Decision makers undervalue multi-option alternatives in two-stage choice

Stephanie M. Smith*[a]
Stephen A. Spiller[b]

[a] Booth School of Business, University of Chicago, Chicago, IL 60637
[b] Anderson School of Management, UCLA, Los Angeles, CA 90095

*Stephanie M. Smith
5807 S Woodlawn Ave
Chicago, IL 60637
stephanie.smith@chicagobooth.edu

Abstract

Initial decisions lead to subsequent decisions. Dominated options in such downstream choices ought to be ignorable in the initial choice for even minimally forward-looking people. Across nine experiments in two domains (consumer goods and risky gambles), we find that adding a less-valuable option (i.e., a less-preferred consumer good or 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. Mouse-tracking reveals that participants 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 decision-making and how decision makers assess the overall value of choices.

Only rarely does a person'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: choosing a home implies choices among where to eat, where to shop, and who to visit. Other times these are explicit: choosing a restaurant implies choosing items from a menu and 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 festivals and conferences, are typically multi-option alternatives, as are game tokens at arcades 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, purchasing a car 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 this paper, we seek to better understand how people integrate across options in a choice set when deciding among multi-option alternatives. We find they regularly sacrifice a chance at maximizing utility by integrating the value of less-attractive options.

Value-Maximizing Decision Rules in Two-Stage Choices

To a value-maximizing individual, a multi-option alternative is worth at least as much as its most-valuable option, because they can ignore the less-valuable option, knowing that they will choose the more-valuable option instead. Formally, we consider the case in which an individual decides 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 in which M_H, "multi-high," is preferred to M_L, "multi-low"). If there is no uncertainty regarding the values of M_H or M_L, then the value of M is $\max\{M_H,M_L\}$. Since M_H is preferred to M_L, $\max\{M_H,M_L\} = M_H$. Thus, 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 than when 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 than when pitted against M_H. (For example, a consumer may generally prefer hot coffee (M_H) to iced coffee (M_L), but make an exception when the temperature exceeds 90° F[1].) In other words, adding M_L as an option to M_H should not decrease the probability of choosing M.

Relatedly, though the analysis above suggests that decision makers should value a multi-option alternative equivalently to the expected maximum of its constituent pieces, some findings from the behavioral literature suggest that people might value a multi-option alternative more than the maximum value of its component options because the multi-option alternative enables choice.[1-4] Because people often value the ability to choose, the added value from the presence of choice itself can lead to an overvaluation of multi-option alternatives.

Valuation of Sets

Alternatively, people might undervalue a multi-option alternative. If valuation processes for multi-option alternatives and bundles are similar, then people would undervalue multi-option alternatives because bundles of goods are valued according to a weighted average of their components' values.[5-9] There is related evidence that consumers estimate the value of a product

[1] Throughout, we consider cases in which the state of the world is exogenously determined: neither the choice set nor the choice affects the preference ordering among S, M_H, and M_L.

to be a weighted average of its features,[10,11] and adding a less-attractive bonus to a product decreases its value.[12] These "weighted averages" are not mathematical averages of numerical properties, but rather central tendencies of subjective values. People can rapidly extract average economic value from a set of products[13] and sometimes use these assessments inappropriately.[14,15] These findings suggest that adding a less-desirable option to a multi-option alternative decreases its value in the same way that adding a lower-valued component to a bundle decreases the bundle's value: the worse the option, the greater the decrease.

Prior Research on Multi-Option Alternatives

Therefore, there are two hypotheses regarding how people might misvalue multi-option alternatives relative to the value-maximizing benchmark. 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.[16,17] These papers focus on how assortments compare to one another and how assortment framing (as sets vs. alternatives) affects valuation. Therefore, this literature tends not to focus on how the value of a multi-option alternative compares to its highest-valued option. To our knowledge, only two prior papers have directly examined this latter phenomenon.[18,19] However, these papers leave many questions unanswered.

Le Lec and Tarroux (2020) tested the phenomenon in a single study in a domain in which the relative values of products are subjective. In this study, participants reported their willingness-to-pay (WTP) for multi-option alternatives regarding what websites to spend time on at the end of the experiment. The authors found evidence for undervaluation of multi-option alternatives, and proposed two explanations: anticipation of future error and holistic evaluation.

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 focused primarily on WTP judgments (though some supplemental studies examined choice). In both cases, the authors found the extent of undervaluation 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 a consumer's focus is primarily on the presence/absence of options rather than on the choice set as a singular entity. Thus, based on the most similar prior literature[18,19] and the arguments above, we propose and test two hypotheses:

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

H2: This reduction in choice share increases with the difference in value between the two component options in the multi-option alternative.

This paper contributes to the literature in multiple ways. First, in contrast to Le Lec and Tarroux (2020) and the primary results of Spiller and Ariely (2020), we use choices to document the effect. This distinction is important, as choices may encourage decision makers to peer down the decision tree to possible outcomes whereas WTP judgments may encourage holistic valuation of a choice set. Second, we document the robustness of the effect across nine online lab experiments (three in the main text, six in the supplement). These nine experiments span two domains, including one domain with objectively dominating/dominated component options, expanding the findings beyond preferential choices in which uncertainty over future preference states may play a larger role. Third, we rule out several possible explanations with additional experiments and features (e.g., trinary choices). Fourth, we examine the potential for transparency of dominance relationships to moderate the effect. Fifth, we document the effect as the number of component options in a multi-option alternative increases. Sixth, we use process tracing to connect the information acquisition process to choices. Finally, we provide evidence of mechanism: adding an inferior sub-option decreases the value of the holistic multi-option alternative, rather than just the superior sub-option.

