Web Appendix:

Consumers Undervalue Multi-Option Alternatives

Web Appendix A: Main Effect, Conditioned on Relative Ratings of S, M_H, and M_L

Web Appendix B: Experiment 1b

Web Appendix C: Experiment 1c

Web Appendix D: Experiment 2b

Web Appendix E: Experiment 3b

Web Appendix F: Extra detail from Experiment 4

Web Appendix G: Experiment 4b

Web Appendix H: Experiment 4c

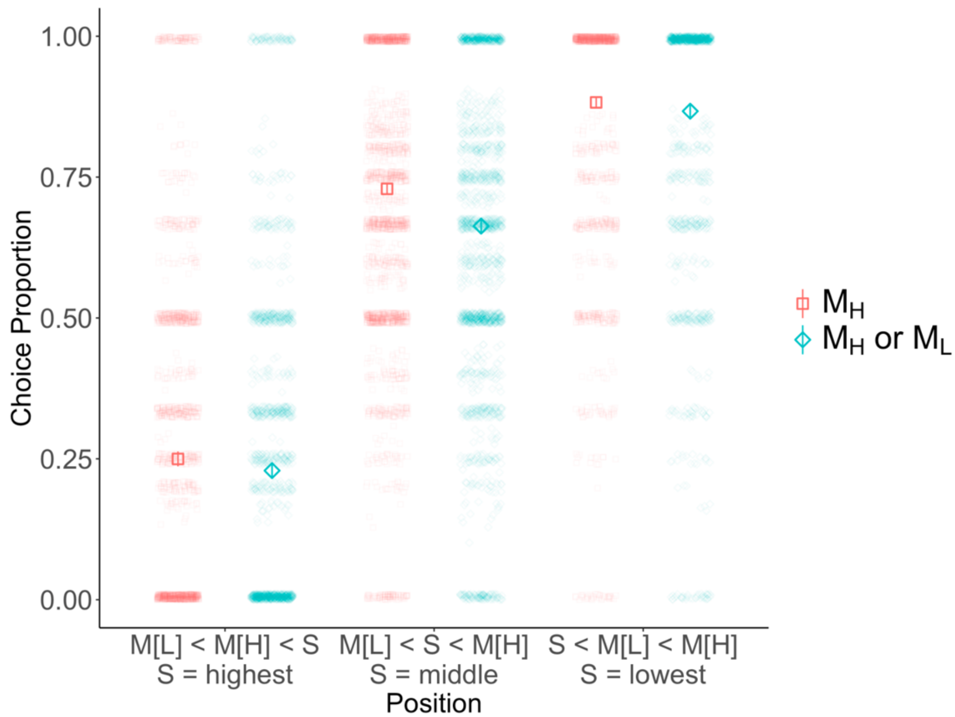
Web Appendix I: Extra detail from Experiment 5

Web Appendix J: Extra detail on Heterogeneity Simulations

Web Appendix K: Supplementary Tables

Appendix A: Main Effect, Conditioned on Relative Ratings of S, M_H, and M_L

Figure S1: Undervaluation by Relative Ratings



Notes: Main effect of undervaluation, separated by the relative rating of S (compared to M_H and M_L). In all situations, 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 (S = highest: $t(2401) = 3.66$, $p < .001$; S = middle: $t(2841) = 12.62$, $p < .001$; S = lowest: $t(2113) = 3.19$, $p = .001$). Data is collapsed across the movie studies (Studies 1, 1b, 1c, 2, 2b, 3, and 3b), which have individual ratings per option. Bars indicate s.e.m. across participants.

Appendix B: Experiment 1b

Method

Participants

For this preregistered experiment (https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT), we collected responses from 604 Amazon Mechanical Turk workers; each earned $1.25 for participating.

Materials and Procedure

The materials and procedure were identical to Experiment 1.

Exclusions and Data Preprocessing

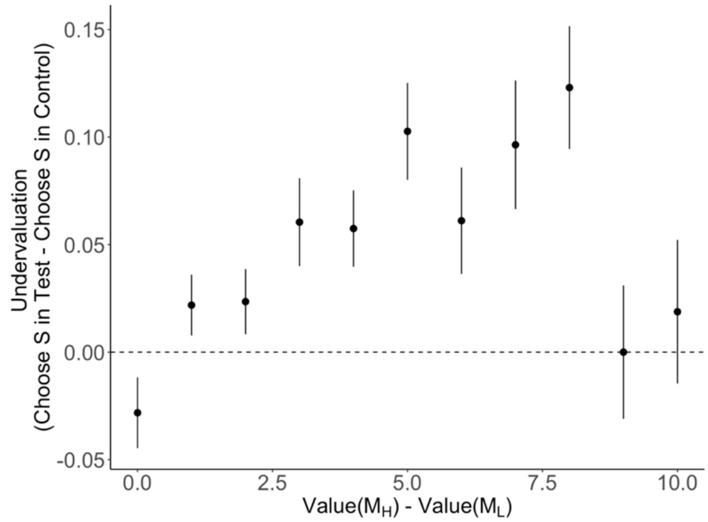
We followed the same exclusion rule used in Experiment 1. These criteria resulted in the exclusion of 80 participants, leaving us with a sample size of 524.

