Human perceptuo-motor decision-making is not optimal

Tyler J. Adkins[1], Richard L. Lewis[1,2,3], Taraz G. Lee[1,*]

[1]Department of Psychology, University of Michigan
[2]Department of Linguistics, University of Michigan
[3]Weinberg Institute for Cognitive Science, University of Michigan

*Corresponding author information:
tarazlee@umich.edu
Department of Psychology, University of Michigan
530 Church st., Ann Arbor, MI, 48109, USA

Abstract

Economic decision-making under risk often fails to maximize expected value, perhaps reflecting cognitive biases and heuristics. A line of recent work argues that when economic decisions are reformulated as decisions about where to aim rapid reaching movements, the resulting decisions are optimal (i.e. maximize expected value), in dramatic contrast to the standard findings obtained with decisions among gambles. These arguments for optimality rely on a comparison between human performance and the performance of an ideal agent in a reaching task with a narrow range of incentive values. Here, we improve on this methodology, both empirically and analytically, by devising a task with a wider range of incentive values, and by performing trial-level comparisons of participant behavior to the Ideal Model *and* two plausible alternative models. The first alternative, the Loss-averse Model, incorporates participant-level loss aversion, and the second alternative, the Heuristic Model, embodies a simple satisficing heuristic that is invariant to incentive magnitudes. Across two experiments, we find that participants' movement outcomes are more likely under the Heuristic Model than the Ideal or Loss-averse models, especially when the risk of financial loss is high. The present work provides evidence against the claim that perceptuo-motor decision making is optimal and demonstrates the fruitfulness of including alternative models in analyses of human behavior.

Decision making often entails the risk of a tangible loss. For example, a brain surgeon who needs to decide whether to operate on a patient with a tumor must weigh the benefits of a successful surgery with both the risks involved in surgical intervention and in choosing a more conservative treatment. However, once the decision to operate is made, the act of removing a small tumor presents another, pereceptuo-motor, decision making problem. During the operation, the surgeon must carefully plan where to aim his movements, since a slip of the scalpel could result in the death or severe impairment of the patient. A recent line of interesting work argues that when choices among risky financial gambles are framed as choices among risky *aimed movements* people make decisions that appear unaffected by the biases and heuristics that often lead to suboptimal economic decisions[1,2,3,4,5]. This claim is supported by analyses which compare human performance in an incentivized reaching task to the performance of an ideal (expected value-maximizing) model. As we review below, these analyses have significant theoretical appeal because, armed with only an estimate of the participants' movement variance, they permit detailed quantitative predictions of human behavior across a multitude of task conditions. Nonetheless, this line of work is limited in that the task employs a small range of incentive values and participants' behavior is compared to a single, ideal model.

In the present work, we improve on the above methodology, both empirically and analytically, by devising a variant of the reaching task with a wider range of incentive values and by performing trial-level comparisons of participants' behavior to the Ideal Model as well as two plausible alternative models deviating from the ideal. Across two experiments, we find that participants' movement outcomes are more likely under a simple heuristic model than the Ideal (expected-value-maximizing) Model, especially when the risk of financial loss is high. The present work provides evidence against the claim that perceptuo-motor decision making is

optimal and demonstrates the fruitfulness of including alternative models in analyses of human behavior.

A large body of research indicates that economic decision making under risk—such as a choice among gambles characterized by probabilities of gains and losses—is influenced by several biases and heuristics that lead people to fail to maximize expected value (EV)[6,7,8 9 10]. In particular, people often misrepresent (that is, introduce biases in) the rewards, penalties, and probabilities involved in economic decision-making problems. For example, loss-averse individuals give outsized weight to potential penalties and often pass on objectively advantageous opportunities[11]. Heuristics, on the other hand, are simple decision-making rules or strategies that people leverage to make quick and minimally effortful, yet often satisfactory decisions[12,9,13]. For instance, in the face of many alternative choices, people may display *satisficing* behavior in which they select the first option that meets a minimum level of acceptability rather than making the optimal choice[13]. Theoretical explanations of various departures from normative optimality differ substantially in the degree to which they reject or embrace utility maximization in their analysis[14], but the empirical departures themselves are widely accepted.

A number of recent studies have argued that human perceptuo-motor decision-making is immune to these biases and heuristics and appears to be normatively optimal[1,2,3,4,5]. In these studies, participants performed speeded reaching movements to small spatial targets and gained or lost money based on the outcomes of their reaches. In this reaching task, each possible aim point can be treated as a gamble with specific probabilities and payoffs. Therefore, the task can be analyzed as a choice among a set of gambles (Figure 1)[5]. The virtue of this task design is that

it is mathematically equivalent to risky economic decision tasks in which people typically display loss aversion or exhibit other biases that lead to failures to maximize EV.

In dramatic contrast to the classic behavioral economics studies of decision-making under risk, prior work suggests that peoples' reaching behaviors appear consistent with the performance of an ideal agent that selects the optimal aim point (or lottery), i.e. the one that maximizes EV[1,2,3,4,5]. In this work, *consistency* with the ideal behavior is assessed in two ways. The first way is to compute an efficiency measure: the percentage of optimal payoff obtained. The second way is to compare participants' aim points (i.e., mean reaching end-points) with the aim points predicted by an ideal model[1,2,3]. However, these quantitative comparisons to a *single* ideal model fail to consider whether the observed behavior might be *better* accounted for by plausible alternative models that deviate from the ideal. The model comparison methodology we advance here addresses this limitation.

In this paper, we aim to rigorously test the claim that human perceptuo-motor decision-making is optimal by comparing three computational models of human performance in an incentivized reaching task, which we refer to as the Ideal Model, the Loss-averse Model, and the Heuristic Model (see Methods for a complete description of all models). The Ideal Model predicts that a participant will select aim-points that maximize EV, given their reaching end-point variability measured during training. The Loss-averse Model is identical to the Ideal Model except that potential losses are weighted according to a participant-level loss aversion parameter assessed in a separate task. We hypothesized that if loss aversion has been eliminated by transforming economic decisions into perceptuo-motor decisions, then the Ideal Model will provide a more accurate characterization of behavior than an otherwise equivalent model that incorporates participant-level loss aversion. The Heuristic Model is motivated theoretically by

the following observation: the reaching task presents participants with an infinite set of aim-points to select from, and the optimal aim-point changes from condition to condition as rewards, penalties, and region placement varies. In this way, the reaching task differs from classic gambling tasks, which present a small finite number of options. One possible heuristic response to the demands of the reaching task is to *satisfice* by adopting a single aim-point that doesn't vary across payoff conditions or across people, that is easily localized perceptually, and that nonetheless achieves satisfactory performance[13]. We formulate the Heuristic Model as one that adopts a simple mid-point calculation between the penalty and reward boundaries.