Process Tracing

Our process tracing (mouse-tracking) substantively contributes to our understanding of the decision process and connects to other work in neuroeconomics. There are multiple reasons why people might under- or overvalue multi-option alternatives; a possible mechanism for these shifts in valuation (and choices) is attention. In binary choices, people tend to choose the option they have looked at longer.[20-27] Moreover, people tend to update their values and choose in line with the choice attribute that they focus on,[28-31] including in situations with more than two alternatives.[32,33] Use of mouse-tracking and information-search paradigms has successfully connected information acquisition and attentional patterns to choice.[34-38]

This body of literature suggests that the attention processes in this environment may inform the decision processes involved in the undervaluation of multi-option alternatives. Specifically, we propose and test the following hypothesis:

H3: Undervaluation is negatively correlated with the difference between: (i) the relative time spent on the multi-option alternative, and (ii) 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 WTPs of multi-option alternatives would be simpler, but 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 well-documented cases in which ratings or WTP judgments deviate from choices.[39-42] Therefore, we measure undervaluation within-subject in a series of carefully constructed choices.

In each of our experiments, we examine undervaluation by comparing choices in two types of decisions: test decisions and binary control decisions. In 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, between which the participant could make a future choice. In all cases, value(M_H) > value(M_L). In the binary control decisions, participants choose between the same single-option alternative (S) and the a priori determined higher-valued component option (M_H). Undervaluation occurs when the proportion of S choices is greater in test decisions than in control decisions. In other words, when participants choose M less often than they choose M_H, this indicates that participants undervalue M relative to its best component option (M_H).

Across three experiments, 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, and 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 dominance of M_H over M_L. In Experiment 3, we use mouse tracking in order to better understand the relationship between information acquisition and the undervaluation phenomenon (H3).

In the supplements, we describe six additional experiments in which we (1) replicate the main results, (2) extend the results beyond alternatives with only 1 or 2 sub-options, (3) provide supporting process evidence for the mechanism, (4) rule out possible alternative explanations (e.g., inattention), and (5) examine the relationships between undervaluation and several individual difference measures (e.g., risk aversion). For a comparison of the methods across all experiments, see Table S1.

Results

Experiment 1

In Experiment 1, we tested for undervaluation in a consumer domain and investigated whether the subjective value difference between the best and worst component options is related to undervaluation, consistent with an averaging process.

Participants first reported their liking for 50 movies. They then made choices from a series of choice sets between different one- or two-movie theaters (Figure 1). These choice sets included test (S vs. $\{M_H, M_L\}$) and control (S vs. M_H) sets, where movie M_H was rated higher than movie M_L in initial ratings by that participant.

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 control choices / number of 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% CI = [0.02, 0.06], $t(267) = 3.73$, $p < .001$. In other words, choice of M was 4 percentage points lower than choice of M_H (see Figure 2a; Table 1).

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 2c). Participants exhibited more undervaluation for multi-option alternatives when they had stronger preferences for M_H over ML, as assessed using pre-choice ratings. We replicated each of these results in a direct replication, Experiment 1b, detailed in the supplements.

In two additional experiments (1c and 1d) detailed in the supplements, we extended these results to a broader array of choice set sizes. Rather than just examining choices between two one-movie theaters (1v1) and choices between one one-movie theater and one two-movie theater (1v2), we also examined choices between one one-movie theater and one three-movie theater (1v3), two two-movie theaters (2v2), and one two-movie theater and one three-movie theater (2v3). This enabled us to consider the effect of adding a marginal option beyond the 1v1 set. Due to statistical power and measurement error concerns (note the x-intercept in Figure 2c), we considered choice sets where the marginal option was at least 2 points lower than the next-lowest option in the same set.

Results were consistent: adding a less-attractive option decreased choice share of that alternative. Adding an option to a 1v1 choice to get a 1v2 choice (as in Experiments 1 and 1b) decreased share of the option including the marginal option (1c: $M = 0.05$, 95% CI = [0.03, 0.07],

$t(655) = 4.95$, $p < .001$; 1d: $M = 0.08$, 95% CI = [0.04, 0.12], $t(283) = 3.83$, $p < .001$). The same effect held for each relevant comparison (1v1 vs. 1v3, 1v2 vs. 2v2, and 2v2 vs. 2v3: all *M*s > 0.03; all *p*s < .05; see Figure 2b and supplements).

A key question is the exact method by which adding M_L to M_H to form the multi-option alternative $\{M_H, M_L\}$ decreases choice probability. One mechanism would be that adding M_L reduces the perceived value of M_H, thereby reducing the derived value of M. A second distinct mechanism would be that adding M_L reduces the perceived holistic value of M. In an additional follow-up study (1e, in the supplements) we asked participants to rate the values of: (1) M_H following a binary choice, (2) M_H following a test choice, and (3) the participant's choice of a movie from the set $\{M_H, M_L\}$ following a test choice. We find a significant difference between M_H following binary and $\{M_H, M_L\}$ following test ($t(174) = 4.01$, $p < .001$). In exploratory analyses controlling for initial ratings of M_H, we find this large difference is significantly larger than the small difference between M_H following binary and M_H following test ($t(174) = 2.02$, $p = .045$). This test provides evidence that adding a lower-value option decreases the value of a set. Moreover, this effect is greater than any effect on the value of M_H alone.

