.1%, $\chi^2(1) = 40.64$, $p < .001$). Thus, in the high-chance condition, the percentage identifying 1 as most likely did not reliably differ from the percentage predicting a 1 ($P_{most\text{-}likely}$ = 96.6% vs. $P_{predict}$ = 91.9%, $\chi^2(1) = 2.25$, $p = .134$). However, in the low-chance condition, there was a sizable and reliable gap between the percentage identifying 1 as most likely and the percentage predicting 1 ($P_{most\text{-}likely}$ = 93.4% vs. $P_{predict}$ = 59.1%, $\chi^2(1) = 42.68$, $p < .001$).

Next, we examined the means of our proposed process measures. We averaged responses to the three "foreseeability" items to create an index of how foreseeable the outcome felt ($\alpha = .93$). Scores were higher, indicating greater perceived foreseeability, in the high-chance condition than in the low-chance condition ($M_{high\text{-}chance}$ = 6.53, $M_{low\text{-}chance}$ = 2.92, $t(261.6) = 15.83$, $p < .001$)[3].

[3] This and following t-tests are Welch's t-tests (without the assumption of equal variances). The degrees of freedom were approximated using the Welch–Satterthwaite equation (Satterthwaite, 1946).

We also averaged responses to the three "bases of prediction" items (with the first item reverse-coded) to create an index that assessed the degree to which participants reported making predictions based on logical reasoning ($\alpha = .93$). Scores were higher, indicating more reported logical reasoning, in the high-chance condition than in the low-chance condition ($M_{high\text{-}chance}$ = 7.87, $M_{low\text{-}chance}$ = 6.04, $t(252.2) = 4.31$, $p < .001$).

We next tested, via serial mediation, whether the change in the absolute likelihood affected people's perceptions of the foreseeability of the outcome, which in turn corresponded to how they reported making their prediction (see Figure 1.3). For this analysis, we fitted the mediation model with the responses from the great majority of participants (272 of 286, or 95.1%) who correctly indicated 1 as most likely.[4] We created a dependent variable, dubbed "judgment-prediction correspondence," that would equal 1 if a participant predicted a 1 and 0 if they did not predict a 1. We included the foreseeability index and the logical reasoning index as potential mediators. Finally, we coded the independent variable, dubbed "low absolute likelihood," as 1 if the most likely number (i.e., 1) had a low chance of being drawn and 0 if it had a high chance.

Results based on 5,000 bootstrapped samples showed a statistically significant total effect of low absolute likelihood on judgment-prediction correspondence ($\beta_{total} = -.202$, bootstrapped SE = .067, $p = .003$). This total effect was serially mediated by foreseeability and logical reasoning, supported by a statistically significant indirect effect ($\beta_{indirect} = -.149$, bootstrapped $SE = .039$, $p < .001$). The remaining direct effect did not reliably differ from 0 ($b_6 = -.053$, bootstrapped SE = .047, $p = .258$). Specifically, the low absolute likelihood made the outcome feel difficult to foresee ($b_1 = -3.692$, bootstrapped $SE = .228$, $p < .001$). This lack of foreseeability corresponded

[4] We did this because our main manipulation focused on "1," and so it seemed cleanest to restrict our analysis to those who reported "1" as the most likely outcome. We also fitted the model with the full sample, as reported in Appendix D. The full-sample results were consistent with the current findings.

to people making a prediction that was more likely to depart from logical reasoning (b_2 = .501, bootstrapped *SE* = .117, *p* < .001). Finally, predictions that were less logical were more likely to diverge from predictors' self-reported most likely outcome (b_3 = .080, bootstrapped *SE* = .006, *p* < .001).

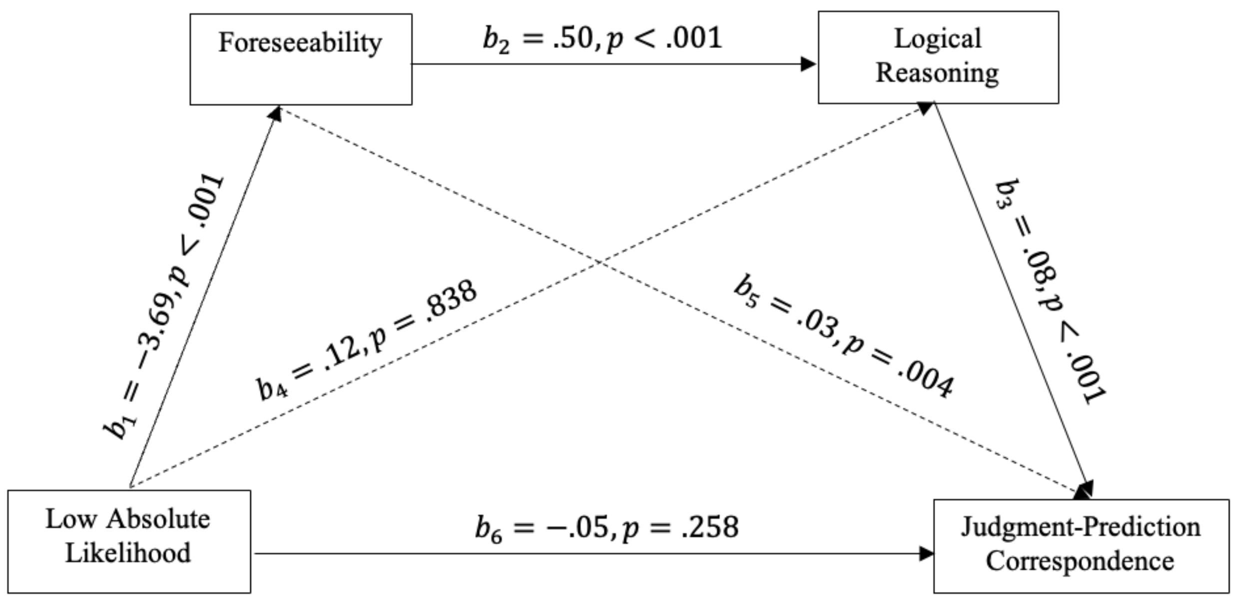

Figure 1.3 Study 5b Mediation

Study 5b provides evidence supporting our hypothesized process for why judgments and predictions diverge when absolute likelihood is low. When the absolute chance of the most likely outcome arising is low, the outcome can feel difficult to foresee. The sense of low foreseeability corresponds to people's tendency to predict arbitrarily. Such arbitrary prediction strategies then correspond to the divergence between prediction and the most likely outcome.

1.9 Study 6: Greater Relative Likelihood

So far, we have shown that people's predictions tend to depart from their perceived most likely outcome when the most likely outcome has a low overall likelihood of arising. To maximize accuracy, people should make a prediction only based on the relative likelihood (i.e., which outcome is more likely to arise than others?). However, we argue that they also consider the absolute likelihood (i.e., how likely is this outcome to arise overall?). Study 5b showed that, when the most likely outcome is overall unlikely to arise, it can feel hard to foresee, which promotes arbitrary guesses that deviate from the optimal prediction.

If it is true that our effects arise because people focus on both relative and absolute likelihood, highlighting the relative likelihood may direct some attention back to this factor and consequently move predictions closer to likelihood judgments. Studies 6 and 7 do this in different ways. One simple way to highlight the relative likelihood is to make it larger. In Study 6, we thus manipulate the relative likelihood of the most likely outcome while holding its absolute likelihood constant. We predict that more people will predict in line with the obvious most likely outcome when it has a greater likelihood relative to others, even when its absolute likelihood remains low.

1.9.1 Method

Participants and design. As preregistered, we recruited 600 participants from MTurk (https://aspredicted.org/65J_8CL). Fourteen of them did not pass an attention check (described below), leaving us with 586 observations (M_{age} = 40.3 years; 52.2% female, 47.3% male, and 0.5% selecting "other"). Participants were randomly assigned to the lower or higher relative likelihood condition.

Procedure. Participants drew from one of two virtual sets of balls as shown in Figure 1.4. In the lower-relative-likelihood condition, participants drew from a set of 10 balls. Two balls were labeled "1," and the other eight were labeled a unique number from "2" to "9." In the higher-relative-likelihood condition, participants drew from a set of 100 balls. Twenty of them were labeled "1," and the other eighty were labeled a unique number from "2" to "81." Thus, although "1" had a 20% chance of being drawn from both sets, it was twice as likely as the other numbers in the lower-relative-likelihood condition but twenty times more likely than the other numbers in the higher-relative-likelihood condition.

Participants both predicted which number they would draw and indicated which number was most likely to be drawn in counterbalanced order. At the end of the survey, participants answered an attention check that asked them to recall how many balls were labeled "1."

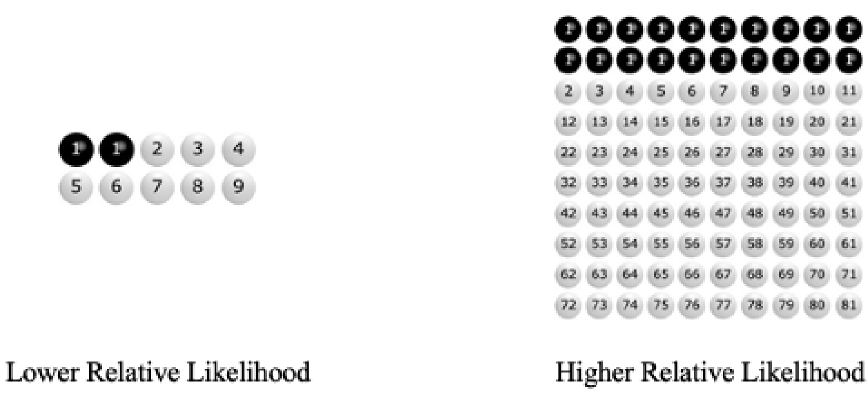

Figure 1.4 Study 6 Stimuli

1.9.2 Results and Discussion

Participants answered both a most-likely question and a prediction question in a counterbalanced order. We focus here on a within-subjects analysis that compares participants' responses to both questions; Appendix E contains a between-subjects analysis that uses only responses to the first question a participant saw. As in previous studies, there was a large most-likely vs. prediction gap within the lower-relative-likelihood condition: 85.6% of the participants indicated 1 as most likely but only 57.9% of them predicted a 1 ($\chi^2(1) = 54.06$, $p < .001$). However, this gap was narrowed and was not reliable in the higher-relative-likelihood condition, with 86.3% identifying 1 as most likely and 75.0% of participants predicting it ($\chi^2(1) = 2.07$, $p = .150$). Furthermore, among the majority of participants who correctly indicated 1 as most likely, 34.0% did not predict a 1 in the lower-relative-likelihood conditions whereas only 9.9% did not predict a 1 in the higher-relative-likelihood conditions ($\chi^2(1) = 41.13$, $p < .001$). Thus, the internal inconsistency was attenuated by the greater relative likelihood. This pattern persisted in both question orders (see Appendix E).

We have suggested that the most-likely vs. prediction gap may arise because people focus on both relative likelihood and absolute likelihood. Study 6 suggests that we can increase the attention paid to relative likelihood, even when absolute likelihood remains low, by increasing the relative likelihood. When the most likely outcome had a greater relative advantage, predictions were less likely to diverge from judgments about what is most likely. That said, because we increased relative likelihood by increasing the number of balls marked "1," perhaps our manipulation also increased perceptions of the absolute likelihood of drawing a 1 (e.g., via the ratio bias, Denes-Raj & Epstein, 1994). This is possible. Study 8 addresses manipulating perceptions of the low absolute likelihood more directly, but first, study 7 uses a manipulation that does not involve changing the likelihoods or how they are presented.

1.10 Study 7: Ranking Alternative Predictions

Study 6 increased the focus on relative (vs. absolute) likelihood by changing the relative likelihood advantage of the most likely outcome. In study 7, we attempt to emphasize the relative likelihood without increasing it. To achieve this, we ask all participants to consider a game where the most likely outcome has a low absolute chance of being drawn. We ask some participants to make a standard prediction, as in prior studies, but ask others to rank different possible predictions that they could make. Specifically, we ask them to list the top two predictions they would make, that is, their best and second-best predictions. Our logic here is that generating a best and second-best prediction might encourage people to directly and deliberately compare the likelihood among the possible outcomes and therefore focus more on relative likelihood, compared to when making a single prediction. Thus, we predict that people will be more likely to choose the most likely outcome as their best prediction in the two-prediction condition, compared to when they just make a single non-ranked prediction, because we predict that people will focus more on relative likelihood in the former condition than in the latter.

1.10.1 Method

Participants and design. We preregistered to recruit 450 participants on MTurk and to exclude those who failed the attention check (https://aspredicted.org/6WD_YG1). Four hundred fifty-two participants completed the study and 13 failed the attention check, rendering a final sample of 439 (M_{age} = 38.9 years; 46.0% female, 53.5% male, and 0.5% selecting “other”). Participants were randomly assigned to one of three conditions.

Procedure. Each participant saw the low-chance set from Figure 1. Participants were randomly assigned to one of three conditions, which either asked them to indicate which number was most likely to be drawn (most-likely condition), make a single prediction of the outcome (prediction condition), or make two ranked predictions (ranked-prediction condition). In the prediction condition, participants would win \$0.50 if they successfully predicted the number they drew. In the ranked-prediction condition, participants were asked to give their best prediction and their second-best prediction. They were not allowed to give the same number for both predictions. If their best prediction matched the number they drew, they would win \$0.50. If their second-best prediction was right, they would win \$0.10. At the end of the survey, participants answered an attention check that asked them to recall how many balls were marked "1."

1.10.2 Results and Discussion

We first replicated the most-likely vs. prediction gap. Nearly everyone (93.5%) in the most-likely condition correctly indicated that 1 was most likely, but only 63.4% in the prediction condition predicted a 1 ($\chi^2(1) = 39.03$, $p < .001$). However, participants were much more likely to choose 1 as their best prediction when making ranked predictions. In the ranked-prediction condition, 83.1% chose 1 as their best prediction, a percentage that was reliably higher than the 63.4% who chose 1 as their single non-ranked prediction ($\chi^2(1) = 13.10$, $p < .001$). (A small minority of participants (7.0%) chose "1" as their second-best prediction.) Notably, still fewer people chose 1 as their best prediction than indicated 1 as most likely ($P_{\text{best-prediction}} = 83.1\%$ vs. $P_{\text{most-likely}} = 93.5\%$, $\chi^2(1) = 6.99$, $p = .008$), suggesting that the most-likely vs. prediction gap was not fully eliminated by ranking the predictions.

Study 7 thus suggests that the practice of ranking possible predictions and finding the best one might promote a focus on relative likelihood and encourage people to predict in line with relative likelihood, even when absolute likelihood is low.

1.11 Study 8: Envisioning 1000 Trials

Studies 6 and 7 narrowed the most-likely vs. prediction gap by increasing people's focus on the relative likelihood. Study 8 examines whether reducing their focus on the low absolute likelihood of the most likely outcome also reduces the most-likely vs. prediction gap. In study 8, we ask some people to first imagine the outcomes of 1000 repeated trials before predicting for a single trial. To understand our predictions, consider drawing from a set of nine balls where the number 1 appears twice. When predicting for one trial, the chance of drawing a 1 seems low. However, when people first imagine the outcomes of 1000 trials, drawing a 1 may no longer seem so unlikely because people will have just envisioned a 1 being drawn over 200 times. Thus, to the extent that people focus on absolute likelihood, the absolute likelihood may not seem as low following this manipulation as when people just consider one drawing in isolation. Similarly, envisioning 1000 trials may prompt people to take an outside view and focus less on the low absolute likelihood of 1 arising on any single trial (Kahneman & Lovallo, 1993).