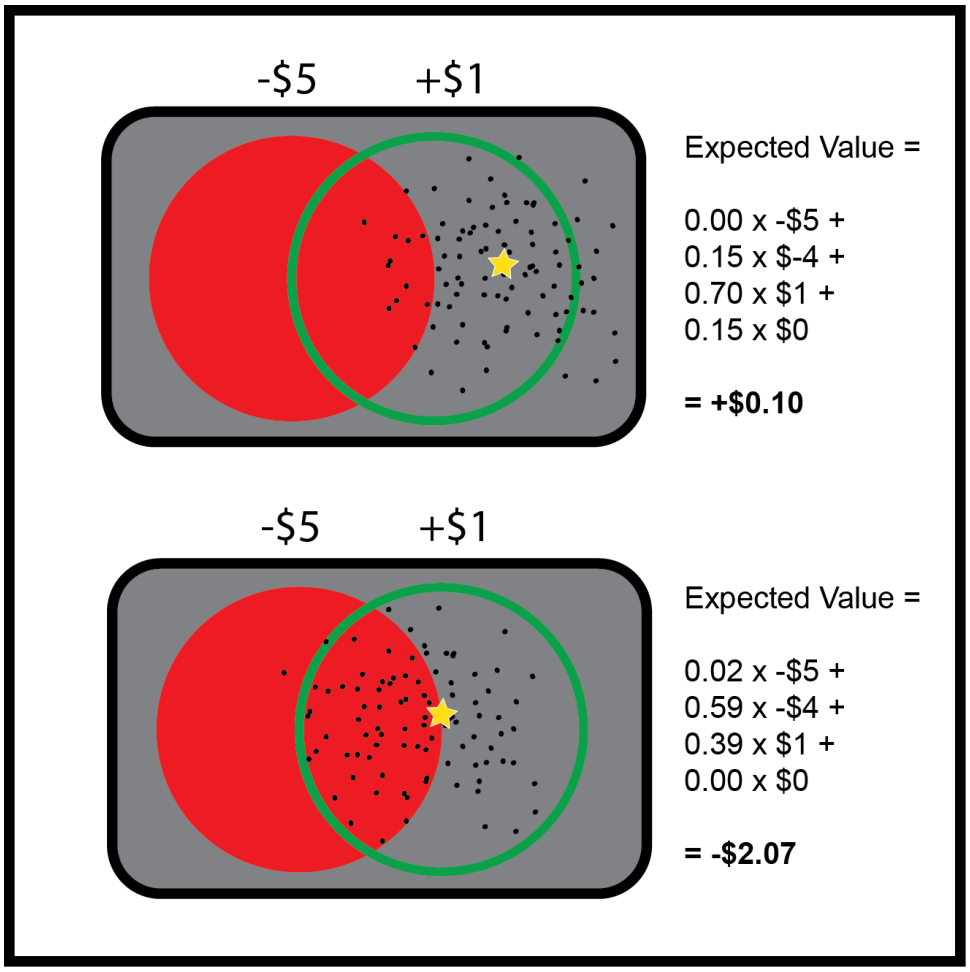

Figure 1 Equivalence between classic decision-making under risk and perceptuo-motor decision-making tasks. Participants must decide where to aim (yellow stars) before rapidly reaching out to touch the screen in order to earn cash bonuses (green circle) while avoiding penalties (red circle). Due to each participant's inherent movement variability, successive reaches (black dots) cluster around the aim point. The expected value (EV) of reaching toward the aim point in (a) is $0.10, while the EV for aiming at the aim point in (b) is -$2.07. Given the participant can aim anywhere on the screen, each with its own EV, the decision about where to aim for a given payoff configuration is equivalent to choosing between an infinite set of lotteries/gambles. The optimal strategy is to select the aim-point that maximizes expected gain, given one's movement variability.

All three models share a crucial analytic property that permits their application to the analysis of individual, trial-level data: each model makes precise predictions of aim points for each condition (and thus each trial). Given an approximation of each participant's variability, it is possible to compute the likelihoods of observing a particular end point in each trial conditioned on the predictions of each model. This serves as the basis for detailed quantitative comparisons between the three models (see Methods). If either the Loss-averse Model or the Heuristic Model provides a significantly better quantitative fit than the Ideal Model, this is evidence against the claim that human perceptuo-motor decision-making under risk is optimal.

In Experiment 1, we found that the Ideal Model provided a better account of the reaching end-points than the Loss-averse Model when the large penalties were at stake, but that the Heuristic Model provided an even better fit to the data than the Ideal Model. However, we developed the Heuristic Model after examination of the Experiment 1 results, and so to confirm the validity of the Heuristic Model, we replicated the result in Experiment 2 using the same task parameters as the original set of studies by Trommershauser and colleagues[1,2]. In Experiment 2, participants' behavior in response to the condition with large penalties was again better described by the Heuristic Model than the Ideal Model, and this was true for a majority of the participants. These findings suggest that people do not consistently rely on an optimal strategy to guide their

decisions about where to reach, but instead, are more likely to use a simple heuristic that is insensitive to the potential payoffs and probabilities imposed by their reaching variability.

Results

In Experiment 1, 19 participants performed a modified iteration of the incentivized reaching task that occasionally presented high-stakes payoff conditions with doubled rewards and penalties (see Figure 2 and Methods). After training on a separate day, participants returned to perform the task with real cash bonuses. They gained and lost money based on where and when they completed their reaching movements. If, before a strict 1s time limit, a participant touched within the boundary of the green "reward" circle or the red "penalty" circle, they received the associated reward or penalty, respectively. There were five payoff conditions throughout Experiment 1: [-$0, +$1]; [-$1, +$1]; [-$5, +$1]; [-$3, +$3]; and [-$15, +$5]. We expected that participants' performance would deviate most from optimal on these high-stakes trials because prior work shows that cognitive and perceptuo-motor abilities are especially prone to failure as incentive values grow large[15,16,17]. Also, the Ideal and Loss-averse models make distinct predictions about behavior on these high-stakes trials, so these trials allow us to empirically distinguish these models.