Experiment 2

In Experiments 2 and 3, we switch to the domain of risky choice, using card draws, die rolls, and coin flips (Figure 1). This domain enables greater control, incentive-compatibility, and testing for cases of objective dominance, while maintaining a similar experimental design. This new domain also reduces the likelihood that participants drew inferences about theater quality or the context of the consumption experience based on the set of movies shown, and further reduces the likelihood that participants undervalue a multi-option alternative due to uncertainty about future preferences.

To test for undervaluation, for every participant, we again 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 2a). Choice of M_H was 5 percentage points higher than choice of M (Table 1).

We find further suggestive evidence that preference strength moderates this undervaluation effect: the coefficient on the difference in expected values ($EV(M_H) – EV(M_L)$) was in the expected direction, though it was not statistically significant, $b_3 = 0.011$, $SE = 0.008$, $p = .16$. However, participants' choices indicated they were sensitive to probabilities directly, not merely probabilities as they impacted expected values. When we use the difference in probabilities instead of the difference in expected values, we find a marginally significant relationship between M_H-M_L dominance and Choice Difference ($b_3 = 0.11$, $SE = 0.06$, $p = .07$, Figure 2d). This speaks in favor of the preference strength explanation (vs. measurement error), as probabilities here are known rather than measured (with noise).

We also compared choices in the test choices (S vs. $\{M_H,M_L\}$) to choices in trinary (S vs. M_H vs. M_L) choices. Using a one-sample t-test of (Choose $\{M_H,M_L\}$ in test choices) – (Choose M_H or M_L in trinary choices), 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, $M = –0.08$, 95% CI = [–0.10, –0.06], $t(240) = –7.95$, $p < .001$ (see Figure S1).

We ran two additional, similar experiments (2b: N = 298, 194 after exclusions; 2c: N = 298, 176 after exclusions). 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], $t(175) = 3.37$, $p < .001$). Additional details are available in the supplements.

Experiment 3

In Experiment 3, we introduce a process-tracing measure to our design: mouse-tracking (Figure 1). We measure the information acquisition process (i.e., whether S, M_H, and/or M_L were examined and for how long) 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- and low-transparency pairings. However, 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. An example of a high-transparency pairing is M_H = "$2 if you roll a 1 or 2 or 3" and M_L = "$2 if you roll a 1." An example of a low-transparency pairing is 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 lead to less undervaluation as the implications for the second stage choice of M_H vs. M_L are clearer for more-transparent pairings.

Choice 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 2a). We tested for moderation by transparency by calculating the within-subject differences in undervaluation on high-transparency trials and undervaluation on low-transparency trials. Using a one-sample t-test we find significant evidence of the moderation ($M = –0.03$, 95% CI = [–0.06, –0.01], $t(208) = –2.76$, $p = .006$). There is less undervaluation when the M_H-M_L dominance is more transparent ($M_{\text{HighTransparency}} = 0.04$, 95% CI = [0.02, 0.06], t(208)

= 3.70, $p < .001$; $M_{\text{LowTransparency}} = 0.07$, 95% CI = [0.05, 0.10], $t(208) = 5.95$, $p < .001$).[2] This moderation indicated attenuation but not elimination. We still find evidence of undervaluation even when the dominant relationship is transparent.

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 binary 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 find a positive correlation between-subject between this average proportion mouse difference and degree of undervaluation, $r = 0.27$, $t(207) = 4.04$, $p < .001$. Participants who spent relatively more time inspecting S when paired with M than when paired with M_H undervalued M more (Figure 3a).[3]

We regressed choice for S (in test 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 $\{M_H, M_L\}$). 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 over M_L, the less likely they were to choose S (i.e., the more likely they were to choose $\{M_H, M_L\}$) (Figure 3b).

Exploratory Results. We compared the proportion of time spent on M_H (relative to M_L) in the high-transparency trials to the proportion of time spent on M_H (relative to M_L) in the low-

[2] Experiments 2b and 2c reported in the supplements also included a transparency manipulation without process tracing. We did not observe moderation by transparency in those two experiments. As described in the General Discussion, we speculate that the focusing enforced by the process tracing increased the moderating effect of transparency on choice.

[3] We also tested for trial-level associations between mouse movements and choice. Across all binary, test, and trinary choices, we used logistic regression to regress choice (of S) on the proportion of mouse-hover time spent on S, with random intercepts and slopes at the subject level. We find a significant relationship between information acquisition and choice within-subject, $b = 4.79$, $SE = 0.18$, $p < .001$.

transparency trials. We found a small but significant difference, such that subjects spent relatively more time on M_H when the dominance was more transparent, $M = 0.01$, 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 = [–0.26, 0.008], $t(206) = -1.86$, $p = .06$. In other words, as subjects spent more relative time on M_H in high-transparency trials (vs. low-transparency trials), they showed a smaller degree of undervaluation in high-transparency trials (vs. low-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. 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).