Results (Preregistered)

Using the same approach as in Experiment 1, we found evidence for undervaluation, M = 0.04, 95% CI = [0.03, 0.05], $t(523) = 6.75$, $p < .001$. Using that approach, we find that 41% of participants exhibited undervaluation, with 37% exhibiting no difference and 22% exhibiting an effect in the opposite direction. As in Experiment 1, we find significant evidence for an effect of M_H-M_L preference strength on undervaluation, $b = 0.011$, $SE = 0.002$, $p < .001$ (Fig. S2).[1]

[1] Due to a since-addressed coding error in an early analysis of Experiment 1, we collected this dataset with additional preregistered exclusions. We specified that we would remove all choices where the M_H-M_L difference was less than 5. With that exclusion, we find a larger effect, $M = 0.09$, 95% CI = [0.06, 0.12], $t(335) = 6.64$, $p < .001$.

Figure S2: Preference Strength and Undervaluation



Notes: 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.

Appendix C: Experiment 1c

Method

Participants

For this preregistered experiment (https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT), we collected responses from 499 Amazon Mechanical Turk workers; each earned $1.25 for participating.

Materials and Procedure

The materials and procedure were identical to Experiment 1, with the following change: instead of a multiple-choice comprehension question about multi-option alternatives, participants were asked to consider a choice between two theaters (one showing a single movie, Beauty and the Beast, and one showing two movies, Fantasia and Jurassic World). Participants then encountered the following prompt: "In the example decision above, suppose you picked theater 2 (Fantasia; Jurassic World). Please briefly describe how which movie you watch would be determined."

As in Experiment 1, participants also responded to comprehension questions to ensure that they understood (1) that the choices were hypothetical and (2) that their choices would not influence the number of choices that they would have to make.

Exclusions and Data Preprocessing

We used the same exclusion criteria as in Experiment 1: participants had to rate at least 20 films and their filler choices had to be directionally predicted by their ratings. Furthermore, as preregistered, we only included participants in analysis if two coders (blind to hypothesis and choice data) agreed that the participants' answer to the open-ended comprehension question unambiguously indicated that the participant would watch their choice of movie from the pair. In total, we excluded 187 participants from the original sample, resulting in a final sample size of N = 312.

Results (Preregistered)

Using the same approach as in Experiment 1, we found evidence for undervaluation, M = 0.02, 95% CI = [0.01, 0.04], $t(311) = 3.18$, $p = .002$. Using that approach, we find that 35% of participants exhibited undervaluation, with 43% exhibiting no difference and 22% exhibiting an effect in the opposite direction. As in Experiment 1, we find significant evidence for an effect of M_H-M_L preference strength on undervaluation, $b = 0.017$, $SE = 0.002$, $p < .001$.

Appendix D: Experiment 2b

Method

Participants

For this preregistered experiment (https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT), we collected responses from 201 Amazon Mechanical Turk (AMT) workers. They earned $1.50 for their participation.

Materials and Procedure

The procedure for this experiment was identical to Experiment 2.

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 24 participants, leaving us with a final sample size of 177. As specified in our preregistration, we excluded any test-control choice pairs that were generated from unrated films.

Results

As in previous experiments, we find evidence of undervaluation, $M = 0.03$, 95% CI = [0.01, 0.04], $t(176) = 2.94$, $p = .003$. Next, we make several comparisons between ratings. First, we look at whether the rating of S changes, depending on whether it is part of a test choice or a control choice, and find that it was rated marginally significantly lower following binary than test ($M = -0.45$, 95% CI = [-0.93, 0.04], $t(175) = 0.069$). Next, we make within-subject comparisons among three values: (i) the rating of M_H following control choices, (ii) the rating of M_H following test choices and (iii) the rating of $\{M_H, M_L\}$ following test choices, using one-sample t-tests.

To reduce sampling variability in the selected movies we controlled for differences in how those movies were rated by those participants in their initial ratings. We again find that $\{M_H,M_L\}$ is rated lower following test than is M_H following binary (M = -0.67, 95% CI = [-1.00, -0.34], $t(174) = 4.01$, $p < .001$). But in this more powerful analysis, M_H is rated marginally significantly lower following test than following binary (M = -0.28, 95% CI = [-0.57, .005], $t(174) = 1.94$, $p = .054$) and $\{M_H,M_L\}$ is rated significantly lower following test than M_H is following test (M = -0.37, 95% CI = [-0.73, -0.01], $t(174) = 2.02$, $p = .044$).