Thus, we predicted that people's predictions would be more in line with what is most likely when they first consider a large number of repeated trials than when they do not.

1.11.1 Method

Participants. As preregistered, we recruited 300 MTurk workers (https://aspredicted.org/VBK_1LH). Ten of them did not pass the attention check, leaving us with a final sample of 290 (M_{age} = 43.6 years; 47.2% female).

Procedure. Participants completed two tasks. In one task, participants were given the low-chance set shown in Figure 1. Before drawing a ball, participants were asked to predict which number they would draw. In the other task, participants imagined that 1000 people were each given that set of balls, and those 1000 people each randomly drew a ball from the set. Participants were asked to imagine that the 1000 people were divided into eight groups based on the number that they drew, such as the group of people who drew a 1, the group who drew a 2, and so on. Participants estimated which group was the largest. The order of the two tasks was counterbalanced, so that some people predicted for a single trial before imagining 1000 people, whereas others imagined 1000 people before predicting for a single trial.

At the end of the study, all participants answered an attention check that asked them to recall how many balls had been marked “1.”

1.11.2 Results and Discussion

Almost everyone (92.4%) estimated that the group of people who drew a 1 was the largest group, and this percentage did not reliably differ between task orders ($P_{1000\text{-before-predict}}$= 93.8% vs. $P_{\text{predict-before-1000}}$ = 91.0%, $\chi^2(1) = .49$, $p = .485$).

Did envisioning 1000 draws bring predictions more in line with the most likely outcome? It did. Reliably more participants predicted that they would draw a 1 after, versus before, they imagined 1000 people drawing a ball ($P_{1000\text{-before-predict}}$ = 74.7% vs. $P_{\text{predict-before-1000}}$ =55.6%, $\chi^2(1) = 10.83$, $p < .001$). We further examined the predictions among the great majority who estimated that the group drawing a 1 was the largest (n = 268). The effect persisted. Although everyone in this subsample explicitly stated that 1 would come up most frequently among the 1000 draws, only

61.1% predicted a 1 before considering the 1000 draws, but 76.6% predicted a 1 after considering the 1000 draws ($\chi^2(1) = 6.89$, $p = .009$).

One may wonder whether the effect of imagining 1000 people arose because doing so mainly reminded people of the most likely outcome. Our previous studies suggest that a simple reminder is not enough to affect predictions. In studies 2a and 2b, people reported the most likely number immediately before making their predictions, and yet, the most-likely vs. prediction gap persisted—and was unaffected by whether predictions came before or after likelihood judgments. Such results suggest that simple reminders of the most likely outcome are not enough to improve predictions and that study 8's manipulation improved predictions by leading people to focus less on the low single-trial likelihood of that outcome.

1.12 Study 9: Giving Advice

So far, we have shown that people predict contrary to what they know to be the most likely outcome when the most likely outcome is unlikely to arise. In our final study, we examine the predictions people recommend to others. Will people also recommend a prediction contrary to what they know to be most likely?

Recall that, in studies 5a and 5b, people reported being less likely to make predictions based on logical factors when the most likely outcome was unlikely. However, the decisions people make on behalf of others are often less biased than the decisions they make for themselves (Andersson et al., 2016; Polman, 2012). Moreover, in the role of an advisor, people are more likely to focus on distributional information relevant to the overall utility of the population (Kray, 2000), such as, "what option would make most people better off?" Thus, we suggest that people will be less likely to depart from the accuracy-maximizing prediction and less likely to predict arbitrarily when advising others, compared to when predicting for themselves.

In addition, previous research has found that the act of giving advice to others can even make people less biased in their own decisions (e.g., Fantino & Esfandiari, 2002), perhaps because people feel hypocritical if they do not follow the advice that they give to others. Thus, we also examine whether giving advice can serve as a debiasing method that brings people's own predictions closer to their likelihood judgments: if people give logical advice to others, will their own predictions follow suit?

1.12.1 Method

Participants and design. Undergraduate students (N = 281; M_{age} = 19.5 years; 54.4% female) from a U.S. university participated for course credit. They were randomly assigned to one of two conditions: predict-first and advise-first. This study was preregistered (https://aspredicted.org/DP6_SPS).

Procedure. Participants played a computerized game in the lab. Each participant saw on the screen the low-chance set from Figure 1. Participants in the predict-first condition first predicted which number they would draw. They could win a small prize (a keychain, a lanyard, or a key-tag of their choice) if their prediction was accurate. After they made their own prediction, they were instructed to give advice to the participant next to them. To do this, they wrote on a piece of paper the number that they would advise their neighbor to predict. Participants in the advise-first condition first wrote advice to their neighbor. Then, they made a prediction for themselves with the same incentive as in the other condition. At this point, all participants had both made a prediction for themselves and had written advice to their neighbor. The experimenter then facilitated the exchanging of written advice.

After everyone received written advice from another participant, they were given a chance to revise their prediction. Then, they were asked to indicate which number was most likely to be drawn. Finally, they clicked to draw a ball.

1.12.2 Results and Discussion

As before, most participants (93.2%) indicated that they were most likely to draw a 1, and this percentage was not affected by question order ($P_{predict\text{-}first\text{-}condition}$= 91.2% vs. $P_{advise\text{-}first\text{-}condition}$ = 95.1%, $\chi^2(1) = 1.13$, $p = .288$).

Although nearly everyone knew that 1 was most likely to arise, their advice differed strikingly from their predictions. To make a clean comparison between advice and predictions, we compared predictions in the predict-first condition to advice in the advise-first condition, as preregistered. Reliably more participants advised others to predict a 1 than predicted a 1 for themselves (P_{advise} = 89.6% vs. $P_{predict}$ = 58.4%, $\chi^2(1) = 34.22$, $p < .001$). In the predict-first condition, there was a large gap between people's predictions and their indicated most likely number ($P_{most\text{-}likely}$ = 91.2% vs. $P_{predict}$ = 58.4%, $\chi^2(1) = 37.50$, $p < .001$). However, in the advise-first condition, the gap between people's advice and their most likely number was much smaller and was not statistically reliable ($P_{most\text{-}likely}$ = 95.1% vs. P_{advise} = 89.6%, $\chi^2(1) = 2.41$, $p = .120$): few participants did not a recommend a 1 if they knew 1 was most likely. Notably, among the participants who correctly indicated 1 as most likely, 37.6% in the predict-first condition did not predict a 1 whereas only 7.6% in the advice-first condition did not recommend a 1 ($\chi^2(1) = 33.50$,

$p < .001$). Thus, although people predicted contrary to their likelihood judgment, their advice to others was much more in line with what they knew was most likely.

Did giving advice improve advisors' own predictions? It did. Significantly more people predicted a 1 after versus before advising others ($P_{\text{advise-first-condition}} = 77.8\%$ vs. $P_{\text{predict-first-condition}} = 58.4\%$, $\chi^2(1) = 11.31$, $p < .001$). The gap between participants' predictions and their reported most likely number was narrower—although not eliminated—for participants who predicted after giving advice ($P_{\text{most-likely}} = 95.1\%$ vs. $P_{\text{predict}} = 77.8\%$, $\chi^2(1) = 17.08$, $p < .001$). Among those who indicated 1 as most likely, 37.6% did not predict a 1 in the predict-first condition whereas only 19.7% did not predict a 1 in the advice-first condition ($\chi^2(1) = 9.46$, $p = .002$). Thus, giving advice brought people's predictions closer to what they knew to be most likely.

Finally, we examined people's revised predictions. Overall, only 11.7% of the participants revised their predictions after receiving the advice, and this percentage did not differ between conditions ($P_{\text{advise-first-condition}} = 11.8\%$ vs. $P_{\text{predict-first-condition}} = 11.7\%$, $\chi^2(1) < .001$, $p > .999$). Appendix F contains additional analyses.

Study 9 demonstrates two things. First, it shows a boundary of the disconnect between prediction and judgment: Although people may not predict what they believe to be most likely, their recommended predictions to others are much more in line with their perceived most likely outcome. Second, it shows that giving advice can be a debiasing method that encourages advice-givers to subsequently predict more in line with their perceived most likely outcome.

1.13 General Discussion

When predicting the outcome of an uncertain event, if one wants to maximize the chances of predicting accurately, one should predict whichever outcome one believes is most likely to arise. However, we show that people's predictions can disagree with their own likelihood judgments. Although people regularly predict their perceived most likely outcome when they think the most likely outcome is overall very likely to arise, they less regularly predict that outcome when it is overall unlikely to arise—even though they still believe that outcome to be most likely. Studies 1 through 4 documented this basic pattern.

This disconnect between prediction and likelihood judgment suggests that people consider not only the relative likelihood (i.e., which outcome is more likely to arise than others?) but also the absolute likelihood (i.e., how likely is this outcome to arise overall?). We argue that when people think that the absolute likelihood of the most likely outcome is very low, they consider the eventual outcome to be rather difficult to foresee, and that this feeling of low foreseeability in turn promotes arbitrary prediction strategies that lead predictions to depart from the perceived most likely outcome. Studies 5a and 5b supported this hypothesis.

It follows that one can encourage people to predict more in line with their perceived most likely outcome by redirecting their focus back to relative likelihood or reducing their focus on the low absolute likelihood. Studies 6 through 8 suggest that this is the case. Nonetheless, although people's predictions tend to diverge from what they believe to be most likely to arise, Study 9 shows that their advice to others is more in line with their believed most likely outcome.

1.13.1 Relation to Previous Research

Previous research on prediction and subjective probability mostly focuses on how people's predictions and judgments depart from formal probability models. As discussed, a long line of research has shown that human predictions and probability judgments can be biased by many different factors, such as heuristics (Tversky & Kahneman, 1974), affect (e.g., Loewenstein et al., 2001), optimism (e.g., Weinstein, 1980), and so on. This body of research usually does not examine the correspondence between people's predictions and their likelihood judgments. Instead, it often reasonably assumes that people's predictions follow from their subjective judgments of likelihood. In contrast, the current research examines the correspondence between people's predictions and their likelihood judgments, putting aside whether those predictions or judgments are biased compared to formal models.

As discussed, previous research has documented a few cases of discrepancies between likelihood judgments and predictions, including cases related to desirability bias (e.g., Park et al., 2022) and probability matching (i.e., Koehler & James, 2009). Researchers have also shown a mismatch between likelihood judgments and choice caused by the ratio bias (e.g., Denes-Raj & Epstein, 1994). The current research adds to the literature by identifying another, arguably even more pervasive, factor that causes a discrepancy between predictions and likelihood judgments: a low absolute likelihood of the most likely outcome. Because the absolute likelihood of the most likely outcome is a basic and inherent property of an uncertain event, prediction distortions caused by it may arise frequently.

This new factor makes predictions that differ from the previous research. First, whereas research on the desirability bias found that prediction and likelihood judgment diverged when one outcome was particularly desirable (e.g., because participants would win money if that outcome obtained), the low absolute likelihood effect does not hinge upon one outcome being more desirable than the others. Rather, our effect might arise whenever people focus on the low absolute likelihood of the most likely outcome, regardless of the desirability of any other outcome.

Second, whereas probability matching is most relevant when people predict a class of events, the low absolute likelihood effect can arise when people predict a single event. Would probability matching predict a similar discrepancy for a single prediction? Not necessarily. Imagine that an individual predicts the outcomes of N repeated draws with a 70% chance of red on each draw and a 30% chance of black. Probability matching would hold that people would predict red for 70% of the draws and black for 30%, despite knowing that red was more likely on each draw. If N = 1,000, they would predict red for 700 draws and black for 300. If N = 10, they would predict red for 7 draws and black for 3. If N = 1 (that is, when they only need to predict a single draw), they again should be more likely to predict red than black, in line with their likelihood judgment. Thus, probability matching would not easily account for the effects seen here, which emerge on a single trial. That said, we acknowledge that there is some similarity between the two types of effects, and we would welcome research that further investigated commonalities between them.

Finally, the low absolute likelihood effect differs from what the ratio bias would predict. In our paradigm, people choose among a set of possible outcomes as their prediction. The most likely outcome from the set has both the highest likelihood and the greatest frequency (i.e., the greatest numerator of a ratio). Therefore, even people showing a ratio bias would still predict the most likely outcome, as predicting it would give them the most chances to win.

1.13.2 Future Directions

On the Absolute Likelihood. What makes an outcome seem likely or unlikely? For most studies in this article, we manipulated the overall likelihood of the most likely outcome by setting the low likelihood near 20% and the high likelihood near 70%. Prior research suggests that people generally perceive such likelihoods as "unlikely" and "likely," respectively (Budescu & Wallsten, 1995; Clark, 1990; Sirota & Juanchich, 2015; Theil, 2002). We also examined more natural settings where people reported their own belief about the absolute likelihood of their perceived most likely outcome, and where that likelihood varied across a wider setting (Studies 3 and 4). Future research could more systematically examine the most-likely vs. prediction discrepancy at different levels of likelihood. Moreover, a certain level of likelihood could seem low or high in different contexts, and so future research could explore the low absolute likelihood effect by using framing or other contextual manipulations to affect whether a given level of likelihood (e.g., 40%) seems high or low.

On Variants of Uncertainty. Previous research has distinguished two types of uncertainty, an internal uncertainty that is epistemic and attributed to a lack of knowledge or information and an external uncertainty that is aleatory and attributed to the properties of the environment (Kahneman and Tversky, 1982; Ulkumen and Fox, 2011). In most studies except for Studies 3 and 4, we displayed all possible outcomes to participants, so the uncertainty involved in those studies was external and aleatory. In Studies 3 and 4, when participants predicted the outcome of a basketball tournament, many might not have complete information or expertise about the tournament. Therefore, the uncertainty in those studies may have been relatively internal and epistemic. In all studies, however, we consistently observed the most-likely vs. prediction discrepancy. Thus, the disconnect between prediction and likelihood judgment seems to arise regardless of whether the uncertainty is more aleatory or epistemic. That being said, we have not systematically gauged how differences between the two types of uncertainty could potentially affect the disconnect between prediction and likelihood judgment. Future research could provide a more thorough understanding on this front.

On Larger Incentives. In Studies 1, 2b, 3, 4, 7, and 9, participants were given a monetary or tangible incentive for accurate predictions, and yet the most-likely vs. prediction discrepancy persisted in reliably large magnitudes. However, these incentives were not large in value, and one might argue that people might not have been adequately motivated to make accuracy-maximizing predictions (that is, the downside of an incorrect prediction was not large). What would happen if the stakes were higher? One might argue that predictions would be more accurate, but one could argue the opposite. When low-chance outcomes are associated with very high stakes, an accurate prediction could feel even more like a matter of luck, and correspondingly, people could be even more drawn to an arbitrary strategy that relies on a lucky number or gut feeling. Future research could explore the potential effects of larger incentives on the low absolute likelihood effect.