Although our main effect of interest (undervaluation) is statistically significant, and the presence of undervaluation is widespread (61% of participants exhibited undervaluation and only 29% exhibited either overvaluation or sensitivity to option value; 10% exhibited no difference), there is variability in effect size across participants ($M = 0.06$, $SD = 0.15$). Some of this variability can be attributed to individual differences: the degrees of undervaluation in low-transparency trials and high-transparency trials are correlated, $r = 0.50$, 95% CI = [0.39, 0.59], $t(207) = 8.30$, $p < .001$. In other words, a high degree undervaluation on one half of the trials corresponded to a high degree of undervaluation on the other half. This 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.

Discussion

Across nine controlled lab 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 $\{M_H,M_L\}$ than they are to choose a single-option alternative M_H. This effect also holds when the multi-option alternative has more than two options. 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, there is less undervaluation. In replication experiments (see supplements), we provide evidence against multiple alternative explanations (noisy responding; strong delayed choice aversion). We also find process evidence which suggests that undervaluation is strongly associated with information acquisition patterns and is driven by devaluating the holistic option M and not just M_H.

These findings connect to work on agenda effects,[43-45] in which the order of decisions influences the option that is ultimately chosen. However, this literature does not offer explanations for the present results. Instead, this undervaluation phenomenon seems to be an additional instance in which agenda effects matter. Moreover, in contrast to work on agenda effects, we find that undervaluation persists even in the simplest case: a single option pitted against a two-option alternative. Our current findings also have certain similarities with the uncertainty effect (where people value a prospect as less valuable than its worst outcome).[46,47] However, two key distinctions are that (a) the uncertainty effect addresses a risky outcome whereas here, the decision maker chooses an option from the multi-option alternative, and (b) the uncertainty effect is assessed relative to the worse outcome whereas undervaluation is assessed relative to the better outcome.

There are several literatures that address similar phenomena, including assortment choice and multi-alternative (i.e., more than 2 alternatives) choice.[16,17] However, these literatures ultimately do not address the question at hand. Assortment choice research focuses on the evaluation of assortments but does not typically enable comparisons to a value-maximizing benchmark, nor does it compare evaluations of assortments to the evaluations of the constituent parts of the assortment in a choice context. Regarding multi-alternative choice research, our test choices (i.e., choices between a single-option alternative S and a multi-option alternative $\{M_H,M_L\}$) are formally equivalent to trinary choices (i.e., choices between S, M_H, and M_L). However, since participants are significantly less likely to choose $\{M_H,M_L\}$ (in a test choice) than they are to choose M_H or M_L (in a trinary choice), these two types of choices are neither psychologically nor practically equivalent.

We have addressed several potential alternative explanations with our data. First, we find evidence in experiment 1e that the value of M in test choices is significantly lower than that of M_H in binary choices. Second, 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 rule out delayed-choice aversion in two ways: (1) participants understand that they will make a fixed number of decisions, and (2) we find that participants choose $\{M_H,M_L\}$ more often than they choose M_L when pitted against S, so they do not have an absolute aversion to delayed choices. In supplemental analyses, using a subset of choices for which M_H is chosen less than 50% of the time, we see that adding a less-attractive option further decreases choice share. This indicates undervaluation is not merely a regression to 50%.

Several unanswered questions about this phenomenon remain. Although our experiments find process evidence for undervaluation, these results are correlational. 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[20,23,26,48], but we do not present causal evidence for that here).

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 (see supplements). 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 in Experiment 3, but not in the others. The more-directed information acquisition method may have better enabled participants to edit out the transparently dominated option.

A promising avenue for future research would include an investigation of individual differences that relate to undervaluation. We began to explore this in supplemental experiments, but our findings did not yield diagnostic results. We hypothesized that individuals who elaborated more extensively on potential outcomes[49] or who use more analytic rather than holistic approaches to thinking through decisions[50] would exhibit less 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.

In addition to individual differences that correlate with undervaluation, a key factor in assessing the presence of undervaluation is the heterogeneity of preferences. With perfect information about each individual's preferences, offering only each person's most preferred option would lead to the highest choice share, regardless of the underlying heterogeneity in preferences. However, choice architects with the power to shape choice sets regularly have incomplete information. In a heterogeneity simulation using data from Experiments 1 and 1b (see supplements), we demonstrate that the range of undervaluation effects that could be observed in the real world will be dependent on both the underlying heterogeneity of preferences and the ability to identify which people have which preferences. Future research could expand on the impact of heterogeneity on observed undervaluation.

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-option alternative. This finding has important implications. Offering a multi-option alternative may appear to be an attractive approach to increase choice for any option(s) as it enables appeal to participants with heterogeneous preferences. But if the options within the alternative are discrepant in value, our results suggest that choice likelihood for that alternative may be lower than it would be otherwise for any given person. It is not clear from the present research whether knowledge of this tendency would reduce or reverse the undervaluation effect, but it is important for choice architects to be aware of this way in which decision makers' choices defy value-maximizing expectations.

Methods

Experiment 1

Participants. For this preregistered experiment (https://aspredicted.org/blind.php?x=j3ex33), we collected responses from 305 Amazon Mechanical Turk (AMT) 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 film. For each film, participants had the option to click a box labeled "I've never heard of this movie." 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 1). 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 were randomly generated for each participant. They fell into 3 categories (with 10 choices in each category) but were presented in randomized order. 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 movies that the participant rated. 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 (i.e., the films for which they selected "I've never heard of this movie"). In this study, only one participant rated fewer than 30 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 was matched to one of the test choices: each control binary choice was a choice between the S and M_H from the matched test choice.