In our preregistration for this study (which was collected before Experiment 2), we did not specify whether we would control for the initial ratings of the films. When we do not control for them, we find that the rating of $\{M_H,M_L\}$ following test choices is significantly lower than the rating of M_H following binary choices (M = -0.60, 95% CI = [-1.06, -0.14], $t(175) = 2.59$, $p = .010$) and that the rating of M_H following test choices is not significantly lower than the rating of M_H following binary choices (M = -0.29, 95% CI = [-0.73, 0.14], $t(175) = -1.33$, $p = .186$). The rating of $\{M_H,M_L\}$ following test choices is not significantly lower than the rating of M_H following test choices (M = -0.31, 95% CI = [-0.74, 0.13], $t(175) = -1.39$, $p = 0.166$).

In sum, it is clear that $\{M_H,M_L\}$ is rated lower following test than M_H is following binary. Given the most powerful test (accounting for sampling variability of differences in movies), $\{M_H,M_L\}$ is rated lower following test than M_H is following test, suggesting that part of the undervaluation effect is indeed due to undervaluation of the holistic option, not just the more-preferred option in the presence of the other option.

Appendix E: Experiment 3b

Method

Participants

For this preregistered experiment (https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT), we collected responses from 517 AMT workers. They earned $1.50 for their participation.

Materials and Procedure

The procedure was identical to Experiment 3.

Exclusions and Data Preprocessing

As specified in our preregistration, we excluded anyone who failed to rate at least 18 films, as they would have fewer usable primary choices and thus noisier estimates. We also excluded anyone whose simple binary 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 158 participants, leaving us with a final sample size of 359.

Also in line with our preregistration, we excluded any choice sets that were generated from unrated films.

Results

In this study, we can test for undervaluation of several theaters. First, analogous to analyses in Experiments 1, 1b, 2, and 2b, we can compare choices in 1v1 and 1v2 trials to test for undervaluation of the 2-movie theater in 1v2 choices. To test for undervaluation in this case, for every participant, we calculated $S_A > \{M_A, M_B\}$ as (Chose S in 1v2 choices / number of valid 1v2 choices). We also calculated $S_A > M_A$ as (Chose S in 1v1 choices / number of valid 1v1 choices). We tested (S_A > {M_A, M_B}) – (S_A > M_A) using a one-sample t-test and found evidence for undervaluation, $M = 0.05$, 95% CI = [0.03, 0.07], $t(655) = 4.95$, $p < .001$. In other words, choice of M_A was 5 percentage points higher than choice of {M_A, M_B}.

We used the analogous approach to compare 1v1 against 1v3, 1v2 against 2v2, and 2v2 against 2v3. In each case, we find undervaluation (1v1 vs. 1v2 trials: $M = 0.08$, 95% CI = [0.04, 0.12], $t(283) = 3.83$, $p < .001$; 1v1 vs. 1v3 trials: $M = 0.08$, 95% CI = [0.03, 0.12], $t(224) = 3.23$, $p = .001$; 1v2 vs. 2v2 trials: $M = 0.07$, 95% CI = [0.03, 0.12], $t(270) = 3.02$, $p = .003$; 2v2 vs. 2v3 trials: $M = 0.06$, 95% CI = [0.01, 0.12], $t(180) = 2.16$, $p = .03$.).

For this experiment (which was collected prior to Experiment 3), we did not preregister the requirement that the added sub-option be rated at two points lower than the existing sub-option(s), as we did in Experiment 3. Without that requirement, only the first two comparisons are significant: 1v1 vs. 1v2 trials: $M = 0.04$, 95% CI = [0.01, 0.06], $t(358) = 2.95$, $p = .003$; 1v1 vs. 1v3 trials: $M = 0.02$, 95% CI = [–0.002, 0.05], $t(358) = 1.80$, $p = .07$; 1v2 vs. 2v2 trials: $M = 0.004$, 95% CI = [–0.02, 0.03], $t(358) = 0.37$, $p = .72$; 2v2 vs. 2v3 trials: $M = –0.03$, 95% CI = [–0.06, 0.01], $t(345) = –1.62$, $p = .11$.