Trial-by-trial performance for each participant was compared to the predictions of the three models. The Ideal Model assumes participants use an optimal strategy whereby they select aim-points with the highest EV, conditional on their own movement variability (estimated from training). If this model has the best fit to our participants' behavior, then it is more likely that our participants used the optimal ideal strategy than the sub-optimal loss-averse or heuristic strategies. The Loss-averse Model was identical to the Ideal Model except that the penalty size was scaled up to account for each participant's trait level loss aversion (see Methods). If this

model has the best fit to our participants' behavior, then it is more likely that our participants used a loss-averse strategy than the ideal or heuristic strategies. The last computational model we test, the Heuristic Model, assumes that participants in this task simply choose to aim at a single satisfactory point when penalties are present and at the center of the reward circle when they are not. This model is insensitive to the participant's movement variability and the exact values displayed and thus gives the same predictions for all participants. If this model has the best fit to our participants' behavior, then it is more likely that our participants used this sub-optimal heuristic strategy than the ideal or loss-averse strategies. (See Figure 3 for the predictions of each model for a representative participant.)

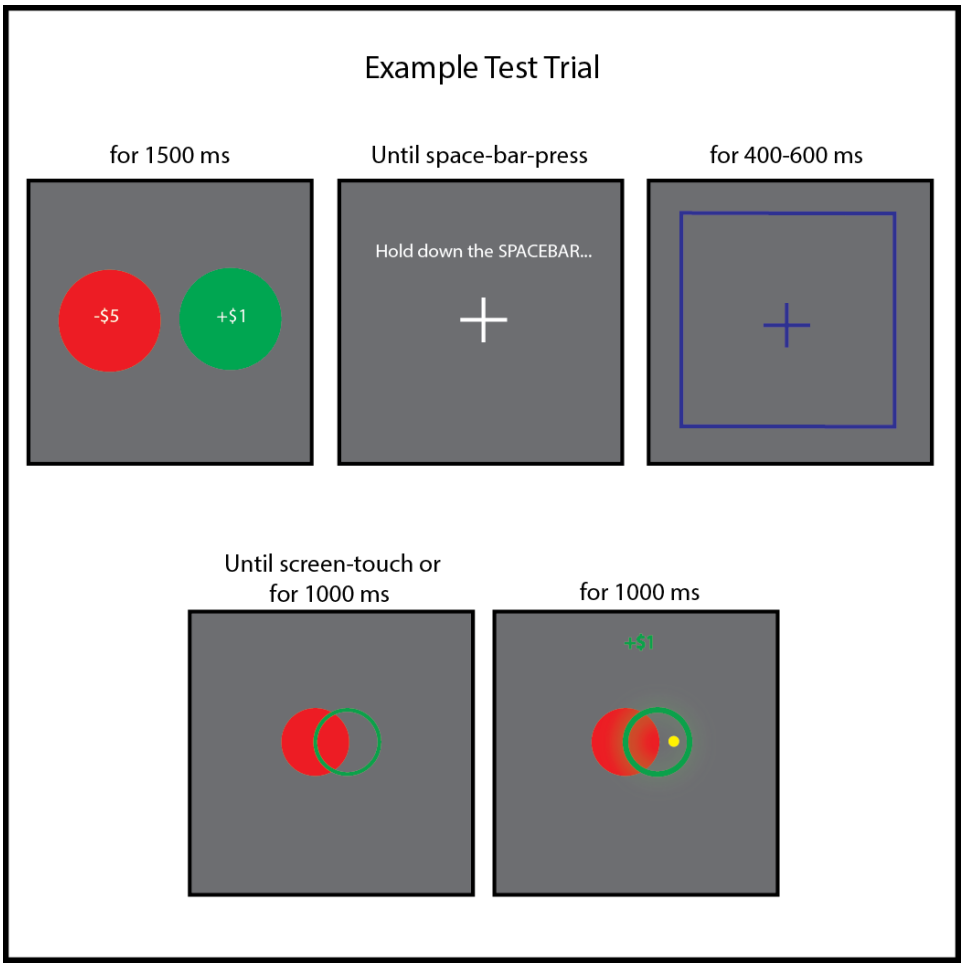

Figure 2 Timeline of an example test trial. The participant was first presented with information about the potential penalty and reward values. Next, they were presented with a fixation cross, which remained until they pressed and held the space bar down. Immediately after the space bar was pressed, a blue bounding box appeared signifying the area within which the payoff configuration could appear. Next, they had 1s to reach and touch the screen. They were then presented with feedback for 1 second about their earnings from the current trial.