Finally, the last category of 10 trials were *filler* choices, each of which comprised two single-movie theaters. The filler trial films were randomly selected from the rated films, and independently from the selection process for the test trials.

Exclusions and Data Preprocessing. As specified in our preregistration, we excluded anyone who failed to rate at least 20 films. We also excluded anyone whose filler choices were not directionally predicted by their ratings in a logistic regression of *ChooseLeft* on (*RatingLeft-RatingRight*). These criteria resulted in the exclusion of 37 participants, leaving us with a final sample size of 268. As specified in our preregistration, we excluded any test-control choice pairs that were generated from unrated films.

Analysis. Because M_H and M_L were randomly selected, we can examine the role of preference strength for M_H over M_L in undervaluation. We used matched choice sets as the unit of observation with the variable Choice Difference as our outcome measure, defined as (Chose S in test – Chose S in matched control), 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).

$$Choice\ Difference_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 (M_H + M_L)_{it} + \beta_3 (M_H - M_L)_{it} + \epsilon_{it} \quad (1)$$

Experiment 2

Participants. For this preregistered experiment (https://aspredicted.org/blind.php?x=ib4vi9), we collected responses from 304 AMT workers. They earned $1.75 (the first 20 participants) or $1.90 (the remaining 284 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 depicted in Figure 1. 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 numbers on a six-sided die. Binary control 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 as shown in Figure 1. 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 $\{M_H, M_L\}$) and binary control choices (e.g., choosing between S and M_H), participants also made trinary choices (e.g., choosing among gambles S, M_H, and M_L). In total, participants made 34 incentivized choices (including three attention check questions and one M_H vs. 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 (Juechems et al. 2017). In this experiment, three of the five participants won money ($6, $7, and $10). See the supplements for additional information on the trial generation process.

As in Experiment 1, participants completed several comprehension check questions, including a question about the meaning of the test choices, i.e., that choosing $\{M_H, M_L\}$ implied that they would get to choose one option 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 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.

Analysis. To examine the role of M_H-M_L dominance, 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).

$$Choice\ Difference_{it} = \beta_0 + \beta_1 EV_{S_{it}} + \beta_2\left(EV_{M_H} + EV_{M_L}\right)_{it} + \beta_3\left(EV_{M_H} - EV_{M_L}\right)_{it} + \epsilon_{it} \qquad (2)$$

We also adapted our preregistered analysis based on exploratory analyses (detailed in the supplements). Instead of regressing Choice Difference on expected values, we regressed it on probabilities.[4] 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. 3).

$$Choice\ Difference_{it} = \beta_0 + \beta_1 P_{S_{it}} + \beta_2 \left(P_{M_H} + P_{M_L}\right)_{it} + \beta_3 \left(P_{M_H} - P_{M_L}\right)_{it} \quad (3)$$

Experiment 3

Participants. For this preregistered experiment (https://aspredicted.org/blind.php?x=8tg3fe) we collected responses from 302 AMT 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. First, the trials were not randomly generated at the subject-level and instead came from a predetermined set. Second, the M_H vs. M_L dominance relationship was manipulated to be either high- or low-transparency. Third, the information about a gamble (i.e., the details of S, M_H, and/or M_L) was not visible unless the participant hovered their cursor over the gamble. Fourth, M_H and M_L were presented horizontally instead of vertically. Fifth, we tracked participants' mouse movements. Specifically, while participants made their choices, we recorded the order in which participants viewed each piece of information (e.g., S, M_H, and M_L) and the durations of

[4] We are unable to conduct this analysis using monetary amounts because M_{M_H} was always the same as M_{M_L}.

these information acquisitions in a MouseLab-like paradigm (Johnson et al. 1989) (Figure 1). Participants made a total of 65 incentivized choices. These choices included: (1) high-transparency test choices: S vs. $\{M_H,M_L\}$ (HT), (2) low-transparency test choices: S vs. $\{M_H,M_L\}$ (LT), (3) high-transparency control choices: S vs. M_H (HT), (4) low-transparency control choices: S vs. M_H (LT), (5) trinary choices: S vs. M_H vs. M_L, and (6) binary choices between M_H and M_L. At the end of the survey, participants rank-ordered 10 $\{S, M_H, M_L\}$ triplets. 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 $\{M_H,M_L\}$ 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.

Acknowledgments

The authors acknowledge research support from the Morrison Family Center for Marketing and Data Analytics at the UCLA Anderson School of Management.