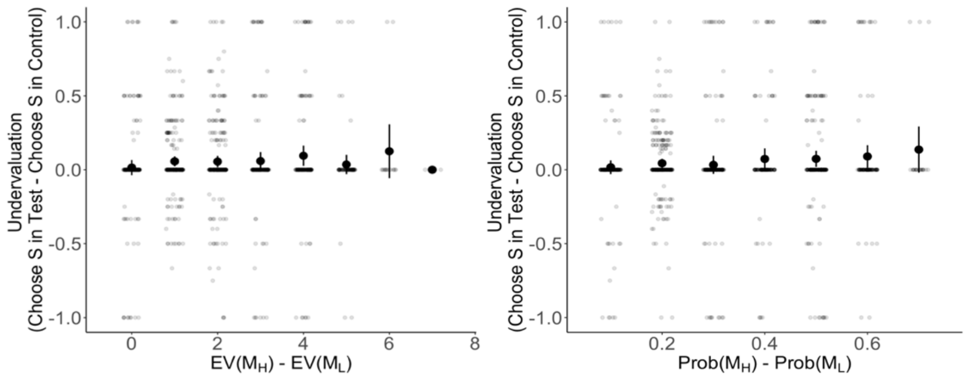
Appendix F: Experiment 4

Trial Generation Process

To generate the 34 trials, we started with the set of all possible events, separated by domain (i.e., cards, dice, and coins) and probability. In the card domain, we only used suits/colors (e.g., drawing a black card, drawing a spade, drawing a red card or a club). In the dice domain, we used every combination of numbers (e.g., rolling a 1, rolling a 1 or 2 or 4 or 6, rolling a 2 or 3 or 5). In the coin domain, we used every outcome of flipping one or two coins (e.g., flipping heads on one coin, flipping tails on two coins, flipping at least one head on two coins). We then randomly selected, for each participant and with replacement, three domains and three probabilities within those domains to fill the roles of S, M_H, and M_L. The only constraint we placed on the assignment of domains/probabilities to the three roles (S, M_H, and M_L) was to require that the probability of M_H be greater than the probability of M_L. If this was not possible (i.e., if the sampled probabilities were equal), then we resampled all three roles. Once the domains and probabilities were established, to populate binary control, test, and trinary choices, we randomly sampled with replacement three events for S, three events for M_H, and two events for M_L. For instance, if the domain for S was dice and the probability for S was 0.5, then we sampled three events (with replacement) from the entire list of possible dice events with probability 0.5 (e.g., rolling a 1 or 2 or 3; rolling a 3 or 5 or 6; rolling a 1 or 3 or 4). To determine the monetary amounts for each option, we randomly selected (with replacement) two dollar amounts ranging from \$2 to \$10; the first was assigned to the S events and the second was assigned to M_H, M_L, and M_H or M_L.

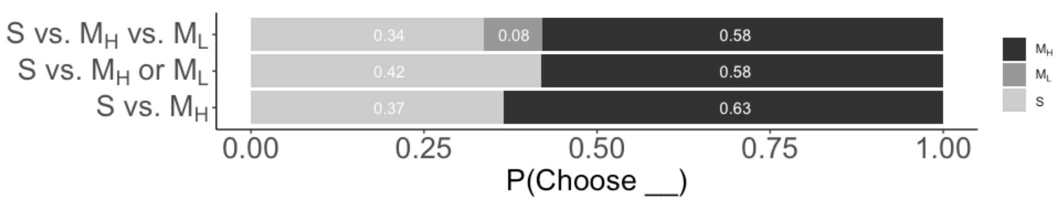
At the end of the survey, participants also completed a corresponding M_H vs. M_L trial from one of the 10 trial sets that they completed. Three attention check trials were used as detailed in Experiment 4b.

Figure S3: Dominance Strength and Undervaluation



Notes: Relationship between (M_H>M_L) dominance strength and undervaluation. Across both definitions of dominance ((a) expected value, (b) probabilities), 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 indicate s.e.m. across participants.

Figure S4: Proportion of Choices



Notes: Proportion of choices for each option in binary, test, and trinary choices for Experiment 4.

Appendix G: Experiment 4b

Method

Participants

For this preregistered experiment (https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT), we collected responses from 298 Amazon Mechanical Turk workers. They earned \$2.50 (the first 20 participants) or \$1.50 (the remaining 278 participants) for their participation. Five randomly-selected participants received the outcome of one of their decisions, as detailed below.

Materials and Procedure

This study was very similar to Experiment 4, except for the following changes. First, the choices were not randomly generated as in Experiment 4. See below for trial generation information. The other important difference is that the M_H vs. M_L dominance relationship was designed to be either high-transparency or low-transparency as in Experiment 5. Participants made a total of 39 or 40 choices. At the end of the study, we randomly selected five participants, but none of them won anything from their randomly-selected decision.

Trial Generation

There are 10 trial sets (i.e., t = {1:10}). Each trial set *t* consists of three option types: S^t, M_H^t, and M_L^t each with an associated monetary winning (*m*, ranging from \$2 to \$10) and probability of winning (*p*, ranging from 0.167 to 0.75). Within a trial set, $m(M_H) = m(M_L)$ and $p(M_H) > p(M_L)$. Within each trial set, there are three trial varieties (*v*) that share *p* and *m* for each of {S^{tv}, M_H^{tv}, and M_L^{tv}}. However, the exact mechanisms may vary across varieties. For instance, if S^{t1} were "\$5; Rolling a 2 or a 3 or a 4" then S^{t2} could be "\$5; Flipping heads on one coin" and S^{t3} could be "\$5; Drawing a black card." The dollar amounts and probabilities for each of the option types are consistent across varieties, but the exact mechanisms (e.g., coin flips vs. card draws vs. die rolls) can vary.