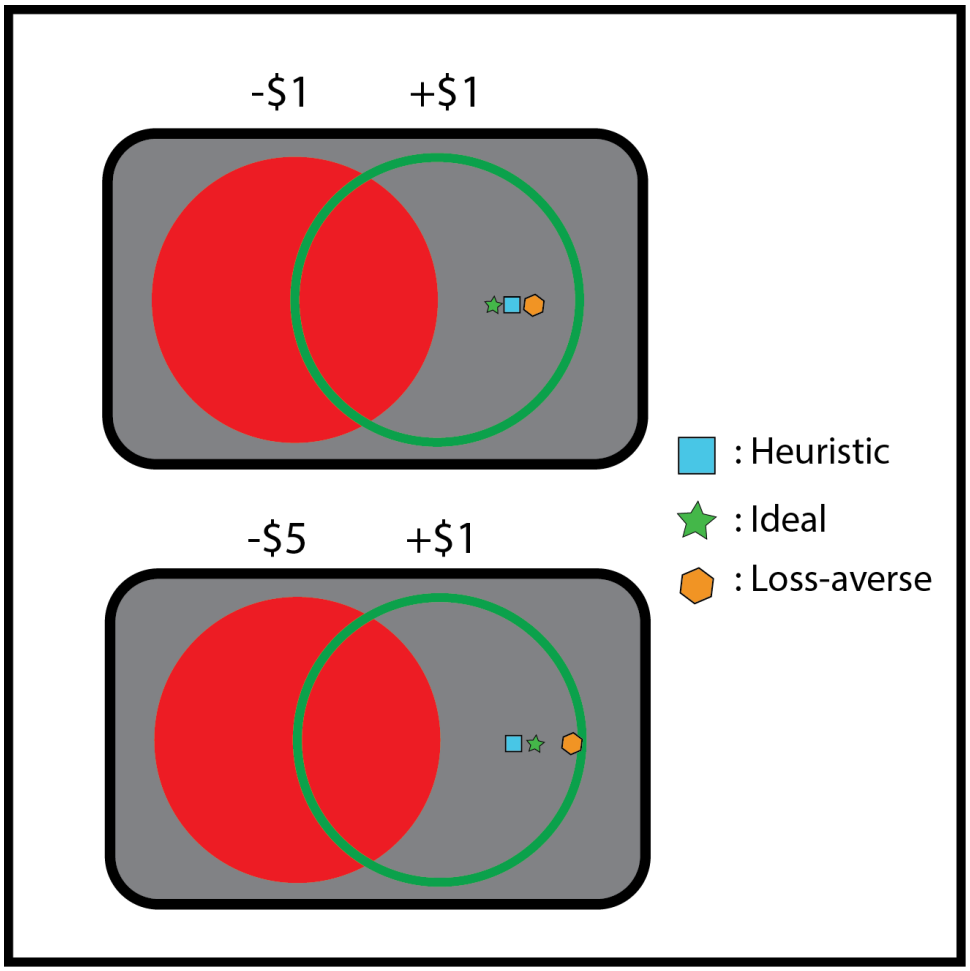

Figure 3: Example aim-point predictions of each model. The predicted aim-points of the ideal and loss-averse models shift away from the penalty circle as the size of the penalty increases, though the Loss-averse Model does so to a greater extent. The predicted aim-points of the Heuristic Model do not change as the penalty increases. All models predict no shift in the vertical direction. Predictions shown here are from a representative participant.

The ideal model provides a better fit to observed end-point data than the Loss-averse Model.

While previous studies have shown that human performance in risky pereceptuo-motor decisions is similar to that of an ideal agent, in order to claim that behavior is optimal it is necessary to show that performance is not *more* similar to that of a sub-optimal agent. A common measure of optimality that has been used in prior work is "efficiency", which is simply the ratio of a participant's actual earnings to the earnings of an ideal agent (with the same movement variability). A high level of efficiency (e.g. 90+%) has been used as evidence that perceptuo-motor decision-making is optimal[1,2,3,20]. In our first experiment, we found that participants were efficient relative to both models, obtaining a mean efficiency of 101% (SD = 2.47) relative to the Ideal Model, and 102% (SD = 4.12) relative to the Loss-averse Model. Given that there was no significant difference between these scores ($t_{18} = 1.36$, $p = 0.19$), this efficiency measure does not help to discriminate between the two strategies.

A more sensitive measure of optimality is the relative quantitative fit of the Ideal and Loss-averse models to our participants' reaching end-points. Accordingly, we found that the Ideal Model provided a significantly better fit than the Loss-averse Model to the aggregated end-point data (1R: Vuong's statistic $V = 14.9$, $p < 0.001$; 1.4R: $V = 4.9$ $p < 0.001$). Analyzing the fit separately for different conditions provides further insight into why the models are performing differently. When there was substantial overlap between the reward and penalty region (1R separation), the Ideal Model was superior regardless of the specific reward and penalty values ([-\$1, +\$1]: $V = 3.7$, $p < 0.001$; [-\$3, \$3]: $V = 3.4$, $p < 0.001$; [-\$5, +\$1]: $V = 12.3$, $p < 0.001$; [-\$15, +\$3]: $V = 12.7$, $p < 0.001$). However, when the separation distance was increased to 1.4R, the Ideal Model was only superior to the Loss-averse Model when the ratio of penalty to reward was -5:1 ([-\$5, +\$1]: $V = 7.0$, $p < 0.001$; [-\$15, +\$3]: $V = 5.4$, $p < 0.001$). In contrast, when the

penalty and reward values were equivalent, the performance of the two models was not statistically distinguishable ([-$1, +$1]: V = -0.5, p > 0.05; [-$3, +$3]: V = -1.1, p > 0.05).

Given these results, we might tentatively conclude that participants' behavior is approximately optimal in this task. We did find that our participants were somewhat sensitive to the reward and penalty values, as paired t-tests showed that their aim-points shifted further from the penalty region as the penalty-to-reward ratio grew (-5:1 vs. -1:1 - t_{15} = 2.1, p = 0.05). However, our participants were not especially sensitive to their own movement variability (as measured in training). There was not a significant rank-order correlation between this variability and their aim-points in the test session as is predicted by the Ideal Model (1R separation: rho = -0.22, p = 0.37; 1.4R separation: rho = -0.03, p = 0.90). Furthermore, when the penalty to reward ratio was -5:1, all participants systematically aimed closer to the penalty region than predicted by the Ideal Model, and no participants aimed further away from the penalty region than predicted by the Ideal Model. It therefore seemed plausible that our participants were simply aiming for a single easy-to-spot location for all conditions that had a nonzero penalty value. We thus devised the Heuristic Model as an alternative account of the strategy that participants use to decide where to reach (see Methods).

The Heuristic Model provides a better fit to observed end-point data than the Ideal Model.

The Heuristic Model provided a significantly better fit than the Ideal Model with respect to the aggregated data (1R: V = -5.8, p < 0.001; 1.4R: V = 5.7, p < 0.001). Again, analyzing the fit separately for different conditions provides further insight into why the models are performing differently. Although both models performed similarly at the lower penalty values when there was a 1 radius distance between the reward and penalty regions ([-$1, +$1]: V = 1.3, p = 0.2; [-

$3, +$3]: V = 1, p = 0.3), the Heuristic Model provided a substantially better fit than the Ideal Model when the penalty-to-reward ratio grew to -5:1 ([-$5, +$1]: V = -8.0, p < 0.001; [-$15, +$3]: V = -8.0, p < 0.001). The mean fits for each model by condition for the 1R separation distance are shown in Figure 4. Since the Heuristic Model provides the best account of our participants' reaching end-points, especially when the penalty to reward ratio is -5:1, we believe our participants in Experiment 1were more likely using a satisficing heuristic than the optimal strategy.

In terms of efficiency, our participants earned 102% of what a satisficing agent would have earned in this task (SD = 3.59). Additional simulations revealed that an individual who followed the Heuristic Model exactly would earn approximately 99.3% of the earnings of an ideal agent. Although previous studies have used high efficiency as evidence for optimality[1,2,3,20], it would be incorrect to conclude that the satisficing individual following the optimal strategy. It is therefore preferable to test for optimality using the relative quantitative fit of the three models to participants' reaching end-points.