On Uniqueness. Could people sometimes predict an outcome that is not the most likely outcome because doing so is fun, exciting, or makes the predictor feel unique? This is certainly possible, but we note that predictions and most likely judgments only diverge when absolute likelihood is low, not high. If our effects were driven purely by fun-seeking or a uniqueness motive, one might expect the effects to also appear in the high-absolute-likelihood conditions, where diverging from the most-likely outcome might even be more potentially exciting and might convey uniqueness especially well. That said, choosing an option to feel unique could, broadly speaking, be a type of non-logical, arbitrary prediction strategy, much like choosing an option one likes or going with a gut feeling. Thus, finding that predictions are influenced by uniqueness in low absolute likelihood settings would not be inconsistent with our account. Making a unique choice is an important motive, especially for people from Western societies (Kim & Markus, 1999; Markus & Kitayama, 1991), and so future research could explore the role of this motive.

1.13.3 Implications

Prediction is everywhere. Voters predict the winner of an election; sports fans predict game outcomes; policy makers predict which alternative policy is most efficient, and so on. There is also a large, growing prediction market of sportsbooks, casinos, and online prediction and gambling platforms. Correspondingly, much research has investigated human predictions and likelihood judgments. Researchers often assume, either explicitly or implicitly, that people's predictions follow from their subjective likelihood judgments: if their likelihood judgments are biased, their predictions will follow suit, but if their likelihood judgments are optimal, so too will be their predictions. However, we show that predictions can easily depart from likelihood judgments. This discrepancy calls for research attention to the (non)correspondence between an individual's prediction and their own subjective probability, and suggests that even when people assess outcome probabilities correctly, their predictions might still not be optimal.

Chapter 2: A Co-Branding Conundrum

2.1 Introduction

Imagine that you are about to book a luxurious seven-night Caribbean cruise. When it is time to check out, you pull out your wallet and see your two credit cards, a Chase Visa and a Walmart Rewards Mastercard. Would you think that the Walmart-branded credit card is the right card to use for a luxury cruise? What if the Walmart-branded card pays more cash back than the Chase Visa for this trip? In this research, we propose that featured brands on co-branded credit cards can discourage consumers from using those cards when the purchase is unrelated to the featured brand, even when a co-branded card offers the best rewards.

Co-branded credit cards such as the Walmart Rewards Mastercard are the product of a partnership between a credit card issuer (e.g., Capital One) and a merchant brand (e.g., Walmart), and are backed by a major payment network (e.g., Visa, Mastercard, or American Express)[5]. These cards are popular. About 29% of U.S. adults (or 73.7 million people) hold co-branded credit cards (Packaged Facts, 2021). Despite this popularity, we do not know a lot about how consumers use their co-branded credit cards. Adopting a co-branded credit card does appear to encourage spending on products from the card's featured brand (Zhao, Gopalakrishnan, & Narasimhan, 2022; see also Blackett & Russell, 2000). However, a potentially even more interesting question centers on how co-branded credit cards influence purchases *outside* the cards' featured brands.

[5] The co-branded credit cards discussed in this paper are open-loop credit cards that can be used anywhere that accepts their payment networks (such as Visa, Mastercard, or American Express), in contrast to closed-loop store credit cards that can only be used in specific stores.

Co-branded credit cards typically offer special rewards for brand-specific purchases, but they also commonly offer rewards for other purchases outside of the featured brand. The rewards outside of the featured brand are frequently comparable to the rewards from non-co-branded cards. For example, Table 1 shows the reward structures of three common co-branded credit cards, all of which offer competitive rewards for purchases outside of their featured brand. Do people use their co-branded credit cards as broadly as they should to maximize rewards, and if not, why not? The current research investigates these questions.

Table 2.1 Examples of Credit Card Rewards

	Reward Structure
Some popular co-branded cards:	
BPme Rewards Visa	• 5% cash back on non-fuel purchases and 15¢ off per gallon of fuel at BP and Amoco • 3% cash back on dining • 3% cash back on groceries • 1% cash back on other qualifying purchases
Walmart Rewards Mastercard	• 5% cash back at Walmart.com • 2% cash back in Walmart stores • 2% cash back at restaurants • 2% cash back on travel • 1% cash back everywhere else Mastercard is accepted
Amazon Rewards Visa Signature Card	• 3% cash back (and 5% back for Amazon prime members) at Amazon.com and Whole Foods Market • 2% cash back at restaurants • 2% cash back at gas stations • 2% cash back at drug stores • 1% cash back on all other purchases where Visa is accepted
A popular non-co-branded card:	
Blue Cash Everyday Card from American Express	• 3% cash back at U.S. supermarkets • 3% cash back on online retail purchases • 3% cash back at U.S. gas stations • 1% cash back on other eligible purchases

Note: The reward information was collected from each card's official website in May 2023.

2.1.1 Credit Cards and Decision Making

When busy and distracted consumers carry multiple credit cards, each with their own terms and conditions, departures from financially optimal decisions are bound to occur. For instance, when consumers cannot completely repay their debts on several cards, they rely on imperfect heuristics to guide their repayment decisions, causing them to leave money on the table (e.g., Amar et al., 2011; Gathergood et al., 2019; Kettle et al., 2016). Likewise, when choosing a card to use in the first place, people do not consistently prioritize the cards' different interest rates, causing them to incur excessive interest charges (Ponce, Seira, & Zamarripa, 2017).

Yet, credit card interest rates, the focus of most investigations about consumers and credit card use, are not particularly relevant to the many credit card users who completely repay their credit card debts each month. A 2023 American Bankers Association report indicates that about 44% of all active U.S. credit card accounts did not carry a revolving balance in the third quarter of 2022.[6] Many credit users are able to completely repay their debts each billing cycle. For these users, the best card to use for any particular purchase is largely determined by their cards' rewards structures.

[6] This percentage is derived from the American Bankers Association's reported statistics. According to the report, of all U.S. credit card accounts in the third quarter of 2022, 43% were "revolvers" (building debt but not completely repaying it), 34% were "transactors" (using their card and completely repaying their debt at each billing cycle), and 23% were "dormants" (posting no activities, including interest charges, in that quarter). The percentage of active users who did not carry a revolving balance is 33.6% divided by the sum of 43% and 34%.

2.1.1 Two Psychological Effects of Co-Branding

In this research, we propose that consumers are reluctant to use co-branded credit cards outside of a card's featured brand for at least two reasons discussed below: (1) co-branding produces assumptions about the card's reward structure, and those assumptions limit effortful attention to the card's actual reward structure, and (2) co-branding makes many potential purchases feel like a bad "fit" with the card.

Assumptions Direct Attention. When a credit card features a merchant brand, users might reasonably assume that the card's reward structure is designed to promote purchases from that brand, meaning that its major rewards will be brand-specific perks. By contrast, rewards outside the featured brand might be an afterthought, and users might assume that those rewards are minimal (thinking, for example, "Why would my Macy's American Express card incentivize purchases at places such as restaurants and gas stations?"). These assumptions may be bolstered when possible users encounter marketing materials for co-branded cards, which generally shine a spotlight on rewards for purchases from the featured brand. For example, on the Sam's Club Mastercard website[7] (Figure 1), rewards on Sam's Club purchases are displayed prominently, whereas rewards on other purchases are in smaller print at the bottom, including 5% back on gas, 3% back on dining, and 1% back on other purchases. This stark visual contrast likely bolsters the assumption that co-branded credit cards are overwhelmingly focused on incentivizing purchases from the focal brand

[7] https://web.archive.org/web/20230608043936/https://www.samsclub.com/content/credit

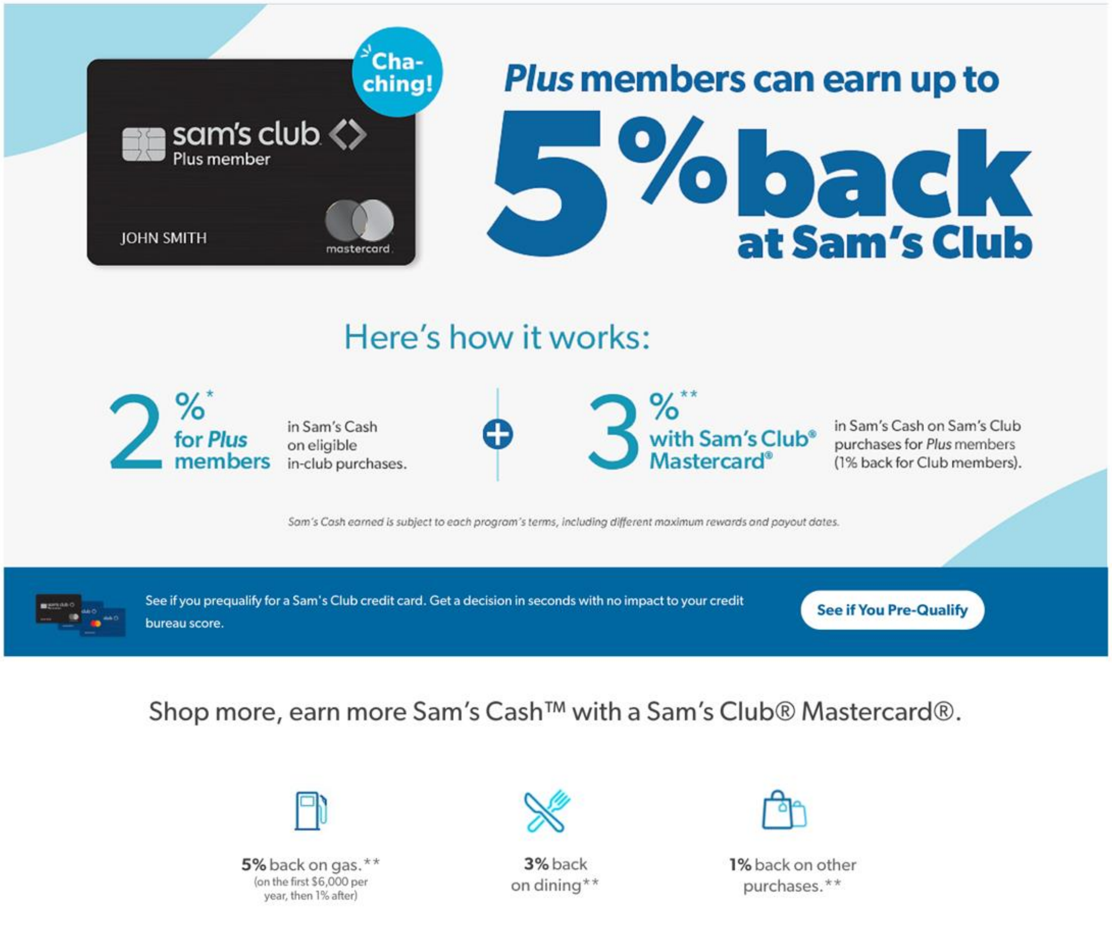

Figure 2.1

Such assumptions could discourage users from paying full attention to reward structures because humans tend to act as "cognitive misers," who seek to conserve scarce cognitive resources whenever possible (e.g., Fiske & Taylor, 1991; Shah & Oppenheimer, 2008). Much like stereotypes are often used as "energy-saving devices" (Macrae, Milne, & Bodenhausen, 1994), strong assumptions about a credit card's reward structure might spare users from having to carefully check the details of that reward structure. As the philosopher David Hull (2001, p. 37) has argued, "the rule that human beings seem to follow is to engage the brain only when all else fails – and usually not even then."

The inattention to reward details that results from these assumptions is unlikely to affect co-branded card usage within the featured brand. Users will assume that co-branded cards strongly incentivize purchases from the featured brand, and they will usually be correct. However, if users also assume that purchases outside the featured brand are less rewarding than they actually are, they are likely to under-use their co-branded card for purchases outside the featured brand.

Perceived Card-Purchase Fit. Consumers are known to rely on categorical cues to decide which purchases they should make with certain types of payments or funds. For example, Kooreman (2000) shows that Dutch parents spend more on their children when they receive child benefits as cash from the government than when they receive other equivalent sources of income. The "child benefit" label helps categorize the otherwise fungible money as a specific fund to spend on their children. Even more closely related to the current research, Reinholtz, Bartels, and Parker (2015) find that consumers use the retail brand of a gift card to determine which purchases they should make with the gift card: people endowed with a brand-specific gift card are more likely to purchase products that are consistent with the brand image than are people endowed with a general-use gift card.

In the current research, we propose that the presence of a brand on a credit card that can be used across retailers and for virtually any type of purchase might actually influence whether people use that credit card or not. Specifically, when a credit card features a merchant brand, people might mentally designate the card for purchases from that featured brand. When they have the option to use it outside the featured brand, they may perceive a lack of fit, or inconsistency, between the featured brand and that purchase. For example, a printer may not feel like it fits with the outdoor brand, REI, when a consumer has an option to use their REI Mastercard.

Inconsistent information is less easy to mentally process than consistent information (Winkielman et al., 2012). Accordingly, it may be less easy for consumers to accept the idea of using a credit card that features an outdoor brand to buy office supplies. Moreover, research suggests that people attend to their metacognitive feelings about how easy a stimulus (e.g., a purchase) is to process as a source of information to evaluate the stimulus (for a review, see Schwarz et al., 2021). Consistent and easy-to-process stimuli are evaluated more positively than inconsistent and difficult-to-process stimuli (Winkielman et al., 2003).

Indeed, in consumer research, it has been widely documented that a lack of fit leads to negative evaluations. For example, consumers react less positively when the face value of a currency does not approximately match the price of an intended purchase (i.e., spending-denomination fit; Li & Pandelaere, 2021). That is, they do not feel right about using bills in large denominations to buy inexpensive products or vice versa.

Brand-specific fit can influence consumer judgment, too. For example, consumers react less favorably to brand extensions when the extension products are considered a low (vs. high) fit with the brand concept (Aaker & Keller, 1990; Park, Milberg, & Lawson, 1991). Research in cause marketing has found that brand-conscious consumers respond less favorably to a cause marketing message when there is a low, versus high, fit between the brand and the social cause (Nan & Heo, 2007). For instance, consumers have less favorable attitudes towards Johnson & Johnson when the company advertises that its purchases support Feeding America versus the American Red Cross; Johnson and Johnson, a pharmaceutical company, is viewed as a worse fit with the mission of Feeding America than with the values of the American Red Cross.

We similarly suggest that consumers might react less positively when they perceive a low, versus high, fit between an intended purchase and a credit card they can use. Specifically, when a consumer has the option to use a co-branded credit card to pay for a purchase outside of its featured brand, that incongruence is experienced as a lack of card-purchase fit. Consumers then might not “feel right” about the incongruence and thus avoid using the co-branded card, even when that co-branded card maximizes the reward for the intended purchase.

In summary, credit card co-branding may exert two unique effects that discourage people from using their co-branded credit cards for purchases outside of the partnering brands. We predict that:

> *H_1:* Co-branding a credit card reduces the likelihood that consumers will use that credit card for purchases outside of the partnering brand, even when the co-branded card is the reward-maximizing option.