Within each trial variety, there are two levels of M transparency (i.e., high and low). For the high transparency level, M_{Hh} and M_{Lh} share a domain (e.g., coin flips vs. card draws vs. die rolls) and the set of probabilistic events that constitute M_{Lh} is contained in the set of probabilistic events that constitute M_{Hh}. For instance, if M_{Hh} were "$6; Rolling a 1 or 2 or 5 or 6" then M_{Lh} could be "$6; Rolling a 1 or 2 or 5." On the other hand, for the low transparency level, M_{Hl} and M_{Ll} did not share a domain and thus, M_{Ll} was not contained in M_{Hl}, but $p(M_{Hl})$ remained larger than $p(M_{Ll})$. For instance, if M_{Hl} were "$6; Rolling a 1 or 2 or 5 or 6" then M_{Ll} could be "$6; Drawing a black card." Therefore, there are 10 (trial sets) * 3 (trial varieties) * 2 (transparency levels) = 60 varieties of each of S, M_H, and M_L. A full list of possible trials is available in our Research Box.

Choices

For each of 10 trial sets *t*, a given participant completed the following trials: (1) low transparency: S^{t1} vs. M_H^{t1} or M_L^{t1} (L), where M_H^{t1} and M_L^{t1} use different mechanisms; (2) high transparency: S^{t2} vs. M_H^{t2} or M_L^{t2} (H), where $M_H^{t2}$and M_L^{t2} use the same mechanism; and (3) control 1: binary choice between S^{t3} and M_H^{t3} (L), where the mechanism (card draw, die roll, or coin flip) for M_H^{t3} is the same as the mechanism for M_H^{t1} (L). If the mechanisms for M_H^{t1} (L) and M_H^{t2} (H) were different, then the participant completed a 4th trial for that set: (4) control 2: S^{t3} vs. M_H^{t3} (H), where the mechanism for M_H^{t3} (H) is the same as the mechanism for M_H^{t2} (H). Across participants, the assignment of varieties within each trial set (i.e., whether the *i* in S^{ti} is 1 or 2 or 3 for each of the four types of choices) was randomized. Additionally, participants were assigned to one of two between-subjects presentation conditions; the odd (even) numbered sets presented the "M_H or M_L" option as "M_H or M_L" and the even (odd) numbered sets presented the "M_H or M_L" option as "M_L or M_H" to control for presentation-order effects. In one case where S^{t3} was identical to M_H^{t3}, the appropriate control trial was excluded (as we could use its noiseless expected value of 0.5 instead).

In addition to the main trials of interest, participants also completed three attention-check questions, in which one option was designed to be (trivially obviously) dominant over the other according to monetary winnings, probability, or both. These three trials were spaced so that they occurred roughly one-fourth, one-half, and three-fourths of the way through the survey. All of the main trials of interest were presented in a fully-randomized order.

At the end of the survey, participants also completed the two corresponding M_H vs. M_L trials (i.e., M_H^{t1} vs. M_L^{t1} and M_H^{t2} vs. M_L^{t2}) from one of the 10 trial sets.

Exclusions and Data Preprocessing

As specified in our preregistration, we excluded anyone who picked a dominated option in any of the three attention-check questions. This resulted in the exclusion of 104 participants, leaving us with a sample size of 194.

Results

Results (Preregistered)

We analyzed the results as in Experiment 5 and found undervaluation, $M = 0.06$, 95% CI = [0.03, 0.09], $t(193) = 4.28$, $p < .001$. We tested for an effect of transparency as in Experiment 5 and find none, $M = 0.002$, 95% CI = [–0.02, 0.02], $t(193) = 0.12$, $p = 0.91$.

Using that approach, we find that 49% of participants exhibited undervaluation, 20% showed no difference, and 31% showed an effect in the opposite direction. On the high-transparency trials, 48% exhibited undervaluation, 21% exhibited no difference, and 30%

showed an effect in the opposite direction. On the low-transparency trials, 51% exhibited undervaluation, 19% showed no difference, and 30% showed an effect in the opposite direction. At the choice set level, we find that all 10 of the choice sets exhibit undervaluation, 9/10 exhibit undervaluation on high-transparency trials, and all 10 exhibit undervaluation on low-transparency trials.

Exploratory Results

We tested for within-subject consistency in undervaluation by correlating the undervaluation in high-transparency trials with the undervaluation in low-transparency trials. We found a significant positive relationship, $r = 0.67$, $t(192) = 12.46$, $p < .001$. Participants who displayed more undervaluation in high-transparency trials also displayed more undervaluation in low-transparency trials.