The prior analysis asks which model provides the best fit across all participants. Notwithstanding our findings at the aggregate level, it is possible that some participants are better described by the Ideal Model than the Heuristic Model. To summarize the models' relative fits across participants, we counted, for each model, the number of participants whose data (across all payoff conditions) were described significantly better by that model than the others using Vuong's test. No participants were best described by the Loss-averse Model, and only one participant was fit best by the Ideal Model when the penalty region was most offset from the reward region (1.4R). In contrast, the behavior of seven participants (≈36.8% of our sample) was best described by the Heuristic Model for this spatial configuration (five

participants for the 1R condition). The remaining participants were described equally well by all models (Vuong's test p's > 0.05). This again suggests that the Heuristic Model provides the best account of our participants' reaching behavior.

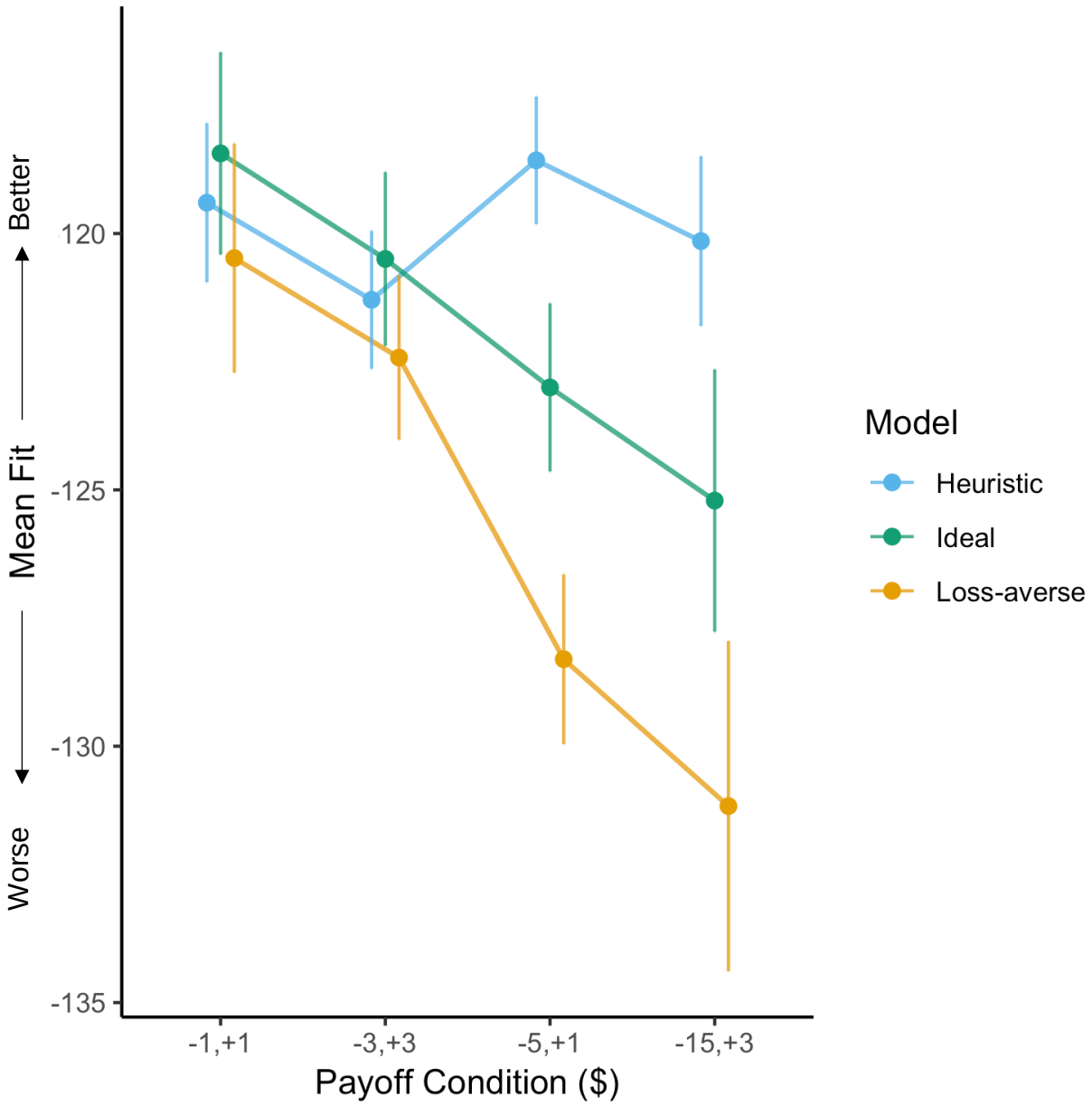

Figure 4 Mean model fits across participants in each payoff condition for Experiment 1 (1R separation distance). A model's "fit" to a participant's dataset in some condition is equal to the sum of the log-likelihoods of all the participant's reaching end-points in that condition, given the model's prediction and that participant's end-point variability from training. Unlike the Heuristic Model, the Loss-averse and Ideal models appear to perform worse as the penalty value increases. All models make identical predictions when the penalty is zero (data not shown). Error bars represent standard errors of normalized mean model fits.

Experiment 2 confirms the Heuristic Model provides the best fit

In Experiment 1, we provided evidence that loss aversion does not affect perceptuo-motor decision-making in this reaching task (as assumed by the Loss-averse Model), given that the Ideal Model was a better fit to behavior than the Loss-averse Model. However, we also provided some evidence that our participants were more likely using a simple satisficing strategy than an ideal, expected-value maximizing strategy. Because we devised the Heuristic Model post-hoc, a proper test of this model required that we replicate the results from Experiment 1 in an independent sample of participants. We therefore ran a second experiment with an additional 16 participants, this time taking using the exact task parameters used in the original incentivized reaching tasks[1,2] to ensure our initial results were not due to small changes made to the stimulus timings or incentive structure (see Methods). If the Heuristic Model again proves to be a better account of our participants reaching behavior, we can conclude that it is unlikely that participants were using the optimal strategy.

We again sought to determine the most likely perceptuo-motor decision-making strategy used by our participants by comparing actual reaching behavior to the behavior predicted by three different computational models. In Experiment 2, the Heuristic Model again outperformed both the Ideal Model ($V = 7.5$, $p < 0.001$) and the Loss-averse Model ($V = 23.0$, $p < 0.001$) on

the aggregated data. When the risk of penalty was relatively low (-100 points), the Ideal Model outperformed the Loss-averse Model ($V = 13.0$, $p < 0.001$) and the Heuristic Model ($V = 6.0$, $p < 0.001$). Yet, when the penalty value was relatively large (-500 points), the Heuristic Model vastly outperformed the Ideal Model ($V = 13.0$, $p < 0.001$) and the Loss-averse Model ($V = 20.0$, $p < 0.001$). The mean model fits across participants by condition are shown in Figure 5. Given these results, it is more likely that our participants used a satisficing heuristic than the optimal strategy, especially when the risk of financial loss was highest.

