Consumers Undervalue Multi-Option Alternatives in Two-Stage Choice

Abstract. Decisions in the present lead to decisions in the future. Dominated options in future choices ought to be ignorable in the present for even minimally forward-looking consumers. Across six experiments (in two domains: consumer goods and risky gambles), we find that adding a less-valuable option (i.e. a less-preferred consumer good; a dominated gamble) decreases the choice share of an otherwise attractive alternative. This difference is moderated by the value difference between the more- and less-valuable options and by the transparency of the dominance relationship. Mouse-tracking reveals that consumers who attend more to the dominated option are less likely to choose the multi-option alternative. This work contributes to our understanding of multi-stage consumer decision-making and how consumers assess the overall value of choices.

Preregistrations, data, and materials are available at

https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT
Introduction

Only rarely does a consumer’s full decision process conclude at the moment of choice. Instead, each node in a decision tree typically leads to more decisions. Sometimes these are implicit, as when choosing a home implies choices among where to eat, where to shop, and who to visit. Other times these are explicit, as when choosing a restaurant implies choosing items from a menu or choosing to watch TV implies choosing a show.

Even this narrower case of explicit multi-option alternatives is ubiquitous. Multi-retailer gift cards constitute multi-option alternatives, since any part of the balance that is spent at a single retailer cannot be spent at another. Airline choices also constitute multi-option alternatives, as a given airline might have multiple routes available from origin A to destination B. Food and drink tickets (common at fairs, festivals, and conferences) are typically multi-option alternatives, as are game tokens at arcades and the like which can be used for a variety of games. Ultimately, many multi-attribute choices can be (and are) characterized as multi-option alternatives. For instance, the decision of which car (make and model) to purchase includes multi-option alternative(s): a consumer might decide on a Ford Mustang, only to be faced with the choice of blue vs. yellow.

In the present paper, we seek to better understand how consumers integrate across options in a choice set when deciding among such multi-option alternatives. We find consumers regularly sacrifice a chance at maximizing their subjective value by inappropriately integrating the value of less-attractive options.

**Normative Decision Rules in Two-Stage Choices**

While we document that consumers regularly undervalue multi-option alternatives relative to the most-valuable option, the normative value of a multi-option alternative is equal to its most-valuable option according to standard economic theory. Formally, we consider the case in which consumers make a decision between a single option S and a multi-option alternative M, where M is the choice set \{M_{H}, M_{L}\} (i.e., the choice between M_{H} and M_{L}), where M_{H} is preferred to M_{L}. The value of M is therefore
determined by \( \max \{M_H, M_L\} \). Given that \( M_H \) is preferred to \( M_L \), \( \max \{M_H, M_L\} \) is equal to \( M_H \). Thus, absent uncertainty regarding preference for \( M_H \) over \( M_L \), the choice between \( S \) and \( M \) simplifies to the choice between \( S \) and \( M_H \). \( S \) should be no more likely to be chosen when pitted against \( M \) as when it is pitted against \( M_H \); if there are some states of the world in which \( M_L \) is preferred to \( M_H \), \( S \) may be less likely to be chosen when pitted against \( M \) (vs. pitted against \( M_H \)). In other words, adding \( M_L \) as an option to the existing alternative \( M_H \) should not decrease the probability of choosing \( M \).

The normative value of \( M \) could exceed \( M_H \) if there were some possibility that in some states of the world \( M_L \) is preferred to \( M_H \). For example, a consumer may generally prefer hot coffee (\( M_H \)) to iced coffee (\( M_L \)), but may make an exception when the temperature exceeds 90°F. In this example and throughout, we consider cases in which the state of the world is exogenously determined, that is, neither the choice set nor the choice affects the preference ordering among \( S \), \( M_H \), and \( M_L \).

**Inherent Value of Choice**

Though standard economic theory suggests that consumers should value a multi-option alternative at a level equal to the expected maximum of its constituent pieces, findings from the behavioral literature suggest consumers often behave otherwise. Some findings suggest that consumers might value a multi-option alternative more than the maximum value of its component options due to the fact that the multi-option alternative enables choice (Brehm 1966; Bown, Read, and Summers 2003; Mochon 2013; Shin and Ariely 2004). In other words, because consumers often value the ability to make choices, this added value from the presence of choice itself can lead to an overvaluation of multi-option alternatives, relative to normative standards.

**Valuation of Sets**

Alternatively, consumers might undervalue a multi-option alternative. If, for instance, the valuation process for a multi-option alternative is similar to the valuation process for bundles, then this would lead to undervaluation, because consumers tend to value a bundle of goods according to a weighted average of the values of the components (Brough and Chernev 2012; Chernev and Gal 2010; Gaeth, Levin, Chakraborty, and Levin 1991; Yadav 1994; Shenhav and Karmarkar 2019). Similarly, there is
evidence that consumers estimate the value of a product to be a weighted average of its features (Weaver, Garcia, and Schwarz 2012; Troutman and Shanteau 1976), and adding a less-attractive bonus to a product decreases its value (Simonson, Carmon, and O’Curry 1994). This literature would suggest that adding a less-desirable option to an existing alternative decreases the value of the multi-option alternative in the same way that adding a lower-valued component to a bundle decreases the value of the bundle.

Indeed, consumers are capable of rapidly extracting average economic value from a set of products (Yamanashi Leib et al. 2020) and consumers sometimes do use such fast natural assessments inappropriately (Frederick and Kahneman 2002; Kahneman 2003).

**Prior Research on Multi-Option Alternatives**

Thus, there are two competing hypotheses regarding how consumers might misvalue multi-option alternatives relative to the normative baseline. Prior research on choice sets and assortments can help to inform the relative relevance of those prior literatures. Several papers have investigated various aspects of assortment choice (Kahn and Lehman 1991; Sood, Rottenstreich, and Brenner 2004). These papers focus on how assortments compare to one another and how assortment framing (as a set vs. an alternative) affects valuation, respectively. Because of its focus on the role of assortments, this literature tends not to focus on how the valuation of a multi-option alternative compares to its highest-valued option. To our knowledge, only two prior papers have directly examined this latter phenomenon (Le Lec and Tarroux 2020; Spiller and Ariely 2020). However, these papers leave many questions unanswered.

Le Lec and Tarroux (2020) tested the phenomenon in a single study in a subjective domain (i.e. a domain in which the relative values of products are subjective rather than objective). In this study, participants stated their willingness-to-pay (WTP) for various multi-option alternatives regarding what websites to spend time on at the end of an experimental session. The authors found substantial evidence for undervaluation of multi-option alternatives, and speculate that two possible explanations, anticipation of future error and holistic evaluation, could account for the finding. Spiller and Ariely (2020) focused entirely on a subset of multi-option alternatives: media of exchange (e.g., gift cards and promotional credit). Much like Le Lec and Tarroux (2020), this paper used a subjective domain and largely focused on
WTP judgments. In both cases, the authors found the extent of undervaluation as assessed via WTP increases with the difference in value between options, consistent with a weighted averaging process.