Appendix H: Experiment 4c

Introduction

In Experiment 4c, we investigate possible connections between undervaluation and various individual difference measures: risk aversion, analytic-holistic thinking (Choi et al. 2007), and elaboration on potential outcomes (Nenkov et al. 2008). We predicted that greater undervaluation would be associated with more holistic thinking (i.e., treating the "M_H or M_L" option as a holistic unit, rather than as its component parts) and less elaboration on potential outcomes.

In addition to investigating individual differences, we also sought to rule out some alternative explanations for the effect. First, it is possible that people choose S more when it is paired with "M_H or M_L" because people have a strong aversion to multi-option alternatives. If this were the case, we would expect them to be more likely to choose M_L (over S) than "M_H or M_L."

Method

Participants

For this preregistered experiment (https://researchbox.org/124&PEER_REVIEW_passcode=WUXZVT), we collected responses from 298 Amazon Mechanical Turk workers. They earned $2.50 for their participation. Five randomly-selected participants received the outcome of one of their decisions, as detailed below.

Materials and Procedure

This experiment was very similar to Experiment 4b, with the following changes. Participants completed S vs. M_L choices and trinary choices in addition to binary and multi-option choices. Moreover, at the end of the experiment, participants completed 5 incentivized risk aversion trials (Holt and Laury 2002). They also completed the analytic-holistic thinking scale (Choi et al. 2007) and the elaboration on potential outcomes scale (Nenkov et al. 2008).

Additional Trial Details

In this experiment, participants made 43 incentivized choices (including 3 attention check questions and 5 risk aversion measures); some were control binary choices (e.g., S vs. M_H), some were trinary choices (e.g., S vs. M_H vs. M_L) and some were test choices (e.g., S vs. M_H or M_L). For each of 10 trial sets *t*, a given participant completed the following trials: (1) low transparency: S^{t1} vs. M_H^{t1} or M_L^{t1} (L), where M_H^{t1} and M_L^{t1} use different mechanisms; (2) high transparency: S^{t2} vs. M_H^{t2} or M_L^{t2} (H), where M_H^{t2} and M_L^{t2} use the same mechanism; (3) trinary: S^{t3} vs. M_H^{t3} vs. M_L^{t3}; (4) SM_L: S^{t1} vs. M_L^{t1} (or S^{t2} vs. M_L^{t2}); (5) M_HM_L: M_H^{t2} vs. M_L^{t2} (or M_H^{t1} vs. M_L^{t1}); (6) control 1: binary choice between S^{t3} and M_H^{t3} (L), where the mechanism (card draw, die roll, or coin flip) for M_H^{t3} (L) is the same as the mechanism for M_H^{t1} (L). If the mechanisms for M_H^{t1} (L) and M_H^{t2} (H) were different, then the participant completed a seventh trial for that set: (7) control 2: S^{t3} vs. M_H^{t3} (H), where the mechanism for M_H^{t3} (H) is the same as the mechanism for M_H^{t2} (H).

Participants were assigned to either complete trial sets 1-5 or trial sets 6-10. For trial types 3 and 4 (i.e., trinary and SM_L), participants were randomly assigned in each set to complete the first or second versions of the trials, as detailed above. The version for M_HM_L was set to be the opposite of the version participants completed for SM_L. The full list of possible trials is available in our Research Box.

At the end of the experiment, participants completed the two corresponding M_H vs. M_L trials from one of the five trial sets that they completed. They also completed 5 risk aversion trials (Holt and Laury 2002). Finally, they completed the analytic-holistic thinking scale (Choi et al. 2007) and the elaboration on potential outcomes scale (Nenkov et al. 2008). As in Experiment 4, at the end of data collection, we randomly selected five participants and then randomly selected one of their choices to play out. In this experiment, two of the five participants won money ($2 and $0.10).

Exclusions and Data Preprocessing

As specified in our preregistration, we excluded anyone who picked the dominated option in any of the three attention-check questions. This resulted in the exclusion of 122 participants, leaving us with a sample size of 176.

Results

Preregistered Results

We tested for undervaluation using the approach in Experiment 4b, and find evidence for undervaluation, $M = 0.06$, 95% CI = [0.02, 0.09], $t(175) = 3.37$, $p < .001$. We tested for moderation by transparency as in Experiment 2b and do not find evidence of the moderation, $M = 0.002$, 95% CI = [–0.04, 0.04], $t(175) = 0.11$, $p = 0.91$. 55% of participants exhibited undervaluation, 19% exhibited no difference, and 27% exhibited an effect in the opposite direction. 41% exhibited undervaluation on high-transparency trials, and 49% exhibited undervaluation on low-transparency trials. At the choice set level, 8/10 of the choice sets exhibit undervaluation overall, 8/10 exhibit undervaluation on high-transparency trials, and 8/10 exhibit undervaluation on low-transparency trials.