While the Heuristic Model fit the aggregate data better than the Ideal Model, it was possible that some participants behaved more similarly to the Ideal Model. Nine participants (≈56.25%) were described best by the Heuristic Model while only one participant (≈6.25%) was described best by the Ideal Model. No participant was described best by the Loss-averse Model and six participants (≈37.5%) were described equally well by all three models. This again suggests that, overall, it is more likely that our participants used a simple satisficing heuristic than the optimal strategy.

For each model and participant, we computed the ratio of the participant's actual earnings to what they would earn if they had followed the strategy entailed by the model. Across all participants, the mean efficiency with respect to the Ideal Model was 95.5% with a standard deviation of 4.4%. The mean efficiency with respect to the Loss-averse Model was 103% with a standard deviation of 14.6%. And the mean efficiency with respect to the Heuristic Model was 97.6% with a standard deviation of 3.61%. A one-way repeated measures ANOVA using Type III sums of squares revealed no significant effect of model-type on efficiency ($F(1.02, 15.36) = 3.86$, $MSE = 168$, $p = 0.07$). This similarly suggests that efficiency is a poor indicator of the strategy a participant is using.

We again failed to find a significant rank-order correlation between aim-point (in test) and movement variability (in training) (rho = 0.15, p = 0.57). This suggests that our participants did not use estimates of their training variability to decide where to reach. Although a negative result, this is consistent with the hypothesis that our participants did not use the optimal strategy to decide where to aim their reaching movements and instead opted for a simpler heuristic strategy that is insensitive to their individual reaching variability.

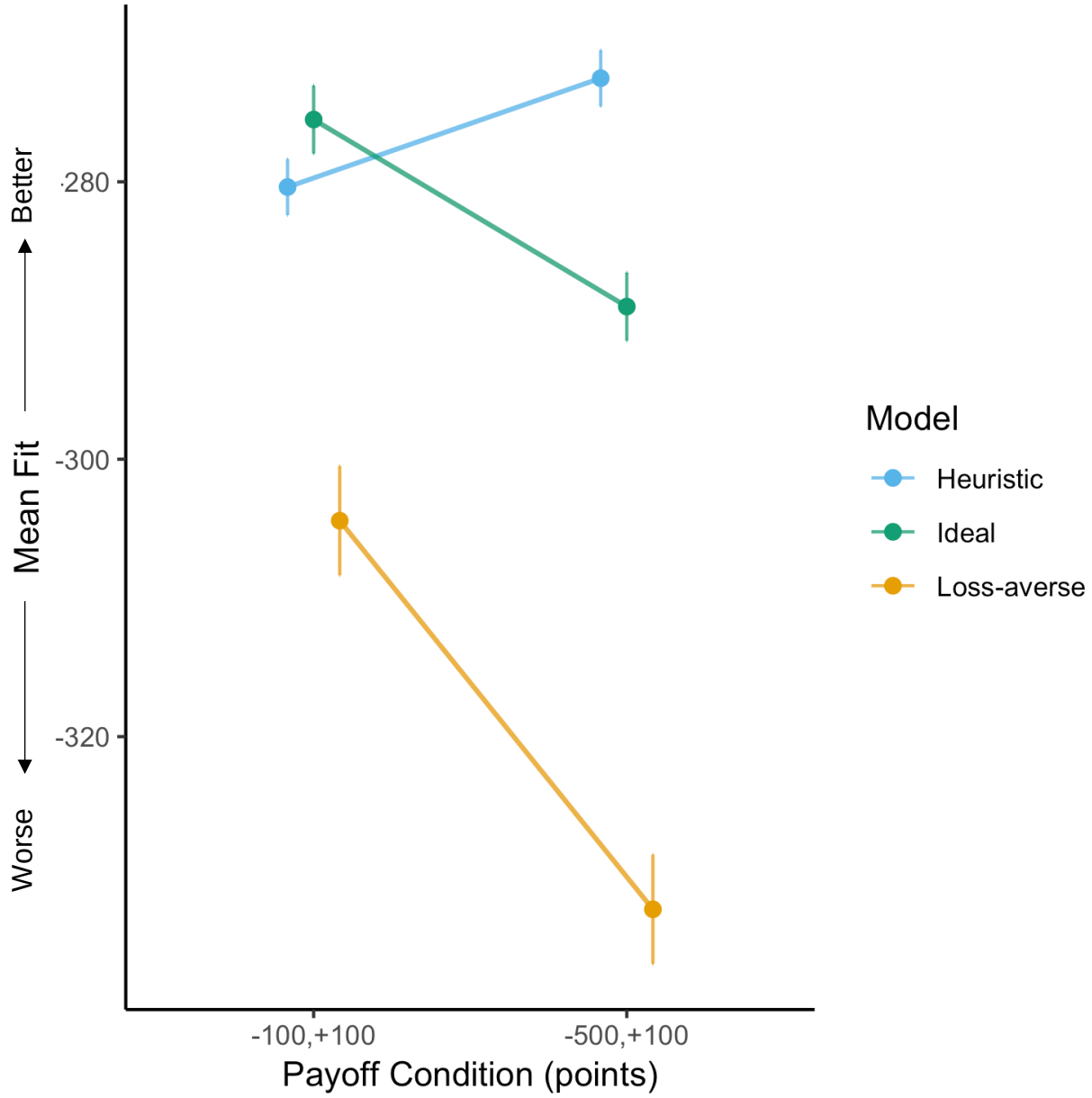

Figure 5 Mean model fits across participants in each payoff condition for Experiment 2. Mirroring the results from Experiment 1, only the Loss-averse and Ideal models perform worse as the penalty value increases while the Heuristic Model does not. All models make identical predictions when there is the penalty is zero (data not shown). The error bars represent standard errors of normalized mean model fits.