References

1. Brehm, J. W. (1966). *A theory of psychological reactance.* (135). Academic Press.
2. Bown, N. J., Read, D., and Summers, B. (2003). The lure of choice. *Journal of Behavioral Decision Making*, *16*(4), 297–308.
3. Mochon, D. (2013). Single-Option Aversion. *Journal of Consumer Research*, *40*(3), 555–566.
4. Shin, J., and Ariely, D. (2004). Keeping Doors Open: The Effect of Unavailability on Incentives to Keep Options Viable. *Management Science*, *50*(5), 575–586.
5. Brough, A. R., and Chernev, A. (2012). When Opposites Detract: Categorical Reasoning and Subtractive Valuations of Product Combinations. *Journal of Consumer Research*, *39*(2), 399–414.
6. Chernev, A., and Gal, D. (2010). Categorization Effects in Value Judgments: Averaging Bias in Evaluating Combinations of Vices and Virtues. *Journal of Marketing Research*, *47*(4), 738–747.
7. Gaeth, G. J., Levin, I. P., Chakraborty, G., and Levin, A. M. (1991). Consumer evaluation of multi-product bundles: An information integration analysis. *Marketing Letters*, *2*(1), 47–57.
8. Yadav, M. S. (1994). How Buyers Evaluate Product Bundles: A Model of Anchoring and Adjustment. *Journal of Consumer Research*, *21*, 342–353.
9. Shenhav, A., and Karmarkar, U. R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice. *Scientific Reports*, *9*(1), 1958.
10. Weaver, K., Garcia, S. M., and Schwarz, N. (2012). The Presenter's Paradox. *Journal of Consumer Research*, *39*(3), 445–460.
11. Troutman, C. M., and Shanteau, J. (1976). Do Consumers Evaluate Products by Adding or Averaging Attribute Information? *Journal of Consumer Research*, *3*(2), 101.
12. Simonson, I., Carmon, Z., and O'Curry, S. (1994). Experimental Evidence on the Negative

Effect of Product Features and Sales Promotions on Brand Choice. *Marketing Science*, 13(1), 23–40.

13. Yamanashi Leib, A., Chang, K., Xia, Y., Peng, A., & Whitney, D. (2020). Fleeting Impressions of Economic Value via Summary Statistical Representations. *Journal of Experimental Psychology: General*. Advance online publication.
14. Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, *49*(49-81), 74.
15. Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697.
16. Kahn, B. E., and Lehmann, D. R. (1991). Modeling Choice Among Assortments. *Journal of Retailing*, *67*(3), 274–299.
17. Sood, S., Rottenstreich, Y., and Brenner, L. (2004). On Decisions That Lead to Decisions: Direct and Derived Evaluations of Preference. *Journal of Consumer Research*, *31*(1), 17–25.
18. Le Lec, F., and Tarroux, B. (2019). On Attitudes to Choice: Some Experimental Evidence on Choice Aversion. *Journal of the European Economic Association*.
19. Spiller, S. A., and Ariely, D. (2020). How does the perceived value of a medium of exchange depend on its set of possible uses? *Organizational Behavior and Human Decision Processes*, *161*, 188–200.
20. Armel, K. C., Beaumel, A., and Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, *3*(5), 396–403.
21. Fiedler, S., and Glöckner, A. (2012). The Dynamics of Decision Making in Risky Choice: An Eye-Tracking Analysis. *Frontiers in Psychology*, *3*.
22. Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and

comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298.

23. Mormann, M. M., Navalpakkam, V., Koch, C., and Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, *22*(1).
24. Newell, B. R., and Le Pelley, M. E. (2018). Perceptual but not complex moral judgments can be biased by exploiting the dynamics of eye-gaze. *Journal of Experimental Psychology: General*, *147*(3), 409–417.
25. Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., and Hertwig, R. (2018). Prospect Theory Reflects Selective Allocation of Attention. *Journal of Experimental Psychology: General*, *147*(2), 147–169.
26. Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., and Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences*, *112*(13), 4170–4175.
27. Vaidya, A. R., and Fellows, L. K. (2015). Testing necessary regional frontal contributions to value assessment and fixation-based updating. *Nature Communications*, *6*, 10120.
28. Busemeyer, J., and Townsend, J. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459.
29. Kim, B. E., Seligman, D., and Kable, J. W. (2012). Preference reversals in decision making under risk are accompanied by changes in attention to different attributes. *Frontiers in Neuroscience*, *6*.
30. Fisher, G. (2017). An attentional drift diffusion model over binary-attribute choice. *Cognition*, *168*, 34.
31. Smith, S. M., and Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General*, *147*(12), 1810–1826.

32. Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857.

33. Gluth, S., Kern, N., Kortmann, M., and Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*.

34. Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 534–552.

35. Konovalov, A., and Krajbich, I. (2020). Mouse tracking reveals structure knowledge in the absence of model-based choice. *Nature Communications*, *11*(1).

36. Stillman, P. E., Krajbich, I., and Ferguson, M. J. (2020). Using dynamic monitoring of choices to predict and understand risk preferences. *Proceedings of the National Academy of Sciences*, *117*(50), 31738.

37. Reeck, C., Wall, D., and Johnson, E. J. (2017). Search predicts and changes patience in intertemporal choice. *Proceedings of the National Academy of Sciences*, *114*(45), 11890–11895.

38. Diehl, K. (2005). When Two Rights Make a Wrong: Searching Too Much in Ordered Environments. *Journal of Marketing Research*, *42*(3), 313–322.

39. Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, *69*(4), 623-638.

40. Lichtenstein, S., and Slovic, P. (1971). Reversals of Preference Between Bids and Choices in Gambling Decisions. *Journal of Experimental Psychology*, *89*(1), 46–55.

41. Amir, O., and Ariely, D. (2007). Decisions by rules: The case of unwillingness to pay for beneficial delays. *Journal of Marketing Research*, *44*, 142–152.

42. O’Donnell, M., and Evers, E. R. K. (2019). Preference Reversals in Willingness to Pay and Choice. *Journal of Consumer Research*, *45*(6), 1315–1330.

43. Plott, C. R., and Levine, M. E. (1978). A Model of Agenda Influence on Committee Decisions. *The American Economic Review*, *68*(1), 146–160.