How are these results reconciled with the findings regarding the inherent value of choice? An important feature of the options-increase-value literature cited above is that the focus is primarily on the notable presence or absence of options rather than on the choice set as a singular, holistic entity. Brehm (1966) focuses on reactance to the restriction of freedom when choice options are removed; Mochon (2013) focuses on the heightened tendency to search for more options (rather than choosing one from an available set) when only a single option is available; and Shin and Ariely (2004) consider the desire to avoid losing options when the options have the potential to disappear. In none of those cases is the choice set evaluated as a single cohesive unit, which is our focus. Bown et al. (2003) find that consumers are more likely to choose options that provide choice, but interestingly, in the one case that aligns most closely with our focus (their Study 2, in which a dominated lure is added to one option), they do not observe statistically significant evidence of higher choice proportion for the option that provides a choice. In contrast, in the bundling and options-decrease-value literatures, choice sets are considered as holistic entities, as they are in the cases on which we focus. Thus, based on the most similar prior literature (e.g. Le Lec and Tarroux 2020; Spiller and Ariely 2020), we propose and test two key hypotheses:

H1: Consumers are less likely to choose a multi-option alternative than they are to choose its higher-valued option (over another, fixed alternative).

H2: The reduction in choice share proposed in H1 increases with the difference in value between the two options in the multi-option alternative.

In the present paper, we extend the literature in multiple ways. First, in contrast to Le Lec and Tarroux (2020) and the primary focus of Spiller and Ariely (2020), we use choices to document the effect. Such choices are arguably more common in consumers’ daily life than WTP ratings and sometimes exhibit qualitatively different effects (Grether and Plott 1979; Lichtenstein and Slovic 1971; Amir and Ariely 2007; O’Donnell and Evers 2019). In the current case, this is particularly important to examine as choices may encourage consumers to peer down the decision tree to possible outcomes whereas WTP
judgments may encourage them to value a choice set holistically. Second, we document the robustness of
the effect across six experiments (three in the main text, three in the supplement) in two domains,
including one domain with objectively dominant and dominated component options, expanding the
findings beyond preferential choices in which where uncertainty over future preference states may play a
larger role. Third, we rule out several possible explanations with additional experiments and features
including examining trinary choices. Finally, we use a process tracing method (mouse-tracking) in order
to connect the information acquisition process to the choices made about multi-option alternatives.

**Process Tracing**

This final feature (mouse-tracking) substantively contributes to our understanding of the decision
process and enables connections to other work in neuroeconomics. As described above, there are multiple
reasons why people might value multi-option alternatives differently from what normative theory
predicts. A possible mechanism for these shifts in valuation (and ultimately, the shifts in choices) is
attention. In binary choices, people tend to choose the option they have looked at longer (Armel,
Beaumel, and Rangel 2008; Atalay, Bodur, and Rasolofiarison 2012; Chandon et al. 2009; Fiedler and
Glöckner 2012; Fisher 2021; Krajbich, Armel, and Rangel 2010; Lim, O’Doherty, and Rangel 2011;
Mormann et al. 2012; Newell and Le Pelley 2018; Pachur et al. 2018; Päramets et al. 2015; Pieters and
Warlop 1999; Vaidya and Fellows 2015). Moreover, consumers tend to update their values and choose in
line with the choice attribute that they focus on (Busemeyer and Townsend 1993; Busemeyer and
Diederich 2002; Roe, Busemeyer, and Townsend 2001; Dai and Busemeyer 2014; Kim, Seligman, and
Kable 2012; Fisher 2017; Cohen, Kang, and Liese 2017; Smith and Krajbich, 2018). The existing
attention-informed decision models have been extended to account for the relationship between attention
and choice in situations with more than two alternatives (Krajbich et al. 2011; Gluth et al. 2020; Thomas,
Molter, and Krajbich 2020). Use of mouse-tracking and information-search paradigms has successfully
connected information acquisition and attentional patterns to consumer choice (Payne 1976; Johnson et al.
1989; Payne, Bettman, and Johnson 1988; Konovalov and Krajbich 2020; Stillman, Krajbich, and
Ferguson 2020; Reeck, Wall, and Johnson 2017; Stillman, Shen, and Ferguson 2018; Diehl 2005).
Altogether, this body of literature suggests that the information acquisition and attention processes in this environment may inform the decision processes involved in the undervaluation of multi-option alternatives. More specifically, because attention predicts choice in binary and trinary choice environments, we propose and test the following hypothesis:

H3: Undervaluation is correlated with the difference between the relative time spent on the multi-option alternative and the relative time spent on its highest-valued option in simple binary choice.

Measuring Undervaluation

Though the concept of undervaluation is straightforward, finding an efficient, effective way to measure it within-subjects is decidedly less so. Eliciting willingness-to-pay (WTP) or liking ratings of single- and multi-option alternatives, for instance, would be a methodologically-simpler way to investigate the phenomenon. However, ratings are less-than-ideal for several reasons. First, they are not as common in the real world as choices; although they provide insight about value, we often care about that value because of how it informs choices. Second, there are several well-documented cases in which ratings or WTP judgments deviate from choices (e.g., Grether and Plott 1979; Lichtenstein and Slovic 1971; Amir and Ariely 2007; O’Donnell and Evers 2019). In the present case, this could plausibly come about if choice increases the likelihood that consumers spontaneously consider the second stage choice (e.g. \( M_H \) vs. \( M_L \)). Therefore, we measure undervaluation both within- and between-subjects with a series of carefully constructed choices.

In all experiments, we examine undervaluation by comparing choice shares in two types of decisions: test decisions and control decisions. In the test decisions, participants choose between a single-option alternative (S) and a multi-option alternative (M). Within the multi-option alternative M, there are two component options: \( M_H \) and \( M_L \), where the value of \( M_H \) is greater than the value of \( M_L \). In the control decisions, participants choose between the same single-option option (S) and the a priori determined higher-valued component option (\( M_H \)). Undervaluation occurs when the proportion of choices of S is greater in the test decisions than in the control decisions. In other words, when participants choose M less
often than they choose $M_H$ (relative to the same option $S$), this demonstrates that participants undervalue $M$ relative to its best component option ($M_H$).

Across six experiments (three in the main text, and three replications in the supplement), we document the undervaluation phenomenon and investigate the underlying mechanism. In Experiment 1, we find evidence of undervaluation (H1) in a consumer domain with subjective values while examining a potential moderator: the strength of the preference for $M_H$ over $M_L$. We find stronger undervaluation as the difference between $M_H$ and $M_L$ increases (H2). In Experiment 2, we move into a more controlled domain (incentivized gambles) with objective — rather than subjective — dominance of $M_H$ over $M_L$. In Experiment 3, we add mouse tracking to the gamble choices in order to better understand the relationship between information acquisition and the undervaluation phenomenon (H3). We also introduce a potential moderator: the transparency of this dominance. In line with previous literature (Spiller and Ariely, 2020), we find a reduction of the effect when the dominance is more transparent. Across all experiments, we find strong, consistent evidence of undervaluation.