Further, because co-branding a credit card might produce assumptions about its reward structure, we expect that consumers will pay less attention to a credit card’s reward structure when the credit card features a merchant brand.

> *H_{2a}:* Co-branding a credit card reduces consumers’ attention to the card’s reward structure.

As a result, partial inattention reduces consumers’ awareness of the card’s rewards outside of its featured brand, which in turn restricts the card use outside of the brand. That is:

H_{2b}: People's attention to a credit card's reward structure and their subsequent awareness of its reward structure serially mediate the effect of co-branding on credit-card usage outside of the featured brand.

We have also suggested that consumers avoid using a co-branded credit card outside of the featured brand even in cases where they are fully aware of the reward structure because it still does not "feel right" when the credit card and the purchase do not seem like a good fit. That is:

H_3: The perceived fit between a co-branded credit card and an intended purchase, and the subsequent evaluative reaction of "feeling right" (or not), serially mediate the effect of co-branding on co-branded credit card usage outside of its featured brand.

Such perceptions of fit and the subjective feeling of "feeling right" do not hinge on people's awareness of the card's rewards, so we expect this mediation path to uniquely explain the underuse of co-branded credit cards even when controlling for awareness of credit card rewards.

In what follows, we first present a pilot survey that provides descriptive insights into how consumers use their co-branded credit cards. Next, we describe four experimental studies that formally test our hypotheses.

Study 1 serves two purposes. First, it tests the main hypothesis that consumers are reluctant to use a co-branded credit card outside of its featured brand even when the co-branded card is the reward-maximizing option (H_1). Second, it provides initial evidence for both proposed mechanisms. Study 2 further examines the importance of co-branding in consumers' decision to use co-branded credit cards, including a novel condition in which a card has the same reward terms as its co-branded counterpart, but is not co-branded. Study 3 relies on a mouse-tracking technique to test the effect of co-branding on attention (H_{2a}). Study 3 also formally examines our proposed process in which people's attention to, and subsequent awareness of, credit card rewards explains part of consumers' reluctance to use co-branded credit cards outside of their featured brands (H_{2b}). Finally, Study 4 examines the second process via perceived card-purchase fit (H_3). All of our study materials, data, codes, and preregistration documents are available at: https://osf.io/frmz2/?view_only=c5527906c1fe449db06c370e8915eb3f

2.2 Pilot Survey

To get a sense of how people use their co-branded credit cards, we surveyed U.S. adults via Prolific.co about their credit card usage. We aimed to survey 200 U.S. adults who possessed at least one co-branded credit card and one other credit card that was not co-branded. To find qualified participants, we prescreened 700 people based on the number of co-branded and non-co-branded cards they held. Two hundred seventy-three participants possessed at least one co-branded card and one non-co-branded card and thus were invited to participate in the pilot survey, which allowed a maximum of 200 participants. Participation was on a first-come first-serve basis.

Across four days, 201 invited participants completed the survey (M_{age} = 40.5 years; 53.7% female, 44.8% male, and 1% non-binary or preferring not to say). Participants indicated the full names of the one co-branded credit card and the one non-co-branded credit card that they used most often. Based on their reported credit card names, we were able to identify most reported credit cards and verify their reward structures. Specifically, we successfully identified 200 of the

201 reported non-co-branded cards and 195 of the 200 reported co-branded cards.[8] We verified and recorded the reward terms and APR[9] of every identified card from their official websites. Our results suggested that the co-branded cards that participants possessed had similar APRs ($M_{non\text{-}co\text{-}branded}$ = 24.5%, $M_{co\text{-}branded}$ = 24.8%, $t(360.7) = .9$, $p = .382$) and offered competitive rewards outside their featured brands, as compared to their non-co-branded cards. Specifically, 99% (vs. 88%) of their co-branded (vs. non-co-branded) credit cards offered a base reward for general purchases that were not tied to any particular brand or product category, which was typically 1% to 1.5% cash back or 1x to 1.5x reward points per dollar spent. In addition, 89% (vs. 70%) of their co-branded (vs. non-co-branded) credit cards offered at least one special perk category that was not brand-specific, including several common categories of everyday spending (Table 2.2).

Table 2.2 Credit Card Reward Summary

	Percentage (%) of participants' credit cards that offer each type of reward	
	Non-co-branded cards (N = 200)	Co-branded cards (N = 195)
Base reward for general spending	88	99
Brand-specific perks	0	98
Special perk(s) for certain category(s) (not brand-specific)	70	89

Notes: The reward information was collected in February 2024.

[8] One non-co-branded card entry was excluded because it was co-branded. Six co-branded card entries were excluded, including four that were store credit cards that could not be used outside the stores and another two that were not co-branded.
[9] For credit cards that had more than one possible APR %, we took the average of all possible APR %'s as an estimate.

The similar reward structures and APRs between the two types of credit cards suggested that participants could use their co-branded credit cards as broadly as their non-co-branded credit cards. However, they did not. Participants indicated how they tended to use their co-branded and non-co-branded credit cards on a scale from 1 (only for specific purposes or at specific stores) to 10 (everywhere). As shown in Figure 2.2, there was a stark difference in how they used the two types of card. They used their co-branded credit cards significantly more restrictively than their non-co-branded credit cards ($M_{\text{non-co-branded}} = 7.2$, $M_{\text{co-branded}} = 3.8$, $t(400) = 11.5$, $p < .001$). They further estimated the percentage of their monthly credit card spending that was done with each of the two cards. Consistently, on average, 51.1% of their monthly credit card spending that was done with their non-co-branded card but only 34.2% with their co-branded card ($t(399.5) = 5.5$, $p < .001$).

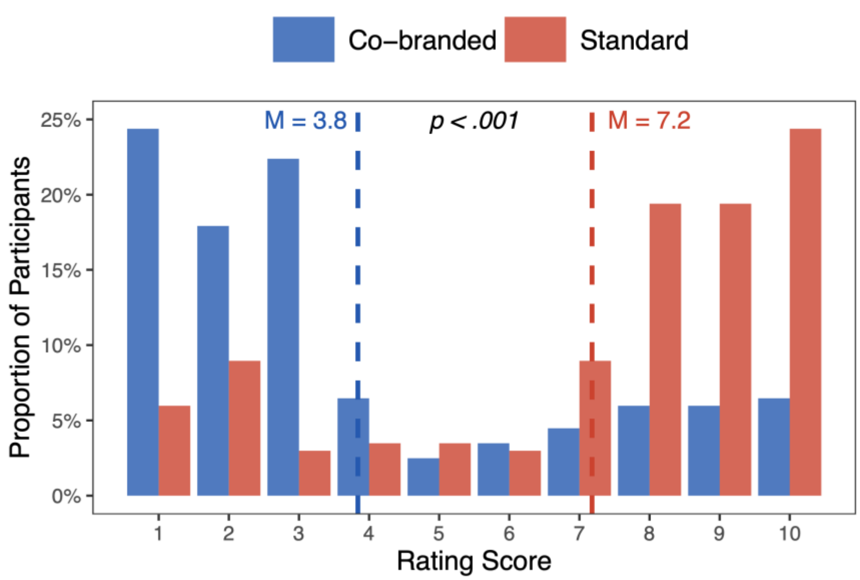

Figure 2.2

Taken together, these patterns provide suggestive evidence that consumers use co-branded cards less than non-co-branded cards, even when both could reasonably be used to take advantage of credit card rewards. Of course, these survey results are purely descriptive and as mentioned earlier, we did not have full access to these participants' full credit card terms (e.g., interest rates, which can vary within card and by user over time) and balances. In what follows, we propose two mechanisms and then describe four experimental studies that formally examine consumers' decisions to use co-branded credit cards. Further, we show that limited use of co-branded credit cards persists even when using a co-branded credit card is clearly the reward maximizing option.

2.3 Study 1

In Study 1, we test our main hypothesis that consumers are less likely to use a credit card for a purchase when the card features a merchant brand unrelated to the purchase, even if the co-branded card maximizes the reward on that purchase (H_1).

Moreover, we seek initial evidence for our proposed mechanisms by testing two more predictions. First, as part of hypothesis H_{2b}, we predict that consumers will have less awareness of the details of a co-branded card's reward structure because they hold pre-existing assumptions about that reward structure and they do not exert substantial effort to confirm or disconfirm those assumptions. Second, because we also propose that the perceived card-purchase fit uniquely explains some of the underuse of co-branded credit cards independent of awareness, we expect that people still underuse a co-branded credit card even if they are fully aware of its rewards outside of the featured brand. Here, we test both predictions.

2.3.1 Method

As preregistered (https://aspredicted.org/CQ3_YFG), we recruited 400 CloudResearch-approved participants from Amazon Mechanical Turk (MTurk). Participants were randomly assigned to either a Control condition or a Co-branded condition. In both conditions, participants imagined possessing two credit cards, and the terms of the two cards were displayed side by side (see Figure 2). In the Control condition, participants imagined that they carried a Chase Visa and a Bank of America (BOA) Visa. These two credit cards had the same terms (the same credit line, current balance, and APR) except for their cash back rewards. The Chase Visa paid 3% cash back at gas stations and restaurants and 1% on all other purchases. The BOA Visa paid 3% cash back on online retail purchases and 2% on all other purchases. In the Co-branded condition, participants imagined possessing the same Chase Visa and a Best Buy Visa. The Best Buy Visa had was identical to the BOA Visa from the Control condition except for one item in its reward structure. The Best Buy Visa paid 3% cash back at Best Buy instead of on online purchases. All other attributes were identical to the BOA Visa from the Control condition including the receipt of 2% cash back on all other purchases.

After learning about their available credit cards, participants proceeded to the next screen where they imagined that they were shopping for groceries. Once participants were ready to check out, they saw their total purchase amount and chose one of their two credit cards to use to make the payment. At this point, participants could click to open a link to review their credit card terms again if they wished to do so before they made their decision.

After choosing a card to pay with, we asked participants which of the two cards offered more cash back on groceries to measure participants' awareness of the cards' rewards structure. Participants could select one of four options: the Chase Visa, the [Bank of America / Best Buy] Visa, "They both gave the same % of cash back," or "I don't remember." As a more general attention check, participants were then asked to indicate the two credit cards that they had in the grocery store scenario. Participants in both conditions could select one of three options: "Chase

Visa and Best Buy Visa," "Chase Visa and Bank of America Visa," or "American Express and Citi Mastercard". We preregistered a plan to exclude participants who did not pass this general attention check. Nine participants failed, leaving us a final sample of 391 participants (M_{age} = 40.0 years; 55% female, 45% male, and 1% indicating "Other").

	Chase Visa	**Bank of America Visa**
Credit Line	$8,000	$8,000
Current Balance (After most recent payment)	$0	$0
Interest Rate (APR)	19%	19%
Cashback Rewards	• 3% at gas stations and restaurants. • 1% on all other purchases.	• 3% on online retail purchases. • 2% on all other purchases.

	Chase Visa	**Best Buy Visa**
Credit Line	$8,000	$8,000
Current Balance (After most recent payment)	$0	$0
Interest Rate (APR)	19%	19%
Cashback Rewards	• 3% at gas stations and restaurants. • 1% on all other purchases.	• 3% at Best Buy . • 2% on all other purchases

Figure 2.3 Study 1 credit card terms in the Control (upper) and Co-branded (lower) conditions

2.3.2 Results and Discussion

In the Control condition, almost all participants (90%) chose to use their reward-maximizing card (the BOA Visa). However, in the Co-branded condition, a significantly lower percentage (60%) chose to use their reward-maximizing card (the Best Buy Visa; $\chi^2(1) = 46.7$, $p < .001$). In other words, when participants' reward-maximizing card featured a retail brand that did not fit with a prospective purchase, participants were less likely to use that card.

Were participants in both conditions equally aware of which card paid more cash back on groceries? They were not. As predicted, fewer participants in the Co-branded condition (vs. Control condition) correctly recalled which card paid more cash back on groceries (67% vs. 84%, $\chi^2(1) = 13.9$, $p < .001$). Moreover, as expected, participants were less likely to use the optimal card if they did not recall its rewards correctly: overall, only 25% of those who failed to recall the reward-maximizing card chose to use it, whereas 91% of those who correctly recalled the reward-maximizing card chose to use it ($\chi^2(1) = 162.6$, $p < .001$). This pattern holds true in both the Co-branded condition (14% vs. 82%, $\chi^2(1) = 81.3$, $p < .001$) and the Control condition (48% vs. 98%, $\chi^2(1) = 67.1$, $p < .001$). We more closely examine the relationship among people's attention to credit card rewards, their awareness of it, and their card choice in Study 3.

We additionally found that fewer participants in the Co-branded condition (vs. the Control condition) clicked to review the credit card terms (16% vs. 28%, $\chi^2(1) = 5.6$, $p = .018$). This pattern corresponds to our proposition that co-branded credit cards discourage people from checking their rewards because they assume they already generally know what the reward structures look like based on their featured brands. As a result, the lack of double-checking in the co-branded condition could potentially contribute to the difference in recalling the reward-maximizing card between conditions condition. That being said, the difference in click rate was relatively small (12 percentage points) and was unable to fully account for the larger difference in the percentage of correctly recalling which card paid more cash back on groceries (30.6 percentage point). To verify this, we ran a logistic regression predicting the likelihood of correctly recalling the reward advantage with a condition indicator (0 = Control condition; 1 = Co-branded condition) and a click indicator (0 = link unclicked; 1 = link clicked). There was a significant effect of the click ($\beta = 1.35$, SE = .42, $p = .001$), suggesting that participants who clicked to review the credit card information were indeed more likely to correctly recall which card paid more cash back on groceries. More important, however, there was also a significant effect of condition ($\beta = -.86$, SE = .25, $p < .001$). Thus, controlling for whether participants clicked on the link to review the credit card terms, we still found that participants were less successful in recalling which card paid more cash back on groceries when the better card features the Best Buy brand. Overall, the evidence from both participants' recall performance and their click patterns provide initial support for our hypothesis that people are less attentive to, and thus less aware of, the full rewards of a co-branded credit card compared to its non-co-branded counterpart.

Next, as a pre-registered exploratory analysis, we examined only participants who correctly recalled which card paid more cash back on groceries (n = 295). Within this group, fewer participants in the Co-branded condition (vs. the Control condition) chose to pay with the reward-maximizing card (82% vs. 98%, $\chi^2(1) = 21.1$, $p < .001$). Thus, even when they knew that the Best Buy Visa paid more cash back, participants still tended to underuse the Best Buy Visa for an unrelated purchase. In other words, participants underused the Best Buy Visa not only because they were unaware of its rewards outside of Best Buy, consistent with our hypothesis that another factor—the perceived card-purchase fit—can also discourage people from using the Best Buy Visa for groceries.

In an additional study reported in Appendix G, we replicate this finding with a more rigorous recall test that required participants to report the exact percentage of cash back that each credit card offered for a grocery purchase (i.e., using fill-in-the-blank questions, rather than multiple choice). Within the group who reported the rewards correctly, we still found that participants were less likely to use the Best Buy Visa card for grocery purchases relative to a non-co-branded card that offered the same rewards for groceries.