Discussion

In contrast to prior work, we provide evidence that people do not use an ideal (expected-value-maximizing) strategy to guide perceptuo-motor decision-making under risk. While previous studies found similarities between human and ideal performance in an incentivized reaching task, we sought stronger and more sensitive tests of the claim that humans use an ideal strategy to perform the task. We therefore compared human performance to the predictions of three different computational models, all of which make specific assumptions about the strategies being used to perform the task. Of particular interest was the fitness of an ideal model compared to two different models that incorporate a known bias (loss aversion) or a known heuristic (satisficing). Across two experiments, we found that the Heuristic Model more closely approximated participants' behavior than the Ideal Model, especially when the there was a risk of a relatively large financial loss. These findings strongly suggest that perceptuo-motor decision-making in this task is not guided by an optimal strategy.

The incentivized reaching task used here was specifically designed to be mathematically equivalent to classical risky economic decision-making tasks[1,2,3,4,5]. Each potential reaching aim-point is equivalent to a gamble with specific probabilities of gains and losses, imposed by the person's reaching variability. Selecting a point to reach to is essentially deciding to accept single gamble from among a virtually infinite set of gambles. It has been previously reported that people select aim-points in this task in a manner consistent with an ideal (expected-value-maximizing) model[1,2,3,4]. However, in tasks in which participants iteratively choose between just a few gambles, they often display biases such as loss aversion[7,11]. And in tasks in which participants simultaneously choose between several gambles, they often rely on sub-optimal satisficing strategies, such as heuristics[13,21]. Here, we provide evidence that people do not behave

optimally in an incentivized reaching task nor do they display loss-averse patterns of behavior. Instead, people appear to rely on a simple satisficing heuristic.

However, participants did display near-optimal patterns of behavior when the potential penalties and rewards were the same magnitude. And prior work has reported that participants' mean reaching end-points are highly correlated with the predicted, ideal aim-points[2]. These findings suggest that some people tend to behave nearly optimally in some contexts, perhaps when the financial stakes are low. Nonetheless, the inconsistency of this optimality across payoff conditions and across individuals suggests that it is not a general feature of human perceptuo-motor decision-making.

The optimal strategy for this task requires that participants consider the exact penalty and reward values, as well as an accurate estimate of their own movement variability. Prior work has shown that people are sometimes sensitive to their own movement variability[22,23,24,25]. Indeed, one previous study using this same task showed that participants adjust their strategy to account for externally imposed variability in their reaching end-points[3]. Yet in our experiments we found that participants' aim-points were not significantly correlated with their own movement variability as would be expected of a group of ideal agents. Our data do not have bearing on whether or not participants have an accurate representation of their own movement variability. However, the lack of a clear relationship between aim-point and movement variability is consistent with the conclusion that decisions about where to reach were not guided by the optimal strategy.

We have provided evidence that perceptuo-motor decision making involving penalties and rewards does not conform to an ideal model as suggested by prior work. The evidence comes from model comparisons that show that participants' behavior in an incentivized reaching

task is better accounted for by a satisficing heuristic than an optimal integration of values and probabilities. This suggests that perceptuo-motor decision-making may not be entirely free of the biases and heuristics that influence behavior in classical economic decision-making tasks.

Future research should examine what underlies the systematic individual variability we found in this task, and advance and test alternative plausible alternatives to the models advanced here—in all cases, using the Ideal Model as a principled baseline, and using a wide range of reward and penalty values. One possible alternative model is a *subjective* ideal model that allows for subjective transformations of probabilities and values, much as Prospect Theory (Kahneman & Tversky, 1979) and related models posit transformations of described probabilities and values. Although the work presented here did succeed in distinguishing the models behaviorally, the analysis revealed a weakness of the existing paradigms: the optimal strategy generates expected payoffs that may not be meaningfully different from the alternative heuristic strategies. Thus, we advocate for the use of empirical paradigms in which the alternative models make measurably distinct payoff predictions. Such paradigms, as well as a rich array of plausible computational models, are required for the most rigorous tests of whether human perceptuo-motor decision-making is optimal.

Methods

Experiment 1

Participants. 19 students (16 female) from the University of Michigan participated in of this study. All participants gave written consent to participate in the study. All procedures were approved by the University of Michigan Institutional Review Board of Health Sciences and Behavioral Sciences. All participants were right handed, free of neurological disorders, and had normal or corrected-to-normal vision. At the end of the experiment, participants were

compensated at a rate of $10/hour plus performance bonuses and their outcome from the loss aversion task described below. A participant's base hourly compensation was not affected by their performance in any task.

Apparatus. Participants sat in an immobile chair in a small room in front of a 23" touchscreen computer monitor, and a keyboard positioned such that the spacebar was 21.5 cm away from the screen. We programmed the reaching task using the Python-based experiment building application, OpenSesame[18]. We programmed the loss aversion task using the PyGame python package.

Reaching task. Participants performed an incentivized reaching task that has been argued to elicit near-optimal movement strategies[1,2,3,4,5]. At the start of each trial (Figure 2), the participant was presented with the reward and penalty values for 1500 ms. Next, the participant looked at a white fixation cross and then pressed and held the space bar with their right index finger when they were ready to begin. Once the space bar was pressed, the fixation cross became blue, and a blue box (88 mm x 88 mm) was presented at the center of the screen for a randomized duration of 400 to 600 ms to indicate the region where the imperative stimulus could appear. Next, overlapping green and red circles (radii = 8.5 mm) appeared at a random location within the blue bounding box. The participant then quickly reached out and touched somewhere on a touch-screen monitor. The participants had 1000 ms to reach out and touch the screen.

Participants gained and lost money based on where and when they completed their reaching movements. If a participant touched within the boundary of the green "reward" circle before the time limit (1000 ms), they earned a reward. If they touched within the boundary of the penalty circle before the time limit, they incurred a penalty. If they touched within the boundaries of both circles before the time limit, they incurred the sum of the penalty and reward.

A touch outside both the penalty and reward circles resulted in neither a penalty nor a reward. If a participant completed their reach after the trial time limit had elapsed, they incurred a large penalty (-$5) for that trial and were shown feedback stating, “Too slow”. If a participant released the space bar before the reward and penalty circles appeared, the trial was aborted, and they were shown feedback stating, “Too soon”.