In the supplements, we describe additional experiments in which we investigate (and rule out) possible alternative explanations (e.g., inattention). We also examine several individual difference measures (e.g., risk aversion) to examine relationships with undervaluation. For a comparison of the methods across all experiments, see Table S1 in the supplements.

**Experiment 1**

In Experiment 1, we sought to examine the effect of interest (undervaluation of multi-option alternatives) in a relevant consumer domain. We also sought to investigate a potential moderator of the effect. Specifically, we investigated whether the subjective value difference between the best and worst component options is related to undervaluation, consistent with an averaging process.
Method

Participants. For this preregistered experiment (https://aspredicted.org/blind.php?x=j3ex33), we collected responses from 305 Amazon Mechanical Turk workers. They earned $1.25 for their participation.

Materials and Procedure. First, participants rated 50 well-known films on a scale of 0-10 (measured in increments of 0.1 via a slider) regarding how much they wanted to watch each of them. If a participant had not heard of a film, they could click a box labeled “I’ve never heard of this movie” (Figure 1a).

![Film Ratings](image)

Figure 1. Experiment 1 design. (a) First, participants rated their desire to watch each of 50 films. (b) Then, they made 30 hypothetical choices between two theaters, each of which was showing one or two films. Participants were told to imagine that if they chose a two-film theater, they would get to choose one of the two films to watch.

Next, participants were asked to imagine that they were planning to go see a movie and were asked (on each of 30 trials) to choose which of two hypothetical theaters they would go to (Figure 1b).
Each of the theaters was described as having either one or two movies playing. Participants were told that if they chose a theater with two movies, they would get to choose one of the two movies to see. Participants completed comprehension questions before the choices to ensure that they understood that (1) the choices were hypothetical, (2) that they would get to choose one (and only one) movie to watch, even if they chose a theater with two movies, and (3) that their choices would not influence the number of choices that they would have to make.

The 30 choices presented to participants were randomly generated for each participant. They fell into 3 categories (with 10 choices in each category). The first category of choices were test choices. In each of these trials, participants chose between a theater with one available movie (i.e. movie S) and a theater with two available movies (i.e. movies $M_{H}$ and $M_{L}$, where the participant’s prior rating of $M_{H}$ was greater than the rating of $M_{L}$). Each test choice trial comprised three unique movies, which did not overlap between trials. (Thus, the 10 test choices consisted of 30 different movies.) These movies were drawn randomly from the rated movies (i.e. the movies for which participants submitted a rating rather than clicking the “I’ve never heard of this movie” button). If a participant rated fewer than 30 films, then we generated as many trials as possible from their rated films before drawing from the unrated films.

The second category of choices were control binary choices. These trials consisted of a choice between two theaters, each with one available movie. Importantly, each of the 10 control binary choices paralleled one of the test choices. Specifically, each control binary choice was a choice between the single-movie theater and the a priori higher-rated film from the two-movie theater from the paired test choice. In other words, using the test choice example above (movie S vs. movies $M_{H}$ or $M_{L}$), the control binary choice would be movie S vs. movie $M_{H}$ (where movie $M_{H}$’s rating was higher than movie $M_{L}$’s rating prior to any choices).

Finally, the last 10 trials were filler trials, each of which comprised two single-movie theaters. The filler trial films were randomly selected from the rated films (independently from the selection process for the test trials, above). If a participant rated fewer than 20 films, then we sampled from the
unrated films when necessary. These trials were used for our preregistered exclusion criteria and to add some variability to the choices being made by participants.

It is important to note that before making any choices (after the ratings), participants demonstrated their understanding of the task by completing several comprehension check questions, including a question about the meaning of the test choices, i.e. that choosing “M_H or M_L” implied that they would get to choose which movie from the set \{M_H, M_L\} they wanted to see. They were required to get all questions correct before moving on to the choices.

**Exclusions and Data Preprocessing.** As specified in our preregistration, we excluded anyone who failed to rate at least 20 films, as they would have fewer usable test and control choices and thus noisier estimates. We also excluded anyone whose filler choices were not directionally predicted by their ratings in a logistic regression of ChooseLeft on (RatingLeft-RatingRight) as that indicates lack of minimal attention during ratings, choice, or both. These criteria resulted in the exclusion of 37 participants, leaving us with a final sample size of 268.

Also in line with our preregistration, we excluded any test-binary choice pairs that were generated from unrated films. For instance, if a participant only rated 24 films, then we were only able to generate eight valid test-binary choice pairs. The other two test-binary choice pairs would be generated from unrated films, and thus, excluded from analysis.

**Results**

To test for undervaluation, for every participant, we calculated S > \{M_H, M_L\} as (Chose S in test choices / number of valid test choices). We also calculated S > M_H as (Chose S in control choices / number of valid control choices). We tested (S > \{M_H, M_L\}) – (S > M_H) using a one-sample t-test and found evidence for undervaluation, \(M = 0.04, 95\% \text{ CI} = [0.02, 0.06], t(267) = 3.73, p < .001\). In other words, choice of M was 4 percentage points higher than choice of M_H (see Figure 2 for choice proportions of M and M_H, Table 1 for mean effect sizes across all experiments, and Table 2 for descriptives across all experiments). Using that approach, we find that 40% (106/268) of participants exhibited undervaluation overall, with 34% (91/268) exhibiting no difference and 26% (71/268) exhibiting an effect in the opposite
direction. Note that the 34% exhibiting no difference were not perfectly consistent, as in many cases it reflected multiple offsetting inconsistencies, implying that with more choices, we expect the proportion of people exhibiting no difference would diminish.

Figure 2. Main effect of undervaluation across experiments. Participants are less likely to choose “M_H or M_L” than they are to choose “M_H” (when compared to the same single-option alternative “S”). Bars indicate s.e.m. across participants.