To summarize, Study 1 and its replication study provide evidence that consumers are less likely to use a co-branded credit card relative to a non-co-branded credit card for purchases unrelated to the co-branded card's retail brand, even when it is reward-maximizing to do so. We further show that this reluctance to use co-branded credit cards persists even when people have the correct reward information in mind.

2.4 Study 2

Study 1 demonstrated that, once a credit card features a merchant brand, people tend to underuse it for purchases that do not fit with the featured brand. We have generally attributed this co-branding effect to the presence of the featured brand. To feature a merchant brand, a co-branded credit card not only bears the brand name and other branding elements, such as the brand logo and brand-themed colors, but also features a special reward term for an exclusive category of purchases from the featured brand. How important are the brand name and other branding elements specifically in causing the underuse of co-branded credit cards? In other words, can an exclusive reward category alone, without brand elements, reduce card use outside of the exclusive category? To answer these questions, we separate the potential effect of an exclusive reward category from the effect of branding in Study 2.

More specifically, in Study 2, we have similar Control and Co-branded conditions as before. The reward-maximizing card featured the fuel brand, Shell, in the Co-branded condition and offered a special reward, namely 5% cash back, for gasoline purchases at Shell stations. By contrast, in the Control condition, the reward-maximizing card featured 5% cash back at department stores. Despite the difference in their 5% cash-back categories, both cards pay more cash back (2%) than their alternatives (1%) on the prospective purchase (movie tickets) and thus are both the reward-maximizing option for that purchase in their respective conditions. As before, we expect that the Shell branding of the reward-maximizing card will discourage people from using it to buy movie tickets. More important, we create a third condition by copying the Co-branded condition and erasing all branding elements of Shell, including the Shell name and logo, from the reward-maximizing card but we keep the special reward category as 5% cash back for all gasoline (not specific to Shell). We examine whether the gas reward alone—without branding—prompts consumers to treat the credit card as a "gas card" and underuse it outside of the gasoline product category.

2.4.1 Method

We preregistered a plan to recruit 600 participants from MTurk (https://aspredicted.org/2SK_2S7) and 601 CloudResearch-approved MTurk participants completed the study. Participants were randomly assigned to one of three conditions: the Control condition, the Co-branded (Shell) condition, and the Gas-card condition. In all conditions, participants imagined possessing two credit cards, and the terms of the two cards were displayed to them side by side (See Figure 2.3). Participants in the Control condition imagined that they had a Visa card and a Mastercard. The Visa card paid 5% cash back at hotels and grocery stores and 1% on all other purchases. The Mastercard paid 5% cash back at department stores and 2% on all other purchases. All other terms (e.g., credit line, balance, and APR) were the same between cards.

Participants in the Co-branded condition had the same Visa card and a Shell Fuel Reward Mastercard. The Shell Fuel Reward Mastercard had the same terms as the Mastercard in the Control condition except for its reward structure. Specifically, the Shell Fuel Reward Mastercard paid 5% cash back on Shell gas instead of at department stores. Just like the Mastercard in the Control condition, the Shell Fuel Reward Mastercard also paid 2% cash back on all other purchases.

Participants in the Gas-card condition had the same Visa card and a Mastercard that was very similar to the Shell Fuel Reward Mastercard, but without the Shell brand name. That is, the Mastercard didn't have the term "Shell Fuel Reward" in its name and paid 5% cash back on all gas purchases, instead of Shell gas only. All its other terms were identical to the Mastercard of the other two conditions.

After reviewing the credit card information, participants proceeded to the next screen where they imagined that they went to a movie theater. To make the scenario vivid, participants selected the movie they wanted to watch from ten movies that were in theaters at the time of the study. After choosing a movie, they indicated the number of tickets they needed. Then, they saw their total purchase amount and proceeded to check out. Participants then chose one of their two credit cards to pay for their movie tickets. At this point, they could click to open a link to review their credit card terms if they wished to do so before making their decision.

After choosing a card, participants were asked to recall which card gave more cash back on movie tickets. They could select one of four options: the Visa, the Mastercard, “They both gave the same % of cash back,” or “I don’t remember.” Then, they were asked to indicate the two credit cards that they had in the hypothetical setting by answering a multiple-choice question. This final question served as a general attention check. Twelve participants answered this question incorrectly and were therefore excluded from analysis, as preregistered, leaving us with a final sample of 589 (M_{age} = 41.2 years; 57% female, 43% male, and 1% indicating “Other”).

	Visa	Mastercard
Credit Line	$8,000	$8,000
Current Balance (After most recent payment)	$0	$0
Interest Rate (APR)	19%	19%
Cashback Rewards	• 5% at grocery stores and hotels. • 1% on all other purchases.	• 5% at department stores • 2% on all other purchases.

	Chase Visa	Shell Fuel Rewards Mastercard
Credit Line	$8,000	$8,000
Current Balance (After most recent payment)	$0	$0
Interest Rate (APR)	19%	19%
Cashback Rewards	• 5% at grocery stores and hotels. • 1% on all other purchases.	• 5% on Shell gas. • 2% on all other purchases

	Visa	Mastercard
Credit Line	$8,000	$8,000
Current Balance (After most recent payment)	$0	$0
Interest Rate (APR)	19%	19%
Cashback Rewards	• 5% at grocery stores and hotels. • 1% on all other purchases.	• 5% on gas. • 2% on all other purchases.

Figure 2.4 Study 2 credit terms in Control (upper), Co-branded (middle), and Gas-card (lower) conditions.

2.4.2 Results and Discussion

First, we replicated findings from Study 1. Although the Mastercard was the reward-maximizing card in all conditions to pay for the movie tickets, reliably fewer participants chose the Mastercard in the Co-branded condition relative to the Control condition (61% vs. 84%, $\chi^2(1)$ =

25.6, $p < .001$). Thus, people were less likely to use the reward-maximizing Mastercard to buy movie tickets when it featured the Shell brand.

In the Gas-card condition, participants had a Mastercard without the Shell brand that still offered 5% cash back on gasoline purchases. In the Gas-card condition, 81% of the participants used their Mastercard to buy movie tickets, a percentage that did not reliably differ from the 84% in the Control condition ($\chi^2(1) = .4$, $p = .550$). Thus, merely an exclusive reward category without any branding elements was not enough to discourage people from using the credit card outside of the exclusive category. By contrast, reliably fewer participants chose the Mastercard once its gas reward was labeled a Shell brand (61% vs. 81%, $\chi^2(1) = 19.3$, $p < .001$).

Moreover, and consistent with Study 1, fewer participants in the Co-branded condition (66%) accurately recalled which card paid more cash back on movie tickets than in each of the other two conditions (Control condition: 83%, $\chi^2(1) = 12.7$, $p < .001$; Gas-card condition: 80%, $\chi^2(1) = 8.4$, $p = .004$). The successful recall rate did not differ reliably between the Control and Gas-card conditions (80% vs. 83%, $\chi^2(1) = .3$, $p = .567$).

Further, all of the credit card choice patterns observed among the full sample persisted among the participants who correctly indicated that the Mastercard gave more cash back on movie tickets (N = 449). Specifically, among those participants who correctly recalled that the Mastercard offered the best cash back percentage for movie tickets, fewer participants chose the Mastercard in the Co-branded condition (82%) than in each of the other two conditions (Control condition: 95%, $\chi^2(1) = 12.0$, $p < .001$; Gas-card condition: 93%, $\chi^2(1) = 7.8$, $p = .005$). The percentage choosing the reward-maximizing Mastercard did not reliably differ between the Control and Gas-card conditions (95% vs. 93%, $\chi^2(1) = .3$, $p = .606$), suggesting again that an exclusive gas-reward category without accompanying branding was unable to reliably discourage people from using the credit card outside of the category. Taken together, the results demonstrate the key role of branding in restricting the usage of co-branded credit cards.

2.5 Study 3

We predicted that co-branding a credit card reduces people's attention to the card's reward structure (H_{2a}). In Study 3, we test this prediction with a mouse-tracking technique. We further hypothesized that the reduced attention decreases people's awareness of the card's rewards outside of its featured brand, which in turn restricts the card use outside of the brand (H_{2b}). We formally examine this potential process pathway via serial mediation.

2.5.1 Method

Procedure. As preregistered, we invited 600 participants from Prolific.co (https://aspredicted.org/WZ9_FNN) to participate. Because we needed to track mouse cursor movements, we required that participants use a computer to complete the study. Smart phones, tablets, or other devices were not permitted, and this restriction was implemented by Prolific.co.

Participants were randomly assigned to either a Control condition or a Co-branded condition. In both conditions, participants imagined possessing two credit cards. In the Control condition, they carried a Visa card and a Mastercard. These two credit cards had the same terms (the same credit line, current balance, and APR) except for their cash back rewards. The Visa card paid 5% cash back at restaurants and grocery stores and 1% on all other purchases. The Mastercard paid 5% cash back at gas stations and drug stores and 1.5% on all other purchases. In the Co-branded condition, participants had the same Visa card and an REI Mastercard that was identical to the Mastercard in the Control condition except for its reward structure. Specifically, the REI Mastercard paid 5% cash back at REI Co-ops instead of at gas stations and drug stores. Just like the Mastercard in the Control condition, the REI Mastercard offered 1.5% cash back on all other purchases.

The information for both credit cards was provided to participants on the same screen with some mouse-tracking features. Specifically, all information was contained in a 2 (card type: Visa and Mastercard) by 4 (terms: credit line, current balance, APR, and reward structure) matrix (see Figure 2.4). Each cell contained a gray cover that concealed the information. Participants could "open" a cell by moving their mouse cursor over that cell. Once the mouse cursor moved into the cell, the gray cover was removed and the information was revealed. As soon as the mouse cursor left the cell, the cell closed and concealed the information under the gray cover once again. Figure 2.4 shows an example where the mouse cursor was in the reward cell of the REI Mastercard, revealing that reward information. This mouse-tracking screen was the only place where participants could learn about their credit card terms in this study. On an initial screen, before encountering the credit card matrix, participants had a chance to practice uncovering cells of a sample matrix with the same mouse-tracking features. The sample matrix contained information about car prices that were irrelevant to this study.

After reviewing the credit card terms, participants proceeded to the next screen where they imagined that they were shopping for office supplies. Once participants were ready to check out, they chose one of their two credit cards to pay.

After choosing a card, participants indicated which of the two cards paid more cash back on office supplies. This question tested participants' awareness of the credit card reward structure. Then, participants indicated which two credit cards that they had in the office-supply scenario by choosing from a list of options, including "Visa and Mastercard," "Visa and REI Mastercard," "American Express and Discover," and "I don't remember." This question served as a general attention check. We preregistered to exclude participants who did not pass this general attention check. Eleven participants failed, leaving us a final sample of 589 (M_{age} = 37.5 years; 49% female, 49% male, and 2% indicating "Other").

Mouse-tracking measures for attention. With the mouse-tracking interface, participants could open only information for one credit card at any given time by hovering their mouse cursor over the corresponding cell to reveal that attribute. This allowed us to measure the information acquisition behaviors for each credit card in isolation. Specifically, we measured how many times participants opened a cell and how long their mouse cursor hovered over the cell every time. As pre-registered, we excluded hovers shorter than 0.18 seconds because of prior work establishing that people cannot accurately perceive anything they see for shorter than that amount of time (Card, Moran, & Newell, 2018). Next, among the remaining hovers over each cell, we exclude extremely long hovers that were greater than 2 standard deviations above the average duration for that cell, also as pre-registered. The remaining hovers were considered valid for our analyses. Following existing research (Johnson et al., 2002), we measured the frequency of opening each cell (i.e., the number of valid hovers over that cell) and the total time spent looking at each cell (i.e., the sum of durations of valid hovers over that cell) as proxies for participants' attention to the credit card information in that cell.

Figure 2.5 An Illustrative Example of the Mouse Tracking Interface

2.5.2 Results and Discussion

Card usage. We replicated the effect of co-branding on card usage outside of the featured brand. Reliably fewer participants chose to use the reward-maximizing Mastercard to pay for the office supplies when the Mastercard featured the REI brand than when the Mastercard did not feature the REI brand (60% vs. 78%, $\chi^2(1) = 21.4$, $p < .001$).

Attention to rewards. As predicted, participants paid less attention to the Mastercard's reward structure when it featured the REI brand. They opened the Mastercard's reward cell less frequently and overall spent less time looking at that cell in the Co-branded condition than in the Control condition (frequency of cell visits: $M_{\text{co-branded}} = 3.1$, $M_{\text{control}} = 3.9$, $t(449.8) = 3.5$, $p < .001$; total time: $M_{\text{co-branded}} = 4.2$ seconds, $M_{\text{control}} = 5.2$ seconds, $t(548.5) = 4.5$, $p < .001$). Notably, the presence of the REI brand did not affect attention to other attributes (credit line, current balance, and APR; see Appendix B). Thus, people tended to pay less attention to the credit card's reward structure in particular when the credit card features a merchant brand.

Following our pre-registration, we focus on the total time spent at each cell as the attention measure for the remaining analyses.

Mediation. We tested, via serial mediation, whether the REI brand of the Mastercard discouraged participants from using it to buy office supplies by reducing their attention to its reward structure, which in turn decreased awareness of its rewards outside of REI purchases (see Figure 2.5). We coded the independent variable, co-branding, as 1 if the Mastercard was co-branded with REI and 0 otherwise. We used the total time each participant spent looking at the Mastercard's reward cell to measure their attention to the Mastercard's reward structure as the first mediator. Then, we coded the second mediator, awareness, as 1 if participants correctly indicated that the Mastercard paid more cash back on office supplies, and 0 otherwise. Finally, we coded the dependent variable, reward-maximizing choice, as 1 if they chose the Mastercard and 0 otherwise.

Our results showed a significant total effect of co-branding on the card choice (β_{total} = -.14, SE = .03, p < .001) that was serially mediated by attention and awareness ($\beta_{indirect}$ = -.01, SE = .01, p = .022). Our results also suggested that the effect of co-branding on people's awareness of the reward outside of the brand was fully mediated by their attention to the reward structure ($\beta_{indirect\text{-}awareness}$ = -.02, *SE* = .01, p = .038), as the direct effect of co-branding on awareness was not statistically reliable (β_4 = -.06, *SE* = .04, p = .117).

Nevertheless, there remained a significant direct effect of co-branding on the card choice (β_6 = -.12, *SE* = .03, p < .001), suggesting that co-branding also affected credit card use via other mechanism(s) beyond attention and awareness. We have suggested that the merchant brand on a co-branded credit card can often reduce perceived card-purchase fit, which restricts usage of the card outside of its featured brand. We further examine this mechanism in Study 4.