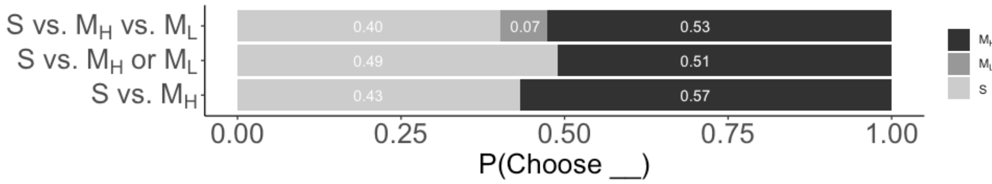
We did not find any evidence of a relationship between undervaluation and risk aversion ($r = 0.003$, $t(174) = 0.03$, $p = 0.97$) or elaboration on potential outcomes ($r = 0.007$, $t(174) = 0.10$, $p = 0.92$). We found a small negative correlation between undervaluation and analytic/holistic thinking ($r = –0.17$, $t(174) = –2.28$, $p = 0.02$), which suggests that participants who reported being more holistic thinkers displayed less undervaluation. However, as this relationship is small, in the opposite direction as anticipated, and was one of several possible measured individual differences (increasing the probability of false positives), we do not strain to interpret it here.

With this data, we were also able to address two possible alternative explanations for our results. First, one possible explanation is that the undervaluation effect is due to blind avoidance of "M_H or M_L" options. Comparing choice proportions in the multi-option alternative choices to choice proportions in the S vs. M_L choices allows us to address this possibility. People choose "M_H or M_L" more often than they choose M_L, which indicates that they do not blindly avoid the "M_H or M_L" options, $M = 0.18$, 95% CI = [0.15, 0.22], $t(175) = 10.93$, $p < .001$.

Exploratory Results

Another alternative explanation is that our results are simply due to noisy and/or careless responding. However, when we limit our dataset to participants who never chose a dominated option (i.e., chose M_H in all of the M_H vs. M_L choices), the effect is still there and is even larger, $M = 0.11$, $t(91) = 5.57$, $p < .001$. Similarly, when we further restrict our dataset to participants who also got all of the comprehension questions correct on the first try, we find a similar (stronger) effect, $M = 0.10$, $t(67) = 4.61$, $p < .001$. Thus, when we limit our dataset to more attentive participants, the effect does not get smaller and, if anything, it gets larger.

Figure S5: Proportion of Choices.



Notes: Proportion of choices for each option in binary, test, and trinary choices for Experiment 4c.

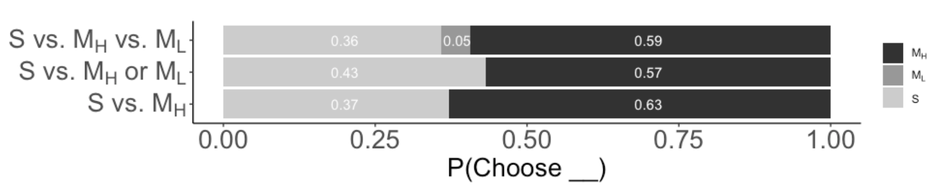
Appendix I: Experiment 5

Trial Generation Process

In this experiment, participants made 65 incentivized choices. Some were control binary choices (e.g., S vs. M_H), some were trinary choices (e.g., S vs. M_H vs. M_L) and some were test choices (e.g., S vs. {M_H, M_L}). For each of 10 trial sets *t*, a given participant completed the following trials: (1) low transparency: S^{t1} vs. M_H^{t1} or M_L^{t1} (L), where M_H^{t1} and M_L^{t1} use different mechanisms; (2) high transparency: S^{t2} vs. M_H^{t2} or M_L^{t2} (H), where M_H^{t2} and M_L^{t2} use the same mechanism; (3) trinary: S^{t3} vs. M_H^{t3} vs. M_L^{t3}; (4) control 1: binary choice between S^{t3} and M_H^{t3}, where the mechanism (card draw, die roll, or coin flip) for M_H^{t3} is the same as the mechanism for M_H^{t1}. If the mechanisms for M_H^{t1} and M_H^{t2} were different, then the participant completed a fifth trial for that set: (5) control 2: S^{t3} vs. M_H^{t3}, where the mechanism for M_H^{t3} is the same as the mechanism for M_H^{t2}.

Participants completed all 10 trial sets. At the end of the survey, participants completed a rank-ordering choice (of S, M_H, and M_L) within each set with the same mechanisms as (1), above. They also completed the two corresponding M_H vs. M_L trials (i.e., M_H^{t1} vs. M_L^{t1} and M_H^{t2} vs. M_L^{t2}) from one of the trial sets that they completed. As in Experiment 4, at the end of data collection, we randomly selected five participants and then randomly selected one of their choices to play out. In this study, all five participants won money ($3, $5, $4, $2 and $4).

Figure S6: Proportion of Choices.



Notes: Proportion of choices for each option in binary, test, and trinary choices for Experiment 5.