Table 1. Main effect of undervaluation across all experiments.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Main effect</th>
<th>95% CI</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$M = 0.04$</td>
<td>[0.02, 0.06]</td>
<td>$t(267) = 3.73$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>1b (supplements)</td>
<td>$M = 0.04$</td>
<td>[0.03, 0.05]</td>
<td>$t(523) = 6.75$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>2</td>
<td>$M = 0.05$</td>
<td>[0.03, 0.08]</td>
<td>$t(240) = 5.04$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>2b (supplements)</td>
<td>$M = 0.06$</td>
<td>[0.03, 0.09]</td>
<td>$t(193) = 4.28$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>2c (supplements)</td>
<td>$M = 0.06$</td>
<td>[0.02, 0.09]</td>
<td>$t(175) = 3.37$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>3</td>
<td>$M = 0.06$</td>
<td>[0.04, 0.08]</td>
<td>$t(208) = 5.64$</td>
<td>$p &lt; .001$</td>
</tr>
</tbody>
</table>

Because $M_H$ and $M_L$ were randomly selected (and thus, the difference between ratings varied across choices and participants), we are able to examine the role of $M_H - M_L$ preference strength (i.e. how preferred $M_H$ is to $M_L$ for that choice for that participant) in undervaluation. We used the variable Choice Difference as our outcome measure, defined as (Chose S in test – Chose S in binary) at the triplet-level,
thereby taking a value of +1, 0, or -1. We regressed Choice Difference on the rating of S, the summed rating of $M_H$ and $M_L$, and the difference in rating between $M_H$ and $M_L$, clustering the standard errors at the subject level (eq. 1).

\[
Choice\ Difference_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 (M_H + M_L)_{it} + \beta_3 (M_H - M_L)_{it}
\]  

We find significant evidence for an effect of $M_H$-$M_L$ preference strength on undervaluation, $b_3 = 0.019$, $SE = 0.003$, $p < .001$ (Figure 3), such that participants exhibited greater undervaluation (greater choice differences) for multi-option alternatives with greater rating differences.

We replicated each of these results in a direct replication with 604 participants, 524 after exclusions. On average, participants again exhibited undervaluation ($M = 0.04$, 95% CI = [0.03, 0.05], $t(523) = 6.75$, $p < .001$), with 41% exhibiting undervaluation, 37% exhibiting no difference, and 22% exhibiting an effect in the opposite direction. We similarly find moderation by preference strength ($b = 0.011$, SE = 0.002, $p < .001$). Additional details are available in the supplements.
Figure 3. Relationship between (M_{H}-M_{L}) preference strength and degree of undervaluation. Undervaluation increases as M_{H} gets progressively better than M_{L}. Analysis reported in text controls for rating of S and sum of ratings of M_{H} and M_{L}. Bars represent s.e.m. across participants. See the supplements for a plot with data overlaid (Figure S2).

**Experiment 2**

In Experiment 1, we established the main effect and the moderation by M_{H}-M_{L} preference strength in a common consumer domain with familiar products in preferential choice. In the following experiments, we switch to the domain of risky choice, using card draws, die rolls, and coin flips. This new domain enables greater control, incentive-compatibility, and testing for cases of objective dominance, while maintaining a similar experimental design. In addition, these new domains reduce the likelihood that participants drew inferences about, for example, theater quality or the context of the consumption experience based on the set of movies shown. Furthermore, the objective dominance in this domain further reduces the likelihood that participants undervalue a multi-option alternative due to uncertainty about future preferences.

**Method**

**Participants.** For this preregistered experiment ([https://aspredicted.org/blind.php?x=ib4vi9](https://aspredicted.org/blind.php?x=ib4vi9)), we collected responses from 304 Amazon Mechanical Turk workers. They earned $1.75 (the first 20
participants) or $1.90 (the remaining 204 participants) for their participation. Five randomly-selected participants also received the outcome of one of their decisions, as detailed below.

**Materials and Procedure.** This experiment was similar to Experiment 1, with the following changes. First, choice options were gambles, as noted above and depicted in Figure 4. These gambles were drawn from a large pool of potential gambles from various combinations of drawing cards of different suits, flipping coins or pairs of coins, or rolling different numbers on a die. Binary choices, test choices, and trinary choices (described below) were matched on winning dollar amounts and probabilities, but could vary (non-systematically) in terms of the specific mechanism used as shown in Figure 4. Second, participants did not rate any options, as we could model each option’s value in terms of its payout and probability. Third, in addition to test choices (e.g. choosing between S and a multi-option alternative of M_H or M_L) and binary control choices (e.g. choosing between gambles S and M_H), participants also made trinary choices (e.g. choosing between gambles S, M_H, and M_L). Therefore, in this experiment, participants made 34 incentivized choices (including three attention check questions and one M_H or M_L choice from a randomly-selected test choice).

At the end of data collection, we randomly selected five participants and then randomly selected one of their choices to play out. Participants were aware of this structure. We opted to reward one trial instead of all of the trials to avoid stockpiling/strategy variability across trials, as described in Juechems et al. (2017). In this experiment, three of the five participants won money ($6, $7, and $10). For additional information on the trial generation process, see the supplements.

As in Experiment 1, participants completed several comprehension check questions, including a question about the meaning of the test choices, i.e. that choosing “M_H or M_L” implied that they would get to choose an alternative from the set \{M_H, M_L\}. They were required to get all questions correct before moving on to the choices.
Binary

$2
Drawing any card except a spade

$5
Flipping tails on one coin

Test

$2
Drawing any card except a heart

$5
A) Flipping heads on one coin
B) Flipping two heads on two coins

Trinary

$2
Drawing any card except a club

$5
Flipping heads on one coin

$5
Drawing a club

Figure 4. Experiment 2 design. Participants made incentive-compatible binary, test, and trinary choices. Each choice option comprised a potential dollar amount to be won and a probabilistic event (or two). If participants chose a multi-alternative option (and that trial was selected for payment), their payout was determined by their choice between the two alternatives.

Exclusions and Data Preprocessing. As specified in our preregistration, we excluded anyone who picked an obviously dominated option in any of the three attention-check questions. This resulted in the exclusion of 63 participants, leaving us with a sample size of 241.

Results

Preregistered Results. To test for undervaluation, for every participant, we calculated $S > \{M_{H}, M_{L}\}$ as (Chose $S$ in test choices / Number of test choices). We also calculated $S > M_{H}$ as (Chose $S$ in binary choices / Number of binary choices). We tested $(S > \{M_{H}, M_{L}\}) - (S > M_{H})$ using a one-sample t-test and found evidence for undervaluation, $M = 0.05$, 95% CI = [0.03, 0.08], $t(240) = 5.04$, $p < .001$ (Figure 2). Choice of $M_{H}$ was 5 percentage points higher than choice of $M$ (Tables 1 and 2). Using that approach, we find that 46% (111/241) of participants exhibited undervaluation overall, with 34% exhibiting no effect and 20% exhibiting an effect in the opposite direction. As in Experiment 1, note that we would expect the number exhibiting no effect to decrease as the number of trials increases.
Because the probabilities of $M_H$ and $M_L$ were randomly selected, we are able to examine the role of $M_H$-$M_L$ dominance (i.e. how much better $M_H$ is than $M_L$) in undervaluation. We used Choice Difference as our outcome measure, defined as (Chose S in test choices – Chose S in binary choices) at the set-level. We regressed Choice Difference on the expected value of S, the summed expected values of $M_H$ and $M_L$, and the difference in expected values between $M_H$ and $M_L$, clustering the standard errors at the subject level (eq. 2).

$$Choice\ Difference_{it} = \beta_0 + \beta_1 EV_{S_{it}} + \beta_2 (EV_{M_H} + EV_{M_L})_{it} + \beta_3 (EV_{M_H} - EV_{M_L})_{it}$$

The coefficient on the difference in expected values was in the expected direction, though it was not statistically significant, $b_3 = 0.011, SE = 0.008, p = .16$ (Figure 5a).