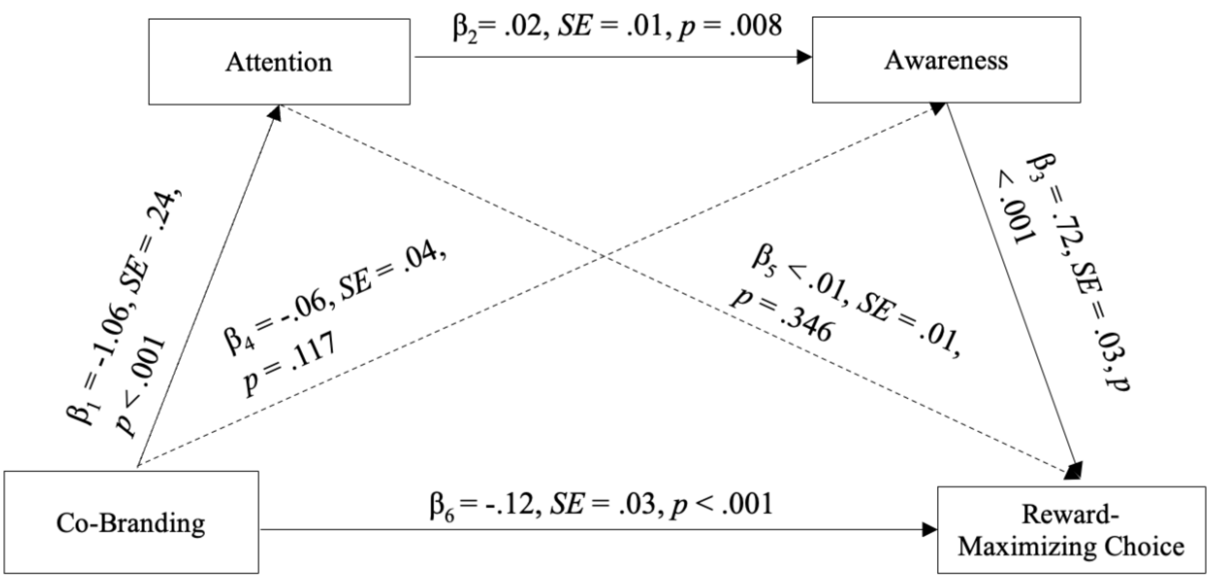

Figure 2.6 Study 3 Mediation

2.6 Study 4

We have so far shown that consumers are reluctant to use co-branded credit cards outside of the cards' featured brands. We further demonstrated that one reason is that credit card co-branding reduces users' attention to the co-branded card's full rewards, which in turn reduces their awareness of card's potential reward advantages outside of its featured brand.

We have also shown that even among participants who were aware of the reward-maximizing card, those participants were less likely to use the reward-maximizing card when it featured a brand that did not fit the prospective purchase than when it did not feature the brand. These findings suggest that co-branding affects credit card usage through at least one other mechanism besides attention and awareness. Indeed, we have proposed that the featured brand on a co-branded credit card makes many potential purchases outside of the featured brand feel like a bad "fit" with the card, which in turn discourages people from using the card for those purchases because they do not feel right about the bad fit. In other words, we predicted that the perceived card-purchase "fit "and the sense of "feeling right" serially mediate the effect of co-branding on card usage (H_3). In Study 4, we test this process pathway.

2.6.1 Method

As preregistered (https://aspredicted.org/1QB_51S), 600 participants on Prolific.co completed our study. The study used the same grocery-shopping scenario and followed the same procedure as Study 1, except for one major difference: After participants recalled which credit card paid more cash back on groceries, they preceded to answer two more questions. The first question measured the perceived fit between the grocery purchase and the reward-maximizing card (i.e., the Best Buy Visa in the Co-branded condition or the BOA Visa in the Control condition), as follows.

> "To what extent did you think the [Bank of America Visa / Best Buy Visa] was a good fit with the grocery purchase?"

Participants responded on an 11-point scale from 0 to 10, with higher scores indicating a better fit. The second question assessed their sense of "feeling right" about using the Visa card to buy groceries, as follows.

> "To what extent did it feel right to use the [Bank of America Visa / Best Buy Visa] to pay for your groceries?"

Participants responded on an 11-point scale from 0 to 10, with higher scores indicating a greater sense of feeling right. At the end of the survey, participants answered a general attention check question identical to that of Study 1. Ten participants answered it incorrectly, and we excluded them as pre-registered, resulting in a final sample of 590 (M_{age} = 34.9 years; 49% female, 49% male, and 2% indicating “Other” or skipping the gender question).

2.6.2 Results and Discussion

First, we replicated the effect of co-branding on card usage outside of the featured brands. Reliably fewer participants chose to use the reward-maximizing card to pay for the groceries when the card featured the Best Buy brand than when the card did not feature any retail brand (46% vs. 69%, $\chi^2(1) = 30.7$, $p < .001$).

Further, as predicted, perceived card-purchase fit was lower in the Co-branded condition than in the Control condition ($M_{\text{co-branded}} = 4.62$, $M_{\text{control}} = 6.87$, $t(560.7) = 9.5$, $p < .001$), suggesting that participants considered the reward-maximizing card a worse fit with the grocery purchase when the card featured the Best Buy brand than when it did not feature a retail brand. Also as predicted, the experience of “feeling right” was lower in the Co-branded condition than in the Control condition ($M_{\text{co-branded}} = 4.53$, $M_{\text{control}} = 6.97$, $t(560.6) = 10.0$, $p < .001$), suggesting that using the Best Buy Visa for groceries did not feel as right as using the BOA Visa, even though the two cards equivalently maximized the reward.

Next, we tested, via serial mediation, whether the Best Buy co-branding discouraged participants from using the reward-maximizing card by inducing a lack of card-purchase fit and producing a sense of not “feeling right” about the lack of fit. Moreover, as discussed earlier, this process pathway should be independent of whether people knew which card paid more cash back. Thus, as pre-registered, to elicit the unique mediated effect via the card-purchase fit and the sense of "feeling right", we included participants' awareness of the reward-maximizing card as a parallel mediator to control for its contribution to the total effect of co-branding (see Figure 2.6). We coded the independent variable, co-branding, as 1 if the reward-maximizing card was the Best Buy Visa and 0 if it was the BOA Visa. We used the reported scores on their perceived fit between the grocery purchase and the reward-maximizing card as the first mediator. Then, we used the reported scores on the sense of "feeling right" as the second mediator. Next, we coded the other parallel mediator, awareness, as 1 if participants correctly indicated which card paid more cash back on groceries, and 0 otherwise. Finally, we coded the dependent variable, reward-maximizing choice, as 1 if they chose the reward-maximizing card to pay for the groceries and 0 otherwise.

Our results showed a significant total effect of co-branding on the reward-maximizing choice (β_{total} = -.117, *SE* = .04, $p < .001$). This total effect was mediated through two separate pathways. Most important, controlling for the mediated effect via awareness ($\beta_{indirect\ (2)}$ = -.03, *SE* = .02, p = .070), a separate piece of the total effect was serially mediated by the perceived card-purchase fit and the sense of feeling right ($\beta_{indirect\ (1)}$ = -.05, *SE* = .02, p = .002). Moreover, our results also suggested that the total effect of co-branding was fully mediated by the two process pathways as the direct effect of co-branding on the card choice was not statistically reliable (β_6 = -.04, *SE* = .03, p = .191). Overall, these results supported our hypothesis that people perceive a lack of fit between a co-branded credit card and a purchase that doesn't match the card's featured brand, which does not feel right and thus discourages people from using the co-branded card for that purchase, even if they are aware that the card is the reward-maximizing option.

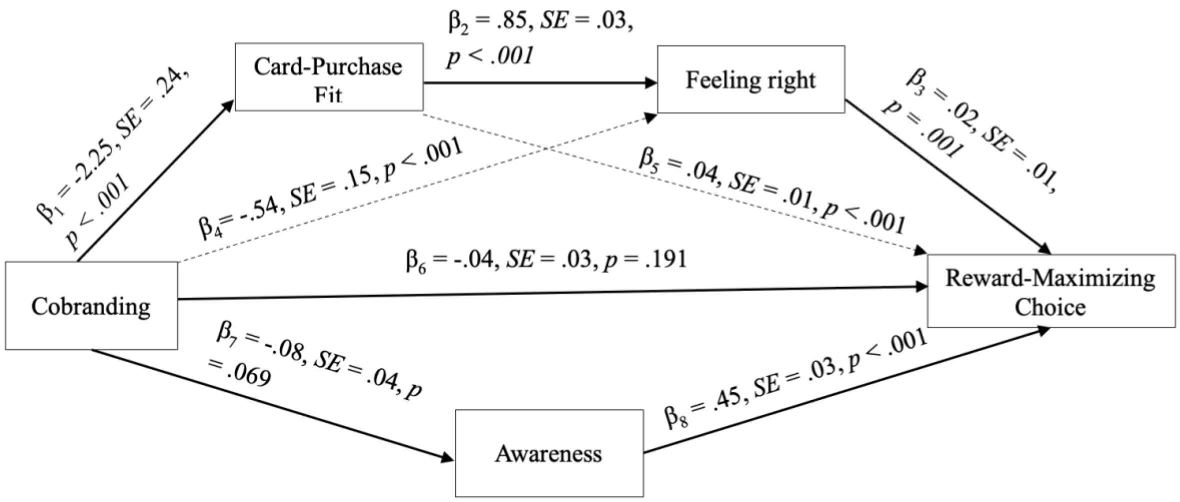

Figure 2.7 Study 4 Mediation

2.7 General Discussion

This research shows that consumers are reluctant to use a co-branded credit card outside of its featured brand. Specifically, consumers are less likely to use their reward-maximizing credit card to pay for a purchase if that card features a merchant brand that is not associated with the intended purchase. This pattern occurs for two reasons. First, when a credit card features a merchant brand, users examine its reward structure less carefully. As a result, they are less aware of the card's reward advantages outside its featured brand and thus are less likely to use it when they should. Further, when shopping outside the card's featured brand, people tend to perceive a low fit between the card and the purchase, which makes people reluctant to use the co-branded card outside its featured brand even if they are aware of its reward advantage. Such patterns are reflected in consumers' reports of their own credit card usage behavior, where they report using co-branded credit cards substantially less often than their non-co-branded counterparts, even when the rewards for most purchases are comparable.

2.7.1 Relationship with Previous Research

On credit card behaviors. As discussed earlier, previous research has investigated various types of credit card decisions that are not financially optimal (e.g., Gathergood et al., 2019; Keys & Wang, 2019; Stewart, 2008; Gross & Souleles, 2002; Amar et al., 2011; Ponce, Seira, & Zamarripa, 2017). Most of these findings center on credit card costs (i.e., interest rates, or APR) and repayment behaviors. The current research is novel in the sense that we investigate credit card purchase behaviors while focusing on credit card benefits, specifically, purchase rewards.

On effects of co-branding. Existing research on co-branded credit cards mainly focuses on consumer behaviors and attitudes within the partnering brand, such as brand loyalty and spending within the brand (e.g., Zhao, Gopalakrishnan, & Narasimhan, 2022). The current research, by contrast, investigates the effect of co-branding on card usage outside of the cards' featured brands.

On choices of payment instruments. A body of research on purchase behaviors examines whether consumers will make certain purchases or what kinds of purchases they will make. The current research contributes to a different body of purchase behavior research that examines what payment instrument consumers choose to use once they have decided to make a purchase. Previous research has investigated consumer choices of payment instruments between cash and cashless instruments (e.g., Bounie & Francois, 2006), between debit and credit cards (e.g., Zinman, 2009), and between credit cards (Ponce, Seira, & Zamarripa, 2017). We add to this literature by investigating payment instrument choices when co-branded credit cards are involved. To that end, the research of Ponce and colleagues (2017) may be most relevant. While they mainly focus on the role of credit card costs (i.e., their interest charges) in credit card choices, we focus on the role of credit card benefits (i.e., their purchase rewards).

On mental accounting. Research in mental accounting has found a similar restricted use of a different payment instrument—gift cards—that is also associated with the specific brands on the cards. Reinholtz, Bartels, and Parker (2015) have shown that people tend to categorize the fund in a brand-specific gift card into a mental account specifically for products typical of the brand, such as jeans at Levi's. As a result, people shopping at Levi's with a Levi's gift card are more likely to buy jeans compared to those shopping with other currency such as debit cards. The current research differs from the findings of Reinholtz and colleagues in four main aspects. First, a credit card is a payment method that can be used broadly wherever credit cards are accepted whereas a gift card is a source of funds that typically can only be spent at featured brand locations. Second, unlike gift cards, credit cards are not sources of funds, and therefore the categorization of funds process identified by Reinholtz et al. (2015) cannot explain our findings. Third, whereas Reinholtz et al. (2015) focus on *what* people choose to buy with a certain fund, we investigate how people pay when they have already decided what to buy. The findings of Reinholtz and colleagues cannot easily generate predictions regarding this topic.

Finally, the restricted use of a gift card for certain types of products does not compromise consumers' ability to maximize rewards. There are no obvious financial benefits for using a gift card for one type of product versus another. In the current research, however, we demonstrate that consumers are reluctant to use co-branded credit cards in some situations where they could maximize reward benefits by doing so.

2.7.2 Attention Spread Across Reward Terms

We have shown that people pay less attention to a credit card's reward structure when the card features a merchant brand. One might wonder how people allocate their limited attention across different reward terms in the reward structure. We speculate that consumers might primarily focus on the brand-specific perks while tending to overlook other unbranded rewards. Research has shown that people often process information in a selective way (Nickerson 1998; Sanbonmatsu et al., 1998), focusing on evidence that confirms their hypotheses while neglecting evidence that is inconsistent with their hypotheses (Kardes et al., 2004; Koriat, Lichtenstein, & Fischhoff, 1980; Snyder, Campbell, & Preston, 1982; Snyder & Swann, 1978). Thus, because people have the assumption that the most important rewards are its brand-specific perks, we expect them to allocate most of their attention to the brand-specific perks that confirms that assumption while paying minimal attention to other unbranded rewards. Future work could examine consumers' attention to reward terms at a more granular level.

2.7.2 Implications

Reluctance to use co-branded credit cards outside of their featured brands can hurt both credit card users and issuers. Co-branded credit card users could lose out on rewards by avoiding their reward-maximizing co-branded card for purchases outside of the brand. In addition, the credit card issuer, including both the issuing bank and the payment network company, such as Visa or Mastercard, may lose revenue when their co-branded credit cards are underused. Credit card issuers may wish to rethink their co-branding strategies to encourage broader usage beyond their featured brands. As discussed, people underuse their co-branded credit cards partly because they minimally attend to, and therefore are unaware of, their potential reward advantages outside their featured brands. Thus, credit card companies could consider emphasizing their co-branded cards' benefits outside their featured brands through various advertising practices. At the same time, some consumers underuse their co-branded credit cards because they feel uncomfortable with the low fit between the purchase and the featured brand. To that end, credit card companies might want to advertise their co-branded cards as a good fit with general uses that should not be limited to their featured brands.