44. Tversky, A., and Sattath, S. (1979). Preference trees. *Psychological Review*, *86*(6), 542–573.

45. Hauser, J. R. (1986). Agendas and Consumer Choice. *Journal of Marketing Research*, *23*(3), 199–212.

46. Gneezy, U., List, J. A., & Wu, G. (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics*, *121*(4), 1283-1309.

47. Simonsohn, U. (2009). Direct risk aversion: Evidence from risky prospects valued below their worst outcome. *Psychological Science*, *20*(6), 686-692.

48. Gwinn, R., Leber, A. B., and Krajbich, I. (2019). The spillover effects of attentional learning on value-based choice. *Cognition*, *182*, 294–306.

49. Nenkov, G. Y., Inman, J. J., Hulland, J., & Morrin, M. (2009). The impact of outcome elaboration on susceptibility to contextual and presentation biases. *Journal of Marketing Research*, *46*(6), 764-776.

50. Choi, I., Koo, M., and Choi, J. A. (2007). Individual Differences in Analytic Versus Holistic Thinking. *Personality and Social Psychology Bulletin*, *33*(5), 691–705.

Figures and Tables

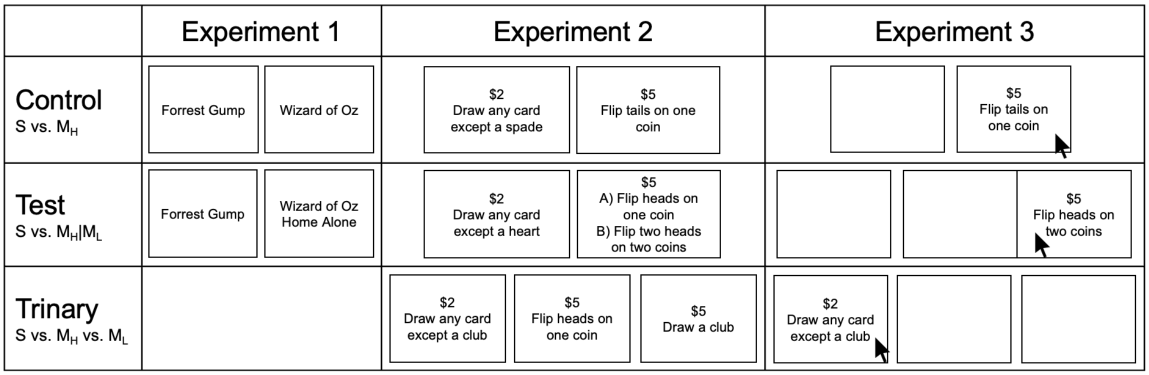

Figure 1. Experiment design. In experiment 1, participants first rated their desire to watch each of 50 films. Participants then made 30 hypothetical choices between two theaters, each of which was showing one or two films. Participants were told to imagine that if they chose a two-film theater, they would get to choose one of the two films to watch. In experiments 2 and 3, participants chose between 2 (or 3) options.