**Exploratory Results.** One possible reason for the lack of significance in the regression above is that participants were less attuned to the expected values of the options and more attuned to the probabilities themselves, especially given how the gambles were depicted. To investigate this possibility, we turned to the binary choices. Using logistic regression, we regressed choice of S on the expected values of S and $M_H$, as well as the probabilities of S and $M_H$ and the monetary amounts associated with each option, clustering the standard errors at the subject level (eq. 3).

$$Choice_{it} = \beta_0 + \beta_1 EV_{S_{it}} + \beta_2 EV_{M_{H_{it}}} + \beta_3 MS_{it} + \beta_4 M_{M_{H_{it}}} + \beta_5 P_{S_{it}} + \beta_6 P_{M_{H_{it}}}$$

We find significant effects of both monetary amounts (representing the simple effects of monetary amount when probabilities are equal to 0; S: $b_3 = 0.19, SE = 0.06, p = .002; M_H$: $b_4 = -0.28, SE = 0.10, p = .003$) and both probabilities (representing the simple effects of probabilities when monetary amounts are equal to 0; S: $b_3 = 2.96, SE = 0.77, p < .001; M_H$: $b_6 = -2.83, SE = 0.91, p = .002$), even after accounting for the effects of expected value (S: $b_1 = 0.42, SE = 0.12, p < .001; M_H$: $b_2 = -0.24, SE = 0.15, p = .11$).

These results suggest that payout and probabilities contributed to evaluation of the gambles above and beyond their contribution to expected value. With these results in mind, we adapted our preregistered analysis above. Instead of regressing Choice Difference on expected values, we regressed it on
probabilities. Specifically, we regressed Choice Difference on the probability of S, the summed probabilities of M_H and M_L, and the difference in probabilities between M_H and M_L, with clustered SEs at the subject level. (eq. 4).

\[
\text{Choice Difference}_{it} = \beta_0 + \beta_1 P_{Si} + \beta_2 (P_{M_H} + P_{M_L})_{it} + \beta_3 (P_{M_H} - P_{M_L})_{it}
\] (4)

Here, consistent with our findings in Experiment 1, we find a marginally significant relationship between M_H-M_L dominance and Choice Difference \((b_3 = 0.11, SE = 0.06, p = .07, \text{Figure 5b})\).

Figure 5. Relationship between \((M_H > M_L)\) dominance strength and degree of undervaluation. (a) Dominance strength is defined using expected value. (b) Dominance strength is defined using probabilities. Across both definitions of dominance, undervaluation increases as M_H gets progressively better than M_L, though this relationship is not significant for expected value \((p = .16)\) and marginally significant for probability \((p = .07)\). Bars represent s.e.m. across participants. See the supplements for plots with data overlaid (Figure S3).

We also compared choices in the test choices (S vs. \{M_H, M_L\}) to choices in the trinary (S vs. M_H vs. M_L) choices (Table 2). Using a one-sample t-test of \((\text{Choose} \{M_H, M_L\} \text{ in test choices}) - (\text{Choose} M_H \text{ or } M_L \text{ in trinary choices})\), we find that participants choose the multi-option alternative \((M_H \text{ or } M_L)\) in test choices significantly less than they choose the same two options \((M_H; M_L)\) in the trinary choices, \(M = -0.08, 95\% \text{ CI} = [-0.10, -0.06], t(240) = -7.95, p < .001\) (see Figure S1). There was not a significant difference between choice for the multi-option alternative \((\{M_H, M_L\})\) in test choices and choice for the higher-valued option \((M_H)\) in trinary choice, \(M = 0.001, 95\% \text{ CI} = [-0.02, 0.03], t(240) = 0.10, p = .92\).
Participants chose $M_H$ significantly more in binary choice than they did in trinary choice ($M = 0.06$, 95% CI = [0.04, 0.07], $t(240) = 5.79, p < .001$).

**Replications**

We ran two additional, similar experiments (2b: N = 298, 194 after exclusions; 2c: N = 298, 176 after exclusions). On average, participants in both experiments exhibited undervaluation (2b: $M = 0.06$, 95% CI = [0.03, 0.09], $t(193) = 4.28, p < .001$; 2c: $M = 0.06$, 95% CI = [0.02, 0.09], $t(175) = 3.37, p < .001$). In the first replication, we find that 49% (96/194) of participants exhibited undervaluation overall, 20% (38/194) showed no difference, and 31% (60/194) exhibited an effect in the opposite direction. In the second replication, we find that 55% (96/176) of participants exhibited undervaluation overall, 19% (33/176) showed no difference, and 27% (47/176) exhibited an effect in the opposite direction. Additional details are available in the supplements.

**Experiment 3**

In Experiment 3, we introduce a process-tracing measure to our design: mouse-tracking. We measure the information acquisition process during all choices and connect it to our main effect. We also manipulated whether the $M_H$ vs. $M_L$ dominance relationship was high-transparency (easy to identify) or low-transparency (difficult to identify). The monetary outcomes and the probabilities of winning were held constant between high-transparency pairings and low-transparency pairings, but in the high-transparency pairings, $M_H$ always contained the event $M_L$ and in low-transparency pairings, $M_H$ never contained the event $M_L$. For example, a high-transparency pairing might be $M_H =$ “$2 if you roll a 1 or 2 or 3” and $M_L =$ “$2 if you roll a 1,” while a low-transparency pairing might be $M_H =$ “$2 if you draw a black card” and $M_L =$ “$2 if you roll a 4.” We hypothesized that a more (vs. less) transparent dominance relationship would result in a lower (vs. higher) degree of undervaluation as the implications for the second stage choice of $M_H$ vs. $M_L$ are more immediately apparent in the high-transparency pairings.
Method

Participants. For this preregistered experiment (https://aspredicted.org/blind.php?x=8tg3fe) we collected responses from 302 Amazon Mechanical Turk workers. They earned $2.50 (the first 20 participants) or $2.75 (the remaining 282 participants) for their participation. Five randomly-selected participants received the outcome of one of their decisions, as detailed below.