References

Aaker, D. A., & Keller, K. L. (1990). Consumer evaluations of brand extensions. Journal of marketing, 54(1), 27-41. https://doi.org/10.1177/002224299005400102

American Bankers Association (2023). Credit Card Market Monitor Q2 2022. Available at: https://www.aba.com/news-research/analysis-guides/credit-card-market-monitor

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengström, E. (2016). Deciding for others reduces loss aversion. *Management Science*, *62*(1), 29–36. https://doi.org/10.1287/mnsc.2014.2085

Bar-hillel, M., & Budescu, D. (1995). The elusive wishful thinking effect. *Thinking & Reasoning*, *1*(1), 71–103. https://doi.org/10.1080/13546789508256906

Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics*, *12*(4), 428–454. https://doi.org/10.1016/0196-8858(91)90029-I

Blackett, T., & Russell, N. (2000). Co-branding—the science of alliance. Journal of Brand Management, 7, 161-170. https://doi.org/10.1057/bm.2000.3

Bounie, D., & François, A. (2006). Cash, check or bank card? The effects of transaction characteristics on the use of payment instruments. Telecom Paris Economics and Social Sciences Working Paper No. ESS-06-05. http://dx.doi.org/10.2139/ssrn.891791

Budescu, D. V., & Wallsten, T. S. (1995). Processing linguistic probabilities: General principles and empirical evidence. In J. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Psychology of Learning and Motivation* (Vol. 32, pp. 275–318). Academic Press. https://doi.org/10.1016/S0079-7421(08)60313-8

Card, S. K. (2018). The psychology of human-computer interaction. CRC Press.

Clark, D. A. (1990). Verbal uncertainty expressions: A critical review of two decades of research. *Current Psychology*, *9*(3), 203–235. https://doi.org/10.1007/BF02686861

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, *66*(5), 819–829. https://doi.org/10.1037/0022-3514.66.5.819

Fantino, E., & Esfandiari, A. (2002). Probability matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *56*(1), 58–63. https://doi.org/10.1037/h0087385

Federal Reserve System (2023). Economic Well-Being of U.S. Households in 2021. Available at: https://www.federalreserve.gov/consumerscommunities/shed.htm

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, *9*(2), 127–152. https://doi.org/10.1007/BF00143739

Fiske, S. T., & Taylor, S. E. (1991). Social cognition. Mcgraw-Hill Book Company.

Fox, C. R., & Ulkumen, G. (2011). Distinguishing two dimensions of uncertainty. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3695311

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), Article 7302. https://doi.org/10.1038/466029a

Hull, D. L. (2001). Science and selection: Essays on biological evolution and the philosophy of science. Cambridge University Press.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, *41*(4), 291–297. https://doi.org/10.1037/h0056878

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. Journal of economic theory, 104(1), 16-47. https://doi.org/10.1006/jeth.2001.2850

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, *39*(1), 17–31. https://doi.org/10.1287/mnsc.39.1.17

Kahneman, D., Slovic, S. P., Slovic, P., Tversky, A., & Press, C. U. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454. https://doi.org/10.1016/0010-0285(72)90016-3

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251. https://doi.org/10.1037/h0034747

Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition*, *11*(2), 143–157. https://doi.org/10.1016/0010-0277(82)90023-3

Kardes, F. R., Cronley, M. L., Kellaris, J. J., & Posavac, S. S. (2004). The role of selective information processing in price-quality inference. Journal of Consumer Research, 31(2), 368-374. https://doi.org/10.1086/422115

Kettle, K. L., Trudel, R., Blanchard, S. J., & Häubl, G. (2016). Repayment concentration and consumer motivation to get out of debt. Journal of Consumer Research, 43(3), 460-477. https://doi.org/10.1093/jcr/ucw037

Kim, H., & Markus, H. R. (1999). Deviance or uniqueness, harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology*, *77*, 785–800. https://doi.org/10.1037/0022-3514.77.4.785

Koehler, D. J., & James, G. (2009). Probability matching in choice under uncertainty: Intuition versus deliberation. *Cognition*, *113*(1), 123–127. https://doi.org/10.1016/j.cognition.2009.07.003

Kooreman, P. (2000). The labeling effect of a child benefit system. American Economic Review, 90(3), 571-583. DOI: 10.1257/aer.90.3.571

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human learning and memory, 6(2), 107. https://doi.org/10.1037/0278-7393.6.2.107

Kray, L. J. (2000). Contingent weighting in self-other decision making. *Organizational Behavior and Human Decision Processes*, *83*(1), 82–106. https://doi.org/10.1006/obhd.2000.2903

Krizan, Z., Miller, J. C., & Johar, O. (2009). Wishful thinking in the 2008 U.S. presidential election: *Psychological Science*. https://doi.org/10.1177/0956797609356421

Li, Y., & Pandelaere, M. (2021). The denomination–spending matching effect. Journal of Business Research, 128, 338-349. https://doi.org/10.1016/j.jbusres.2021.02.020

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267–286. https://doi.org/10.1037/0033-2909.127.2.267

Marks, R. W. (1951). The effect of probability, desirability, and “privilege” on the stated expectations of children. *Journal of Personality*, *19*(3), 332–351. https://doi.org/10.1111/j.1467-6494.1951.tb01107.x

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Massey, C., Simmons, J. P., & Armor, D. A. (2011). Hope over experience: Desirability and the persistence of optimism. *Psychological Science*, *22*(2), 274–281. https://doi.org/10.1177/0956797610396223

Nan, X., & Heo, K. (2007). Consumer responses to corporate social responsibility (CSR) initiatives: Examining the role of brand-cause fit in cause-related marketing. Journal of advertising, 36(2), 63-74. https://doi.org/10.2753/JOA0091-3367360204

Neimark, E. D., & Shuford, E. H. (1959). Comparison of predictions and estimates in a probability learning situation. *Journal of Experimental Psychology*, *57*(5), 294–298. https://doi.org/10.1037/h0043064

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2(2), 175-220. https://doi.org/10.1037/1089-2680.2.2.175

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

Packaged Facts (2021). Co-Branded Credit Cards in the U.S., 8th edition. Available at: https://www.freedoniagroup.com/packaged-facts/co-branded-credit-cards-in-the-u-s-,-8th-edition

Park, C. W., Milberg, S., & Lawson, R. (1991). Evaluation of brand extensions: The role of product feature similarity and brand concept consistency. Journal of consumer research, 18(2), 185-193. https://doi.org/10.1086/209251

Park, I., Windschitl, P. D., Miller, J. E., Smith, A. R., Stuart, J. O., & Biangmano, M. (2022). People express more bias in their predictions than in their likelihood judgments. *Journal of Experimental Psychology: General*, *152*(1), 45–59. https://doi.org/10.1037/xge0001258

Polman, E. (2012). Self–other decision making and loss aversion. *Organizational Behavior and Human Decision Processes*, *119*(2), 141–150.

Ponce, A., Seira, E., & Zamarripa, G. (2017). Borrowing on the wrong credit card? Evidence from Mexico. American Economic Review, 107(4), 1335-1361. DOI: 10.1257/aer.20120273

Postman, L., Jenkins, W. O., & Postman, D. L. (1948). An experimental comparison of active recall and recognition. The American Journal of Psychology, 61(4), 511-519. https://doi.org/10.2307/1418315

Reinholtz, N., Bartels, D. M., & Parker, J. R. (2015). On the mental accounting of restricted-use funds: How gift cards change what people purchase. Journal of Consumer Research, 42(4), 596-614. https://doi.org/10.1093/jcr/ucv045

Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. Psychonomic Bulletin & Review, 5, 197-220. https://doi.org/10.3758/BF03212944

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. Biometrics Bulletin, 2(6), 110–114. https://doi.org/10.2307/3002019

Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. Consumer Psychology Review, 4(1), 4-25. https://doi.org/10.1002/arcp.1067

Simmons, J. P., & Massey, C. (2012). Is optimism real? *Journal of Experimental Psychology: General*, *141*(4), 630–634. https://doi.org/10.1037/a0027405

Simmons, J. P., & Nelson, L. D. (2006). Intuitive confidence: Choosing between intuitive and nonintuitive alternatives. *Journal of Experimental Psychology: General*, *135*(3), 409–428. https://doi.org/10.1037/0096-3445.135.3.409

Sirota, M., & Juanchich, M. (2015). A direct and comprehensive test of two postulates of politeness theory applied to uncertainty communication. *Judgment and Decision Making*, *10*(3), Article 3.

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285. https://doi.org/10.1126/science.3563507

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, *24*(2), 311–322. https://doi.org/10.1111/j.0272-4332.2004.00433.x

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352. https://doi.org/10.1016/j.ejor.2005.04.006

Snyder, M., Campbell, B. H., & Preston, E. (1982). Testing hypotheses about human nature: Assessing the accuracy of social stereotypes. Social cognition, 1(3), 256-272. https://doi.org/10.1521/soco.1982.1.3.256

Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. Journal of personality and social psychology, 36(11), 1202. https://doi.org/10.1037/0022-3514.36.11.1202

Theil, M. (2002). The role of translations of verbal into numerical probability expressions in risk management: A meta-analysis. *Journal of Risk Research*, *5*(2), 177–186. https://doi.org/10.1080/13669870110038179

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101–118. https://doi.org/10.1111/1467-6419.00106

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

Windschitl, P. D., Smith, A. R., Rose, J. P., & Krizan, Z. (2010). The desirability bias in predictions: Going optimistic without leaving realism. *Organizational Behavior and Human Decision Processes*, *111*(1), 33–47. https://doi.org/10.1016/j.obhdp.2009.08.003

Winkielman, P., Huber, D. E., Kavanagh, L., & Schwarz, N. (2012). Fluency of consistency: When thoughts fit nicely and flow smoothly. Cognitive consistency: A fundamental principle in social cognition, 89-111.

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. The psychology of evaluation: Affective processes in cognition and emotion, 189, 217.

Zhao, N., Gopalakrishnan, A., & Narasimhan, C. (2022). The Impact of Co-Branded Credit Card Adoption on Customer Loyalty. Available at SSRN 3820361.

Zinman, J. (2009). Debit or credit?. Journal of Banking & Finance, 33(2), 358-366.

Appendices

Appendix A. Response Distributions of Chapter 1 Studies

Table A1. Response Distributions of Studies 1, 2a, and 2b of Chapter 1

		% Choosing Each Number							
		1	2	3	4	5	6	7	8
Study 1	**Identify the most likely**								
	High chance of drawing a 1	96.7%	1.3%	1.3%	0.7%	n/a	n/a	n/a	n/a
	Low chance of drawing a 1	92.6%	1.5%	0.7%	0.7%	0.7%	1.5%	1.5%	0.7%
	Predict the outcome								
	High chance of drawing a 1	92.4%	2.3%	3.8%	1.5%	n/a	n/a	n/a	n/a
	Low chance of drawing a 1	63.3%	8.8%	10.9%	2.7%	1.4%	8.1%	2.7%	2.0%
Study 2a	**Identify the most likely first**								
	Identify the most likely	93.3%	0.0%	1.3%	0.7%	0.0%	0.0%	3.3%	1.3%
	Predict the outcome	58.0%	4.0%	3.3%	6.7%	6.0%	7.3%	9.3%	5.3%
	Predict the outcome first								
	Identify the most likely	82.7%	3.3%	4.0%	1.3%	3.3%	2.7%	2.0%	0.7%
	Predict the outcome	61.3%	3.3%	9.3%	5.3%	2.7%	10.0%	7.3%	0.7%
Study 2b	**Without the incentive/ reminder**								
	Identify the most likely	97.9%	0.0%	0.6%	0.0%	0.0%	1.4%	0.0%	0.0%
	Predict the outcome	56.6%	0.0%	6.2%	9.0%	6.9%	11.0%	9.0%	1.4%
	With the incentive/ reminder								
	Identify the most likely	97.9%	0.0%	0.0%	0.0%	0.0%	1.4%	0.7%	0.0%
	Predict the outcome	66.0%	2.8%	4.2%	3.5%	2.1%	9.7%	9.7%	2.1%

Appendix B. Additional Analyses of Chapter 1 Studies 2a & 2b

The within-subjects design of Studies 2a and 2b allows us to go further and to match each participant’s predictions to their likelihood judgments to get a sense of how many participants were internally (in)consistent. In Study 2a, most participants (88.0%) correctly identified 1 as most likely. If participants were perfectly internally consistent, those who correctly identified 1 might all predict that 1 would arise. However, 34.1% of these participants did not predict that they would draw a 1. Question order had no reliable effect on this proportion. Among people who identified 1 as the most-likely number first, 38.9% did not predict a 1; among those who identified 1 as the most-likely number second, 29.0% did not predict a 1, $\chi^2(1) = 2.26$, $p = .133$. Thus, recognizing and stating the most likely number before prediction did not bring predictions in line with that number.

Similarly, in Study 2b, 43.0% of people who correctly indicated 1 as most likely in the no-incentive-and-no-reminder condition did not subsequently predict 1; this number was only reduced marginally reliably with the incentive and reminder (32.6%, $\chi^2(1) = 2.79$, $p = .095$).

Appendix C. Additional Analyses of Chapter 1 Study 4.

A dichotomous analysis. We also pre-registered a dichotomous analysis. For this, we split the 514 participants into two groups based on whether the reported likelihood of the most likely winner was “unlikely” (less than 50% (n = 374)) or “likely” (greater than or equal to 50% (n = 140)). As predicted, fewer participants predicted their own perceived most likely winner as the title winner in the unlikely group than in the likely group ($P_{unlikely} = 74.3\%$ vs. $P_{likely} = 87.1\%$, $\chi^2(1) = 9.0$, $p = .003$). We conducted a logistic regression (with likelihood group, order, and their interaction as predictors) to investigate whether this effect was qualified by question order, but the interaction was not reliable ($\beta = -.996$, $SE = .577$, $p = .084$).

Analyses with the full sample. We re-conducted the main analyses of Study 4 with the full sample (N = 602). Although most participants were consistent between their directly indicated most likely winner and their most likely winner as revealed through percentage likelihoods, 88 of them (14.6%) were not. This presented us with two additional options for our analysis.

First, we examined the directly indicated most likely winner and used as our predictor the percentage estimate of how likely that team was to win. This meant that in some cases, the percentage estimates may have revealed a different most likely winner, but for this analysis, we simply focused on the percentage estimate of the directly indicated most likely winner.

We created a dependent variable that equaled 1 if a participant's prediction matched their directly indicated most likely winner and 0 if it did not. We ran a logistic regression regressing this dependent variable on the reported percentage likelihood of how likely the most likely team was to win, the question order (1 = participants first indicated the most likely winner, -1 = participants predicted first), and their interaction.

Predictions were sensitive to the reported likelihood of the most likely team: participants were less likely to predict their own most likely team when they indicated the likelihood of that team winning to be lower ($\beta = .045$, $SE = .008$, $p < .001$). Participants were overall more likely to predict their own most likely team to win when they first indicated the most likely winner than when they first made a prediction ($\beta =.034$, $SE = .246$, $p < .001$), but the order x likelihood interaction was not significant ($\beta = -.007$, $SE = .008$, $p = .344$).

For a dichotomous analysis, we split the sample based on whether the reported likelihood of the directly indicated most likely winner was "unlikely" (less than 50% (n = 449)) or "likely" (greater than or equal to 50% (n = 153)), and we examined whether people predicted their own most likely winner as the winner of the title. As predicted, fewer participants predicted their own perceived most likely winner as the title winner in the unlikely group than in the likely group

($P_{unlikely}$ = 68.4% vs. P_{likely} = 81.0%, $\chi^2(1) = 8.4$, $p = .004$). To examine the effect of question order, we ran a logistic regression that predicted whether participants' predictions matched their own most likely team (matched = 1, not matched = 0) by likelihood (unlikely = -1, likely = 1), question order (predict first = -1, most likely first = 1), and their interaction. The interaction was not statistically significant ($\beta = -.193$, $SE = .123$, $p = .116$), suggesting that the effect of likelihood was not affected by question order.