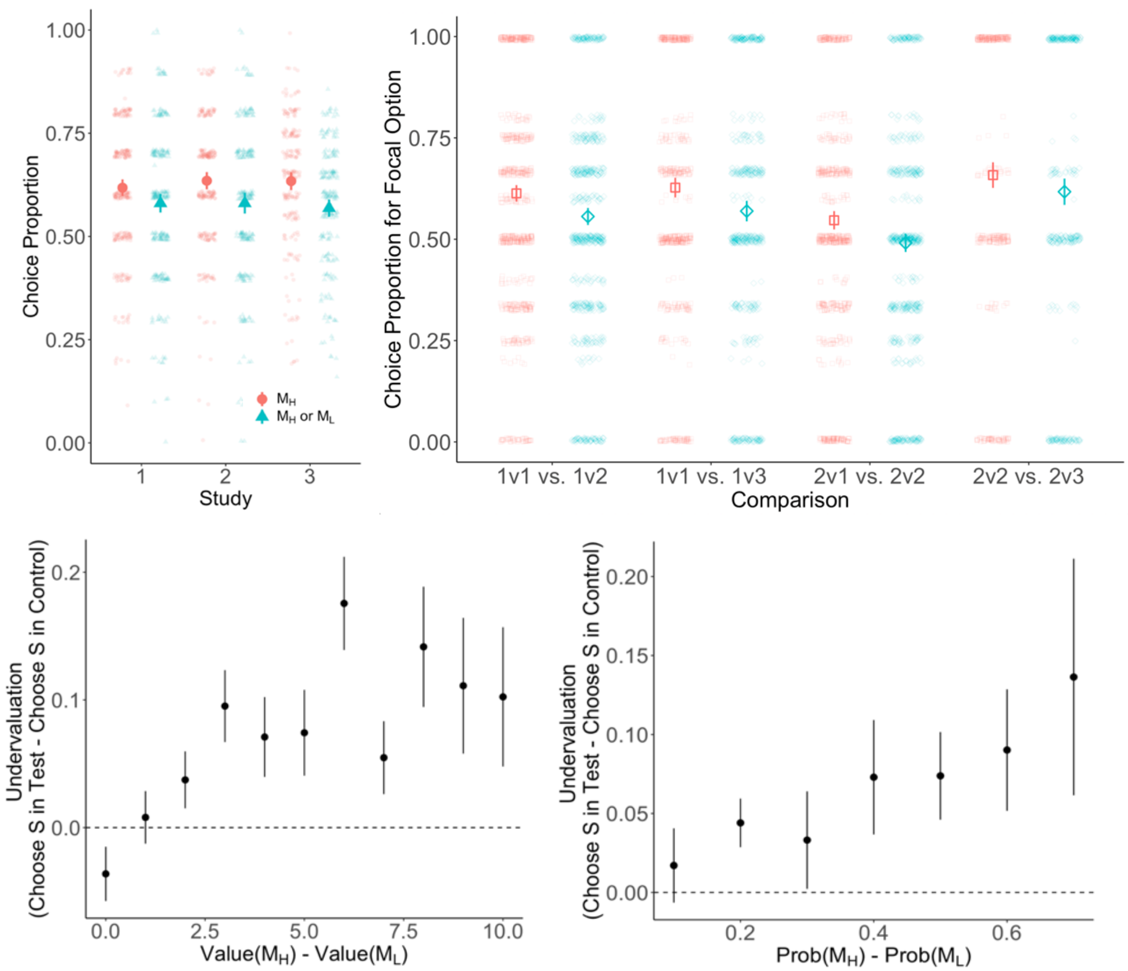

Figure 2. Behavioral results. a) Main effect of undervaluation across experiments. Participants are less likely to choose $\{M_H, M_L\}$ than they are to choose M_H when compared to the same single-option alternative S. b) Main effect of undervaluation in studies 1c and 1d, using multi-option alternatives with different numbers of options. Participants are less likely to choose the alternative with an additional (inferior) sub-option. Results are collapsed across 1c and 1d. c) 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. d) Relationship between (M_H>M_L) dominance strength (defined using probabilities) and degree of undervaluation. Undervaluation increases as M_H gets progressively better than M_L, though this relationship is only marginally significant (p = .07).

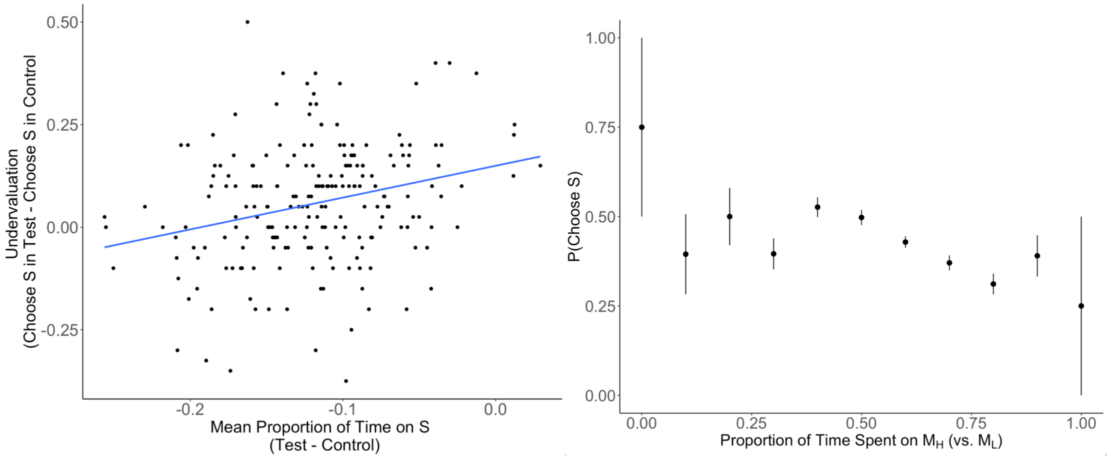

Figure 3. Mouse-tracking results. a) Relationship between subject-level mouse-tracking and undervaluation. As participants spend relatively more time on S in test choices (vs. control), they display more undervaluation, r = 0.27, t(207) = 4.04, p < .001. b) Relationship between mouse-tracking and choices in test choices. As participants spent more time on M_H (relative to M_L), they were less likely to choose the single-option alternative (S), *b* = –1.70, *SE* = 0.24, *p* < .001. 60% of all proportion of time spent on M_H lie in the interval [0.5, 0.75]. Within that interval, the coefficient is even more sharply negative (-3.05).

Table 1. Main effect of undervaluation across all experiments.

Experiment	M_H	M_{HL}	Main Effect	95% CI	t statistic[b]	p value
1	0.618	0.580	0.04	[0.02, 0.06]	3.73	< .001
1b (supplements)	0.632	0.590	0.04	[0.03, 0.05]	6.75	< .001
1c (supplements)[a]	0.587	0.544	0.05	[0.03, 0.07]	4.95	< .001
1d (supplements)[a]	0.590	0.521	0.08	[0.04, 0.12]	3.83	< .001
1e (supplements)	0.630	0.605	0.03	[0.01, 0.04]	2.94	.004
2	0.635	0.580	0.05	[0.03, 0.08]	5.04	< .001
2b (supplements)	0.572	0.508	0.06	[0.03, 0.09]	4.28	< .001
2c (supplements)	0.576	0.510	0.06	[0.02, 0.09]	3.37	< .001
3	0.634	0.569	0.06	[0.04, 0.08]	5.64	< .001

[a] Based on preregistered analysis of 1v1 and 1v2 choices where $Rating_{M_H} - Rating_{M_L} \geq 2$.
[b] Each of these tests are consistent with the results of a corresponding Wilcoxon signed-rank test.