Materials and Procedure. This experiment was similar to Experiment 2, with the following changes. The biggest differences between this experiment and the prior experiments are that (1) the trials were not randomly generated at the subject-level and instead came from a predetermined set, (2) the $M_H$ vs. $M_L$ dominance relationship was manipulated to be either high- or low-transparency as detailed above, (3) the information (i.e. the details of $S$, $M_H$, and/or $M_L$) was not visible unless the participant hovered their cursor over the gamble, (4) $M_H$ and $M_L$ were presented horizontally instead of vertically, and (5) we tracked participants’ mouse movements in this experiment. Specifically, while participants made the choices described above, we recorded the order in which participants viewed each piece of information (e.g. $S$, $M_H$, and $M_L$) and the durations of these information acquisitions in a MouseLab-like paradigm (Johnson et al. 1989) (Figure 6). In this experiment, participants made 65 incentivized choices. These choices included (1) high-transparency test choices: $S$ vs. $M_H$ or $M_L$ (H), (2) low-transparency test choices: $S$ vs. $M_H$ or $M_L$ (L), (3) high-transparency control choices: $S$ vs. $M_H$ (H), (4) low-transparency control choices: $S$ vs. $M_H$ (L), (5) trinary choices: $S$ vs. $M_H$ vs. $M_L$, and (6) binary choices: $M_H$ vs. $M_L$. At the end of the survey, participants completed 10 rank-ordering choices (of $S$, $M_H$, and $M_L$). Please see the supplements for more specific trial information. In this experiment, all five randomly-selected participants won money ($3, $5, $4, $2 and $4).

As in the previous experiments, participants completed several comprehension check questions, including a question about the meaning of the test choices, i.e. that choosing “$M_H$ or $M_L$” implied that they would get to choose an alternative from the set {$M_H, M_L$}. They were required to get all questions correct before moving on to the choices.
**Exclusions and Data Preprocessing.** As specified in our preregistration, we excluded anyone who picked the obviously dominated option in any of the three attention-check questions and we excluded anyone who did not mouse-over the boxes in the instructions as instructed. This resulted in the exclusion of 93 participants, leaving us with a sample size of 209.

For the mouse-tracking data, we converted the hover-times (i.e. the times that participants spent hovering over the available information) into proportions from 0 to 1 at the trial level.

![Figure 6. An example of a test choice in Experiment 3. Information about each probabilistic event (S, M\textsubscript{H}, or M\textsubscript{L}) was hidden until the participant moved their mouse to the box containing the information. Participants were allowed to look at the information as many times and for as long as they wanted before making their decision.](image)

**Results**

**Preregistered Behavioral Results.** We tested for undervaluation using the method in Experiment 2, and we find evidence for undervaluation, $M = 0.06$, 95% CI = [0.04, 0.08], $t(208) = 5.64$, $p < .001$ (Figure 2). Choice share of M was 6 percentage points lower than choice share of M\textsubscript{H}. To test for an effect of transparency, for every participant, we calculated $S > \{M\textsubscript{H}, M\textsubscript{L}\}$ (H) and $S > \{M\textsubscript{H}, M\textsubscript{L}\}$ (L) corresponding to high-transparency and low-transparency test trials using the approach above. We also similarly calculated $S > M\textsubscript{H}$ (H) and $S > M\textsubscript{H}$ (L). We tested for moderation by transparency by analyzing ($S > \{M\textsubscript{H}, M\textsubscript{L}\}$ (H) – $S > M\textsubscript{H}$ (H)) – ($S > \{M\textsubscript{H}, M\textsubscript{L}\}$ (L) – $S > M\textsubscript{H}$ (L)) using a one-sample t-test, and we
find significant evidence of the moderation, $M = -0.03$, 95% CI = $[-0.06, -0.01]$, $t(208) = -2.76$, $p = .006$. That is, the degree of undervaluation is smaller when the transparency of $M_H > M_L$ is high. This moderation indicated attenuation, though not elimination, as we still find evidence of undervaluation even when the dominant relationship is transparent; $M = 0.04$, 95% CI = $[0.02, 0.06]$, $t(208) = 3.70$, $p < .001$. We investigate this finding in more detail in the exploratory results below.

Using that approach, we find that 61% (127/209) of participants exhibited undervaluation overall (46% (97/209) exhibited undervaluation on high-transparency trials and 59% (123/209) exhibited undervaluation on low-transparency trials); of the remaining 39%, 10% (20/209) exhibited no effect and 29% (62/209) exhibited an effect in the opposite direction. As in prior experiments, we expect the percentage exhibiting no effect would decrease with more choices. At the choice set level, we observe undervaluation for 9/10 of the choice sets overall, 8/10 of the high-transparency choice sets, and 9/10 of the low-transparency choice sets.

**Preregistered Mouse-Tracking Results.** We tested for a relationship between aggregate information acquisition (i.e. mouse movements) and undervaluation. For each subject, we computed the difference in average proportion of time spent on S between test and control trials (i.e. average proportion spent on S in S vs. $\{M_H, M_L\}$ choices – average proportion spent on S in S vs. $M_H$ choices). We expected to find a positive correlation between-subject between this average proportion mouse difference and degree of undervaluation, and we did: $r = 0.27$, $t(207) = 4.04$, $p < .001$. Participants who spent relatively more time inspecting S when paired with M (than when paired with $M_H$) undervalued M more.

We also tested for two trial-level associations between mouse movements and choice. First, across all binary, test, and trinary choices, we regressed choice (of S) on the proportion of mouse-hover time spent on S, with random intercepts and slopes at the subject level. Indeed, we find a significant relationship between information acquisition and choice within-subject, $b = 4.79$, $SE = 0.18$, $p < .001$. Second, we regressed choice for S (in S vs. $\{M_H, M_L\}$ choices only) on the proportion of time spent on $M_H$ relative to $M_L$ (i.e. time spent on $M_H$ / total time spent on $M_H$ or $M_L$). We find a significant negative relationship ($b = -1.70$, $SE = 0.24$, $p < .001$), which implies that the more time participants spent
looking at $M_H$ (i.e. the better outcome out of the $M_H$ or $M_L$ option), the less likely they were to choose $S$ (i.e. the more likely they were to choose $M_H$ or $M_L$) (Figure 7).

Finally, we examined mouse-tracking differences between trinary and multi-option alternative choices. First, we examined the difference between average proportions of time spent on $S$, $M_H$, and $M_L$ in each trial type (i.e. trinary vs. test) and compared these differences to undervaluation. No correlations were significant (all $p$s > 0.1). Additionally, we looked at the difference in average proportion of transitions between $M_H$ and $M_L$ in the trinary trials compared to the test trials (i.e. number of transitions between $M_H$ and $M_L$ / total number of transitions). We did not find any significant correlations there, either (all $p$s > 0.1).

![Figure 7. Relationship between mouse-tracking and choices in test choices.](image)

**Figure 7.** Relationship between mouse-tracking and choices in test choices. As participants spent more time on $M_L$ (relative to $M_H$), they were less likely to choose the multi-option alternative ($M_H$ or $M_L$), $b = -1.70$, $SE = 0.24$, $p < .001$. Bars represent s.e.m. across participants.