Next, we re-conducted these analyses ignoring the directly indicated most likely winner and simply using the revealed most likely winner (i.e., the team with the highest reported percentage likelihood of winning). If a participant assigned the same highest estimate for more than one team, we considered all teams with that highest estimate as the most likely winner for that participant. We re-ran the logistic regression, using the same predictors and coding as before (a prediction was coded as matching the most likely winner if the prediction matched any of that participant's most likely winners). The regression revealed similar results. Participants were less likely to predict the most likely winner as the winner when the percentage estimate of that team winning was lower ($\beta = .032$, $SE = .007$, $p < .001$). As before, participants were overall more likely to predict their own most likely team to win when they first indicated the most likely winner than when they first made a prediction ($\beta =.539$, $SE = .257$, $p = .036$), but the order x likelihood interaction was not reliable ($\beta = -.006$, $SE = .007$, $p = .382$).

Then, we conducted the dichotomous analysis, also using the same predictors and coding as before. As in the previous analyses, fewer participants predicted (one of) their own perceived most likely winner(s) as the winner in the unlikely group than in the likely group ($P_{unlikely}$ = 71.9% vs. P_{likely} = 86.9%, $\chi^2(1) = 13.2$, $p < .001$). This effect remained reliably large when participants first predicted the title winner ($P_{unlikely}$ = 64.5% vs. P_{likely} = 85.5%, $\chi^2(1) = 11.7$, $p <$

.001) but was not significant when they first indicated the most likely winner ($P_{unlikely} = 78.7\%$ vs. $P_{likely} = 88.6\%$, $\chi^2(1) = 2.8$, $p = .095$).

In addition, we also preregistered to compare participants' response consistency regarding their reported and revealed most likely winners in the likely versus unlikely group; this is less relevant to our main research focus. More participants gave the highest percentage likelihood to their reported most likely winner in the likely group than in the unlikely group ($P_{likely} = 91.5\%$ vs. $P_{unlikely} = 83.3\%$, $\chi^2(1) = 5.5$, $p = .019$).

Appendix D. Additional Analyses of Chapter 1 Study 5b

Parallel Mediation. In addition to the serial mediation analysis, we fitted the same data to a parallel mediation model where the two mediators, foreseeability and logical reasoning, each independently mediate the effect. Figure A1 shows the results. Each mediator had a significant mediating effect ($\beta_{logical\ reasoning} = -.139$, bootstrapped $SE = .037$, $p < .001$; $\beta_{foreseeability} = -.112$, bootstrapped $SE = .039$, $p = .004$). The parallel mediation model assumes that parallel mediation effects are independent of each other. However, the serial mediation results (Figure 2.3) suggested that the effect of the low absolute likelihood on logical reasoning was fully mediated through foreseeability ($b_1 * b_2 = -1.850$, bootstrapped $SE = .447$, $p < .001$, where b_1 and b_2 are from the serial mediation as in Figure 2.3), and there was no reliable direct effect ($b_4 = .117$, bootstrapped $SE = .575$, $p = .204$, as shown in Figure 2.3). In other words, the parallel mediation path through logical reasoning was not independent of foreseeability. Thus, the serial mediation model better fits the data.

Serial Mediation with the Full Sample. We also fitted the serial mediation model of Study 5b with the full sample (n = 286). We constructed the independent variable and mediators the same way as described in Study 5b. The dependent variable, "judgment-prediction correspondence," would equal 1 if a participant predicted their self-reported most likely number and 0 if they predicted a different number. As shown in Figure A2, results based on 5,000 bootstrapped samples suggest that our predicted mediating path remains reliable.

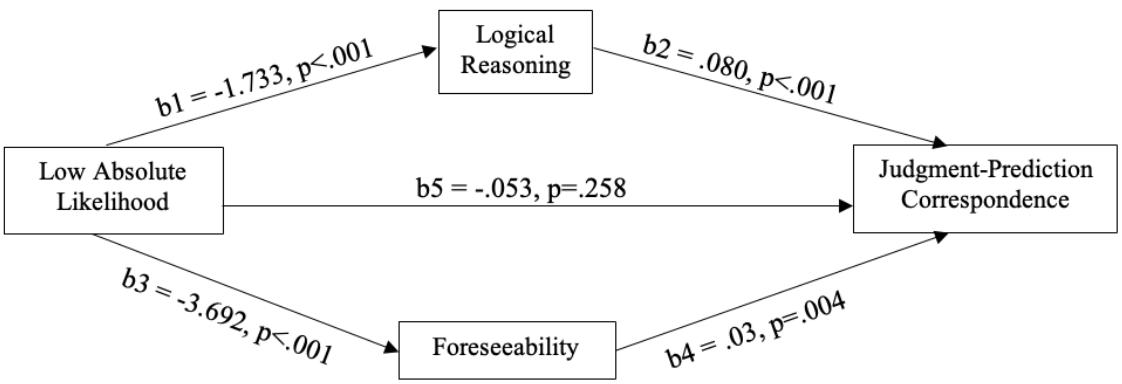

Figure A1

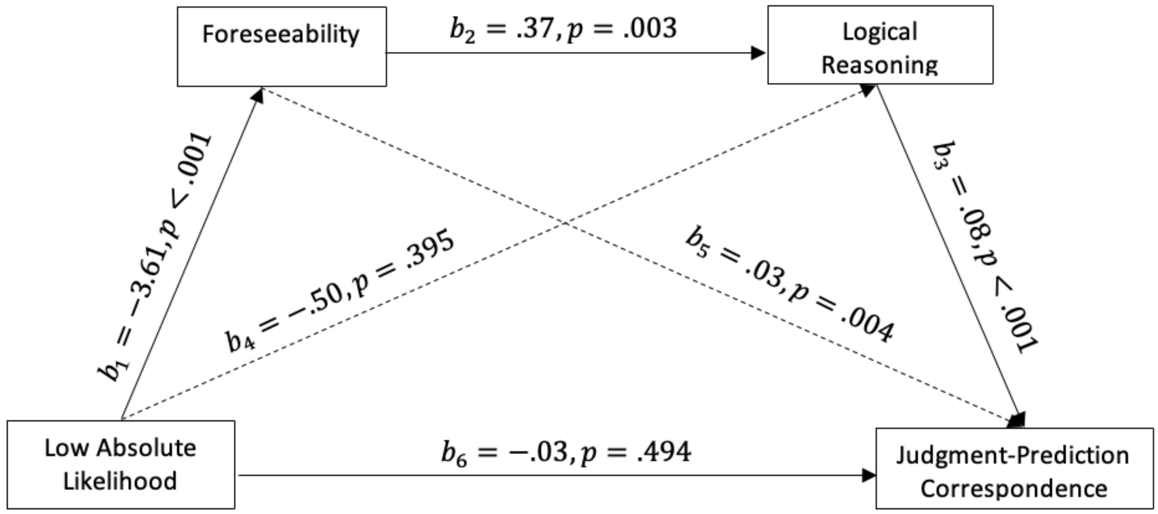

Figure A2

Appendix E. Additional Analyses of Chapter 1 Study 6

Between-Subject Analyses. We conducted the between-subject analyses by focusing on the first question each participant saw. Considering participants who first identified the most likely number, participants could easily identify the most likely number from both sets, and they did it equally well: when participants saw 10 balls, 88.6% correctly indicated that 1 was most likely to be drawn, and when they saw 100 balls, 86.3% correctly indicated this ($\chi^2(1) = .17$, $p = .676$). Thus, relative likelihood did not affect judgments about the most likely outcome. Predictions, however, were sensitive to relative likelihood. Considering only participants who gave their predictions first, more participants predicted a 1 when they drew from the set of 100 than from the set of 10 (75.0% vs. 53.1%, respectively; $\chi^2(1) = 14.19$, $p < .001$). Put differently, the vast majority (88.6%) indicated that 1 was most likely to be drawn from the set of ten but only about half (53.1%) predicted a 1 ($\chi^2(1) = 43.02$, $p < .001$). The gap was narrowed – though not eliminated – when the relative likelihood of drawing a 1 was greater, with 86.3% identifying 1 as most likely and 75.0% predicting it ($\chi^2(1) = 5.30$, $p = .021$).

Examining the Order Effect. In the main analysis, we observed an attenuating effect of the larger relative likelihood on the disconnect between predictions and reported most likely outcomes. This effect persisted in both question orders. When participants indicated 1 as most likely before predicting, 30.3% of those in the lower-relative-likelihood condition did not predict a 1, compared to only 4.8% in the higher-relative-likelihood condition ($\chi^2(1) = 26.99$, $p < .001$). When participants predicted before indicating 1 as most likely, these numbers were 38.1% and 15.1%, respectively ($\chi^2(1) = 15.57$, $p < .001$).

We also conducted a logistic regression to examine whether question order qualified the attenuating effect of the relative likelihood. Because we manipulated the relative likelihood of the most likely outcome (1), it was most relevant to examine whether the predictions of those who identified 1 as most likely were affected by that manipulation. Thus, we focused on the majority of participants (502 of the total 586) who correctly indicated 1 as most likely. We created a dependent variable that equaled 1 if a participant predicted a 1 and 0 otherwise. As preregistered, we ran a logistic regression regressing this dependent variable on relative likelihood (1 = higher relative likelihood, -1 = lower relative likelihood), question order (1 = participants first indicated the most likely number, -1 = participants predicted first), and their interaction. Consistent with the main results, there was a main effect of relative likelihood (β =.852, SE = .139, $p < .001$) and a main effect of question order ($\beta = .404$, SE = .139, $p = .004$) but no reliable interaction ($\beta = -.230$, SE = .139, $p = .098$).

Appendix F. Additional Analyses of Chapter 1 Study 9

In what follows, we describe several additional preregistered analyses of study 9. First, we ran a logistic regression to compare the most-likely vs. prediction gap and the most-likely vs. advice gap. Recall that in the first two questions of the study participants made a prediction and gave advice in counterbalanced order. So, their response to the first question was either their own prediction or their advice to others. For the current analysis, we only considered this first question to get the cleanest predictions and advice that were unaffected by each other. The last question, on the other hand, was always to indicate the most likely outcome. For the regression, we created a new dependent variable that consisted of participants' pooled responses to the first question (prediction or advice, between subjects) and last question (most-likely judgment, given by all participants) as the dependent variable: every participant thus had two entries in this new dependent variable: one for their advice or prediction, and one for their most-likely judgment. We recoded this new dependent variable as 1 if the response was a 1, and 0 otherwise.

We regressed this recoded dependent variable on whether the response is to their first or last question—that is, whether it is a most-likely judgment (most-likely judgment = -1, prediction or advice = 1), the type of their first question (prediction = -1, advice = 1), and their interaction. There was a significant main effect of the first question type (β =.611, *SE* = .147, $p < .001$). [Note that the dependent variable combined people's responses to the first question (i.e., advice or prediction) and the last question (i.e., most likely judgment). The main effect of the first question type was mainly driven by the difference in their first question (i.e., their advice versus predictions) even though their last question (i.e., their most likely judgments) did not reliably differ by the first question type, as we already demonstrated in the main analysis.] More important, there was a significant main effect of whether the response was a most-likely judgment (β = -.707, *SE* = .147, p < .001) that was qualified by a reliable interaction (β =.300, *SE* = .147, p = .044). This means that people were less likely to predict a 1 than to indicate 1 as most likely, but this gap was significantly smaller between their advice and their reported most likely outcome than between their prediction and their most likely outcome.

We also examined the order effect on advice. More participants recommended a 1 before versus after making a prediction for themselves ($P_{advise-first-condition}$= 89.6% vs. $P_{predict-first-condition}$ = 78.1%, $\chi^2(1) = 6.05$, p = .014).

Finally, we examined whether participants' predictions were identical to their advice. More participants chose the same number for both their prediction and advice in the advise-first condition than in the predict-first condition ($P_{advise-first-condition}$= 76.4% vs. $P_{predict-first-condition}$ = 57.7%, $\chi^2(1) = 10.34$, p = .001).

Appendix G. Chapter 2 Supplementary Study

In Chapter 2 Study 1, we asked participants to indicate which credit card gave more cash back on groceries after they chose a card to pay. In this appendix experiment, we use a stronger recall test: we ask participants to enter the exact cash back percentages of both credit cards. As a result, it is more difficult to answer this question correctly by chance (Postman, Jenkins, and Postman 1948). Thus, correct answers here give us more confidence that participants truly understand their cards' reward structures.

Method. We pre-registered a plan to recruit 800 participants from MTurk (https://aspredicted.org/YJX_2KL). Eight hundred two participants completed our study. The procedure was identical to Chapter 2 Study 1 except for one change. After participants chose a credit card to pay for their groceries, instead of indicating which card gave more cash back on groceries, they were asked to enter the exact cash back percentages of both cards on groceries. At the end of the survey, participants answered a general attention check question identical to that of Study 1. Fourteen answered it wrong, and we excluded them as pre-registered, resulting in a final sample of 788 (M_{age} = 40.6 years; 54% female, 44% male, and 2% indicating “Other”).

Results and Discussion. We replicated the key finding from experiment 1. Fewer participants chose the reward-maximizing credit card in the Co-branded condition (i.e., with the Best Buy Visa) than in the Control condition (i.e., with the BOA Visa; 54% vs. 84%, $\chi^2(1) = 84.3$, $p < .001$).

Next, we examined the recall test. Conceptually replicating experiment 1’s findings, fewer participants in the Co-branded (vs. Control) condition were aware that the reward-maximizing card paid more cash back on groceries: 70% in the Co-branded condition versus 82% in the Control condition entered a higher cash back percentage for the reward-maximizing card ($\chi^2(1) = 84.3$, $p < .001$). In addition, only marginally fewer participants in the Co-branded (vs. Control) condition entered the exactly correct percentages for both cards (62% vs, 68%, $\chi^2(1) = 3.2$, $p = .075$), suggesting that co-branding has a weaker effect on awareness of precise cash back percentages.

Those participants who correctly entered the precise cash back percentages on groceries (N = 509, out of 788) in both conditions should know unmistakably which card paid more cash back on groceries. Yet, fewer of them still chose the reward-maximizing credit card in the Cobranded-card versus Control condition (78% vs. 96%, $\chi^2(1) = 34.5$, $p < .001$).

This replication study provides evidence from a conservative recall test that people are less likely to use a co-branded credit card outside its featured brand even when they clearly understand that the co-branded credit card offers more benefits than other available cards.

Appendix H. Attention to Other Attributes of the Mastercard in Chapter 2 Study 3

Table A2 Attention

	Mean		Hypothesis Testing	
	Control condition	Co-branded condition	t-statistic	*p*-value
Frequency of opening the cell				
Credit Line	5.0	5.0	-0.4	.676
Current Balance	1.3	1.3	-1.1	.270
APR	1.4	1.4	-0.3	.748
Total time spent looking at the cell (seconds)				
Credit Line	6.6	6.7	-0.4	.718
Current Balance	0.7	0.7	-0.8	.406
APR	0.8	0.8	-0.3	.773