Appendix J: Heterogeneity Simulation

To simulate behavior in multi-option alternative decisions, we used the data from Experiments 1 and 1b. First, using all control binary choices from the data, we estimated the probability of choosing a film in a binary choice, based on the rating of the film and the rating of its competitor (i.e., using logistic regression we regressed ChooseFilmM_H on RatingS and RatingM_H). Then, using all test choices from the data, we estimated the probability of choosing a multi-option alternative (i.e., M_HM_L), based on the ratings of the single-option alternative and both options of the multi-option alternative (i.e., using logistic regression we regressed ChooseM_HM_L on RatingS, RatingM_H, and RatingM_L). We use these results to estimate the probabilities of choosing single-option and multi-option alternatives as detailed below.

Next, for each unique pair of films (50*49/2 = 1225 pairs), we identified the more-popular film (i.e., the film that was rated higher by more participants, which we will call "Film A") and the less-popular film (i.e., whichever film is not Film A, which we will call "Film B"). For each unique pair of films for each participant, there is also a personally-preferred film (which we call "Film M_H") and a personally-unpreferred film (which we call "Film M_L").

We also identified (using participant-level ratings of S, M_H, and M_L and the estimates from the regressions identified above) at the participant level: (1) the probability that Film A would be chosen, (2) the probability that Film B would be chosen, (3) the probability that Film M_H would be chosen, (4) the probability that Film M_L would be chosen, and (5) the probability that a multi-option alternative comprising both films (Films A and B; a.k.a. Films M_H and M_L) would be chosen. We then averaged each of the five measures above across participants to get average probabilities for each pair of films.

Next, we considered the results assuming different levels of targeting capabilities. First, we consider the situation where we have no ability to target consumers based on their preferences. In this case including both films as options maximizes probability of choice. In other words, we expect "overvaluation" of the multi-option alternative, relative to the more-popular option (Figure 8). Second, we consider the situation where we are able to identify the personally-preferred film for a given consumer with some probability (*w*). In this case, we expect to observe probability (3) above in *w**100% of consumers and observe probability (1) above in (1-*w*)*100% of consumers on average. To represent this in our simulations, we randomly selected which consumers' preferences would be accurately identified (using a binomial distribution with probability *w*). We repeated this random selection 1000 times, averaging at the movie-pair level. As *w* increases, and as offerings are tailored to individuals, expected undervaluation of the multi-option alternative increases as well.

Appendix K: Supplementary Tables

Table S1: Summary of Choice Trials Across All Experiments.

Exp.	Total Number of Choices	S vs. M_H or M_L	S vs. M_H	S vs. M_H vs. M_L	S vs. M_L	M_H vs. M_L	Attention Check Questions	Other
1 1b 1c	30	10	10	0	0	0	0	10 filler binary questions
2 2b	30	10	10	0	0	0	0	10 filler binary questions; post-choice ratings: S, M, M_H, M_L
3 3b	42	5	5	0	0	0	0	17 filler binary questions, 5 1v3, 5 2v2, 5 2v3
4	34	10	10	10	0	1	3	
4b	39-40	20	14-15	0	0	2	3	
4c	43	10	10	5	5	5	3	risk aversion, analytic / holistic thinking, elaboration on potential outcomes
5	65	20	20	10	0	2	3	10 rank-order questions

Table S2: Full Regression Results.

Exp.	Equation Number	Equation	Coefficients	SE	t value	p-value
1	1	Choice Difference ~ b0 + b1*S + b2*(MH+ML) + b3*(MH-ML)	b0 = –0.05 b1 = 0.01 b2 = –0.0004 b3 = 0.02	0.02 0.002 0.001 0.003	–2.64 3.40 –0.27 5.70	.008 <.001 .79 <.001
4	2	Choice Difference ~ b0 + b1*EV(S) + b2*(EV(MH)+EV(ML)) + b3*(EV(MH)-EV(ML))	b0 = 0.04 b1 = –0.005 b2 = 0.001 b3 = 0.01	0.03 0.004 0.003 0.008	1.64 –1.22 0.38 1.39	.10 .22 .71 .16
4	3	Choose S ~ b0 + b1*EV(S) + b2*EV(MH) + b3*M(S) + b4*M(MH) + b5*P(S) + b6*P(MH)	b0 = –0.25 b1 = 0.42 b2 = –0.24 b3 = 0.19 b4 = –0.28 b5 = 2.96 b6 = –2.83	0.72 0.12 0.15 0.06 0.10 0.77 0.91	–0.35 3.49 –1.61 3.09 –2.98 3.83 –3.12	.73 <.001 .11 .002 .003 <.001 .002
4	4	Choice Difference ~ b0 + b1*P(S) + b2*(P(MH)+P(ML)) + b3*(P(MH)-P(ML))	b0 = 0.02 b1 = –0.01 b2 = 0.004 b3 = 0.11	0.04 0.04 0.03 0.06	0.57 –0.40 0.14 1.80	.57 .69 .89 .07