**Exploratory Results.** An advantage of mouse-tracking (and other process tracing measures) is the ability to look further into the decision process than simple choice outcomes (i.e. “participants chose option A”) allow. In this experiment, the significant transparency moderation lends itself particularly well
to investigation through the lens of process tracing. Thus, we compared the proportion of time spent on M_L (relative to M_H) in the high-transparency trials to the proportion of time spent on M_L (relative to M_H) in the low-transparency trials. We found a small but significant difference, such that subjects spent relatively less time on M_L when the transparency was higher, $M = 0.01$, 95% CI = [0.002, 0.02], $t(207) = 2.46, p = 0.01$. Moreover, this subject-level difference in hover-times correlates with the subject-level difference in undervaluation between high- and low-transparency trials, $r = 0.13$, 95% CI = [–0.008, 0.26], $t(206) = 1.86, p = .06$. In other words, as subjects spent more relative time on M_L in low-transparency trials (vs. high-transparency trials), they showed a larger degree of undervaluation in low-transparency trials (vs. high-transparency trials).

As in Experiment 2, we also compared choices in the test choices (S vs. {M_H, M_L}) to choices in the trinary (S vs. M_H vs. M_L) choices (Table 2). Using the same method as in Experiment 2, we find that participants choose the multi-option alternative ({M_H, M_L}) in test choices significantly less than they choose the same two options (M_H; M_L) in the trinary choices, $M = –0.07$, 95% CI = [–0.09, –0.05], $t(208) = –6.77, p < .001$ (fig. S2). Moreover, in this experiment, participants choose the multi-option alternative (M_H or M_L) in test choices significantly less than they choose the higher-valued of the two (M_H) in trinary choice, $M = –0.02$, 95% CI = [–0.05, –0.001], $t(208) = –2.09, p = .04$. Also, like in Experiment 2, participants choose M_H significantly more in binary choice than in trinary choice, $M = 0.04$, 95% CI = [0.02, 0.05], $t(208) = 3.62, p < .001$.

Although our main effect of interest (undervaluation) is statistically significant, and the presence of undervaluation is widespread across participants, there is ample variability in the size of the effect across participants ($M = 0.06$, $SD = 0.15$). Evidence suggests that the noisiness of the effect is due (at least in part) to individual differences. Specifically, the degree of undervaluation in low-transparency trials is significantly correlated with the degree of undervaluation in high-transparency trials, $r = 0.50$, 95% CI = [0.39, 0.59], $t(207) = 8.30, p < .001$. In other words, if a participant displayed a high degree of undervaluation on one half of the trials, they also displayed a high degree of undervaluation on the other
half, which is consistent with individual differences in the tendency to exhibit undervaluation, rather than a uniform extent of undervaluation combined with pure noise in the choice data.

Unlike Experiments 1 and 2, there was a limited number of unique choices and little variability in the relative dominance of $M_H$ over $M_L$ in Experiment 3 across choice sets, so that test had very low power and we do not discuss it here.

To connect these results to other findings in the decision-making literature, we also examined the relationship between decision difficulty (i.e. how similar the options are in value) and response times (RTs). We did not have a direct measure of RTs, so we used total time spent hovering over any of the options as a proxy. Past literature (cites) has demonstrated a positive relationship between response times and decision difficulty, such that the most difficult decisions take the longest to make, on average. Here, we observe a similar effect in our data. We regressed logged RTs on the absolute difference in expected value between the options. For multi-option alternative choices, we set the expected value of the $M_H$ or $M_L$ option to be the expected value of $M_H$, and for trinary choices, we set the absolute expected value difference to be the maximum expected value minus the second-highest expected value (as in Krajbich et al. 2011). In this data, we find that as absolute expected value difference increases (i.e. decisions become easier), the choices become faster, on average, $b = -0.06$, $SE = 0.02$, $p < .001$.

**General Discussion**

Across six experiments in two domains, we document a consistent effect: decision-makers undervalue multi-option alternatives. More specifically, they are less likely to choose a multi-option alternative “$M_H$ or $M_L$” (where the value of $M_H$ exceeds the value of $M_L$) than they are to choose a single-option alternative “$M_H$.” This behavior is non-normative from both expected value and expected utility theory perspectives. Moreover, we find that the strength of undervaluation is related to the difference in values between $M_H$ and $M_L$: as the difference in values increases, undervaluation increases as well. However, when the dominance of $M_H$ over $M_L$ is more transparent or obvious, then the degree of
undervaluation decreases. In replication experiments (available in the supplements), we provide evidence against multiple alternative explanations (noisy responding; strong delayed choice aversion). Finally, we find process evidence which suggests that undervaluation is strongly associated with information acquisition patterns.

**Related Work and Alternative Explanations**

These findings connect to work on agenda effects (Plott and Levine, 1978; Tversky and Sattath 1979; Hauser 1986), in which the order in which subsets of decisions are made influences the option that is ultimately chosen. However, this prior literature does not offer explanations for the results we have discussed here, and therefore, this undervaluation phenomenon appears to be one further instance in which agenda effects matter. Moreover, in contrasts to prior work on agenda effects, we have (at most) one multi-option alternative per decision (rather than two or more, which make the choice more complex and plausibly susceptible to non-normative influences), we do not have comparisons across multiple multi-option alternatives (as in agenda effects and assortment choice research) and instead focus on how participants compare multi-option and single-option alternatives.

There are several literatures that address similar-yet-distinct phenomena, including assortment choice and multi-option (i.e., more than 2 options) choice. However, these literatures ultimately do not address the question at hand. Assortment choice research focuses on the evaluation of assortments in and of themselves, but does not typically enable comparisons to a normative benchmark, nor does it compare evaluations of assortments to the evaluations of the constituent parts of the assortment in a choice context. With regard to multi-option choice research, our test choices (i.e., choices between a single-option alternative S and a multi-option alternative \{M_{hi}, M_L\}) are formally equivalent to trinary choices (i.e. choices between S, M_{hi}, and M_L). However, since participants are significantly less likely to choose \{M_{hi}, M_L\} (in a test choice) than they are to choose M_{hi} or M_L (in a trinary choice), it is clear that these two types of choices are neither psychologically nor practically equivalent.

We have addressed several potential alternative explanations with our data. First, we rule out the possibility that our results are due to noisy or poor responding by showing that the effect is robust to
multiple attention checks and comprehension questions; if anything, the effect is larger among participants who show signs of being more attentive. We also rule out delayed-choice aversion as a possible explanation in two ways: (1) participants understand that they will make the same number of decisions in the experiment, regardless of which decisions they make, and (2) we find that participants choose \{M_{HI}, M_L\} more often than they choose M_L when pitted against S, so they do not have a strong aversion to delayed choices. See supplement for details.

Within the literature on assortment and multi-option choice, a natural point of comparison is Sood et al. (2004)’s findings that lone options are more likely to be chosen than grouped items when grouped items are initially compared to one another. This is because the intra-group comparisons lead to focusing on downsides of grouped options more than single options. Yet in the current examination in the risky domain, there are no tradeoffs among the grouped options, suggesting such comparative loss aversion does not explain these results. Another point of comparison is the model proposed by Kahn and Lehman (1991). This model suggests that the value of an assortment will decrease as the number of unacceptable items in the assortment increases. However, our results are not dependent on this characterization. Indeed, M_L is not preferred to M_{HI}, but it is preferable to receiving nothing, so it is not an unacceptable item. Applied to that model, our findings suggest it may extend beyond unacceptable items and into acceptable-but-less-desirable items.

**Marketing Implications**

The present research has a number of implications for marketing managers. First, it reinforces the importance of understanding the choice environment as construed by consumers. The same (e.g., trinary) choice may be construed as a two-stage decision by some consumers and as a single-stage decision by others. The current findings shed light on how dominated options differentially influence choice share depending on whether they are considered as options within multi-option alternatives or isolated alternatives. Thus, understanding how consumers structure the choice environment will permit more accurate predictions of consumer choice.
In addition to enhancing predictions for consumer choice, it also enables structuring the choice environment with formal choice structures or informal choice cues to shape how the choice is framed. If marketers are able to establish a priori what a consumer would likely prefer within a given choice set, they may strategically structure a set to enhance options by excluding less-attractive options or plausibly reduce choice share by including less-attractive options. Of course, the extent to which such practices are likely to be effective will depend on the precision with which marketers are able to predict consumers’ choice shares; better prediction models will enable greater tailoring. Further, note that this structuring is at odds with what one might predict from research on the attraction effect (Huber, Payne, and Puto, 1982). It appears that whether the asymmetrically dominated option appears as a third option or as an additional component option within one alternative may reverse its effect on relative choice shares. More specifically, in a trinary choice setting, the attraction effect benefits M_H because M_H dominates M_L, and thus, the choice share for M_H increases. However, when M_H and M_L are framed as component options of M (as in all of our experiments), choice share for M_H (i.e. M) decreases, relative to the simple binary choice between S and M_H. In Experiment 3, we observed that M_L affected choice share of S relative to binary choice more in test choices than trinary choices. It remains to be seen whether adding a well-calibrated M_L as a third option to S and M_H would increase the share of M_H while adding M_L as an alternative component option of M_H would increase the share of S.

Finally, these findings reinforce the potential benefit for decision aids for consumers, and likely for managers, to encourage peering down the branches of the decision tree (Shafir 1994). As noted above, one approach to improve choice outcomes is to simply exclude options from consideration. Alternatively, decision aids may prompt consumers to consider second-stage choices when choosing a first-stage option (cf. Spiller and Ariely 2020). While our research thus far has focused on consumer decisions, it is plausible that similar effects manifest among managers when considering paths forward as well.

**Future Directions**

Several unanswered questions about this phenomenon remain. For instance, although our experiments find process evidence for the effect regarding attention to options, these results are
correlational. In order to establish a causal path from information acquisition to decisions in this domain, future research would need to manipulate the information acquisition. (There is evidence from other domains that attention is causally related to choices, but we do not present causal evidence here; see Gwinn et al. 2019; Armel et al. 2008; Mormann et al. 2012; Pärnamets et al. 2015). Another unanswered question is the exact method by which adding $M_L$ to $M_H$ to form the multi-alternative option \{\(M_H, M_L\)\} decreases choice probability. We contend this operates by decreasing the value of the set \{\(M_H, M_L\)\} rather than the value of $M_H$ itself. Does the evidence support this account? It appears to, as the addition of $M_L$ to the binary choice set, resulting in the trinary choice set, does not increase choice of S (whereas adding $M_L$ as a component option of M does increase choice of S). Finally, an interesting puzzle in the present paper is the moderation by transparency. This effect was strong and precisely estimated in Experiment 3, but was not significant in Experiments 2b and 2c reported in the supplement. Of course, there were several key differences between the supplemental experiments (2b-c) and the main text experiment (3). Most notably, information about the options was only visible when the participants moused over each box (via our MouseLab-esque programming) in Experiment 3, but not in the others. It is possible that the more directed information acquisition method altered the consideration process.

A promising avenue for future research would also include consideration of which consumers are more likely to exhibit stronger effects. We began to explore this in experiments reported in the supplements, though our initial findings did not yield diagnostic results. In particular, we hypothesized that consumers who elaborated more extensively on potential outcomes (Nenkov et al. 2009) or who use more analytic rather than holistic approaches to thinking through decisions (Choi et al. 2007) would exhibit lesser degrees of undervaluation. However, neither of these hypotheses held. Given the finding of systematic heterogeneity in Experiment 3, identifying the source of such heterogeneity would be informative.

**Conclusion**

The present research identifies and investigates a curious pattern of choice results: decision makers often choose against their best interests when confronted with a multi-alternative option. This
finding has important implications for consumers and practitioners. Offering a multi-alternative option may seem like a good idea to increase choice for one’s product(s) as it enables appeal to multiple segments of consumers with heterogeneous preferences, but if the alternatives within the option are discrepant in value, our results suggest that choice likelihood for that option may be lower than it would be otherwise for any given consumer. It is not clear from the present research whether knowledge of this tendency would reduce or reverse the undervaluation effect, but we contend it is important for marketing managers and choice architects to be aware of this heretofore underappreciated way in which consumers’ choices defy normative expectations.
References


https://doi.org/10.1177/0146167206298568

https://doi.org/10.1016/j.cogpsych.2017.08.001

https://doi.org/10.1037/a0035976

https://doi.org/10.1509/jmkr.2005.42.3.313

https://doi.org/10.3389/fpsyg.2012.00335

https://doi.org/10.1016/j.cognition.2017.06.007

https://doi.org/10.1287/mnsc.2020.3732

https://doi.org/10.1007/BF00435195

https://doi.org/10.1038/s41562-020-0822-0

https://doi.org/10.1017/CBO9780511618031.006


https://doi.org/10.1016/S0167-8116(98)00022-6


https://doi.org/10.1073/pnas.1707040114

https://doi.org/10.1037/0033-295X.108.2.370

https://doi.org/10.1038/s41598-019-38927-7

https://doi.org/10.1287/mnsc.1030.0148

https://doi.org/10.1287/mksc.13.1.23

https://doi.org/10.1037/xge0000482

https://doi.org/10.1086/383420
[https://doi.org/10.1016/j.obhdp.2020.07.005](https://doi.org/10.1016/j.obhdp.2020.07.005)


Weaver, K., Garcia, S. M., and Schwarz, N. (2012). The Presenter’s Paradox: Figure 1. *Journal of Consumer Research, 39*(3), 445–460. [https://doi.org/10.1086/664497](https://doi.org/10.1086/664497)