The experiment consisted of training and test sessions completed on two separate days and lasting approximately one hour each. The training session was divided into three parts. In Part 1, which consisted of 12 blocks of 36 trials, participants reached at a single unchanging yellow dot. This provided a distribution of end-points from which to estimate each participant’s reaching end-point variability. Part 2 consisted of 6 blocks of 36 trials. During Part 2, participants were exposed to six conditions: 3 payoff conditions ([-$0, +$1]; [-$1, +$1]; [-$5, +$1]) with one of two separation distances between the penalty and reward regions (1R; 1.4R, where R = 8.55 mm). Participants were informed that performance in training did not affect their total earnings in the task and was just for practice. Conditions varied randomly on a trial-by-trial basis. Part 3 of training consisted of 2 additional blocks of 36 trials each in which two additional payoff conditions were added: [-$3, +$3] and [-$15, +$5]. These ‘high-stakes trials’ were essential to the comparison between the Loss-averse Model and the Ideal Model as the former would predict more drastic shifts away from the penalty region on these trials.

On Day 2, before the test session, participants took a “pre-test” to ensure they understood the task instructions. This pre-test presented a series of hypothetical reaching end-points for every possible payoff condition and asked the participant how much would be earned for each item. In order to advance to the test phase of the experiment, participants needed to complete the

pre-test with 100% accuracy. If participants did not understand an item, we explained it to them. Only one participant failed to complete the test with 100% accuracy on their first attempt.

In the test session itself, participants performed 10 blocks of 60 trials each, with all 10 conditions varying randomly on a trial-by-trial basis. At the end of the test session, one trial was selected at random from each block. Participants gained or lost bonus earnings based on their performance on the randomly selected trials. After the reaching task was complete, participants performed the loss aversion task described below.

Pre-processing. We analyzed our data, ran simulations, and visualized our results using several packages in R (R Core Team, 2013). As in prior work, we pre-processed our raw data for the ease of data analyses, to correct for small reaching biases, and to remove temporal/spatial outliers (see Supplementary Methods).

We defined a participant's movement variability as the standard deviation of their horizontal reaching end-points in during the first part of training. As in previous work using this task, we defined a participant's aim-point within each condition as the mean of their reach end-points for that condition. We assume for simplicity that each participant has a single aim-point for a given condition and that all deviations of end-points from this aim-point are a result of the participant's movement variability.

The ideal model. The first model of behavior in this task—which we will call the Ideal Model—assumes that participants use an optimal strategy whereby they select aim-points with the highest relative EV, conditional on their own movement variability. This model is identical to the MEGaMove model commonly used when modeling this task[1,2]. This model can be represented as the claim that participants select aim-points (x, y) which maximize EV via the equation (1) below, where P_S is the penalty size, $P(P_C \mid x,y)$ is the probability of a reaching outcome falling

within the penalty circle, R_S is the reward size, $P(R_C \mid x,y)$ is the probability of a reaching outcome falling within the reward circle:

$$EV(x,y) = P_S * \mathrm{P}(P_C|x,y) + R_S * \mathrm{P}(R_C|x,y). \quad (1)$$

The probability of touching a location on the screen (x', y') given some aim-point (x,y) is determined by the participant's movement variability (see Supplementary Methods for more complete details).

In sum, this model provides a prediction of where a participant will aim in each payoff condition if they were to maximize objective EV. (See the Supplementary Methods for a description of how we numerically derived the predictions of the Ideal and Loss-averse Models.). If the Ideal Model has the best fit to our participants' behavior, then it is more likely that our participants used the optimal ideal strategy than the loss-averse or heuristic strategies.

The Loss-averse Model. The second computational model that we considered—which we will call the Loss-averse Model— assumes participants' EV estimations will be distorted by their loss-averse representations of the penalty values. This model is identical to the first model, except for one change: the penalty size, P_S, in equation (1) is multiplied by the participant's trait-level loss aversion, as approximated the loss aversion task completed in a separate session described below. The Loss-averse Model predicts that participants select aim-points for each condition which maximize the following equation:

$$EV'(x,y) = \lambda * P_S * \mathrm{P}(P_C|x,y) + R_S * \mathrm{P}(R_C|x,y). \quad (2)$$

The lambda coefficient (if greater than 1) effectively increases the size of the penalty, and thereby leads the Loss-averse Model to predict aim-points that are further away from the penalty region than the aim-points predicted by the Ideal Model. If the Loss-averse Model has the best fit

to our participants' data, then it is more likely that our participants used a loss-averse strategy than the ideal or heuristic strategies.

The Heuristic Model. The last computational model we test—which we will call the Heuristic Model—assumes that participants in this task *simply choose to aim at a single satisfactory point when penalties are preset.* This model is insensitive to the exact penalty and reward values displayed or a participant's movement variability. According to this model (illustrated in Figure 3), participants will perform the task using the following heuristic, where P is the penalty value, (x, y) is the participant's aim-point, (0, 0) is the center of the target circle, and φ is the horizontal midpoint between the right edges of the penalty and target circles at y = 0:

$$\text{If } Penalty == 0: \quad (x, y) = (0, 0), \quad \text{else:} \quad (x, y) = (\varphi, 0). \tag{5}$$

According to this model, the participant will aim at the center of the reward circle if there is no penalty. If the penalty is greater than zero, the participant will aim midway between the outside edge of the reward circle and the edge of the penalty circle (See Figure 3 for an illustration of the Heuristic Model's predictions). Since the Heuristic Model does not incorporate exact penalty and reward values or the participant's movement variability, it gives the same predictions for the different payoff conditions, as well as the same predictions for all participants. If the Heuristic Model has the best fit to our participants' behavior, then it is more likely that our participants used this sub-optimal heuristic strategy than the ideal or loss-averse strategies. Figure 3 shows the predictions of the three models for a representative participant in Experiment 1.

Model comparison. We defined the fit of a model to a set of reaching end-points as the summed log-likelihoods of the data, given the model's predictions (see Supplementary Methods). We tested for significant differences of fit between models using Vuong's test for comparing non-nested models[19]. All statistically significant differences were Bonferroni corrected for multiple comparisons.

Additional analyses. In line with prior work, we examined how similar participants' earnings in the test session were to the expected earnings predicted by each model. To test this, we computed, for each participant and model, the ratio of the participant's actual earnings to what would be expected if the participant followed the strategy entailed by the model (i.e., for a given subject, expected earnings equal to the expected-value of the model's aim-point for a condition times the number of completed trials for that condition, summed over all conditions). In prior work, participants earned nearly 100% of what was predicted by the Ideal Model and this has been cited as evidence for optimal performance in this task[1,2,3,20].

Additionally, prior work has argued that, in this reaching task, participants are sensitive to the penalty values, reward values, and their own movement variability when making decisions about where to reach[3,4]. If our participants were sensitive to penalty and reward values, then the distance between their aim-point and the penalty region should increase as the penalty to reward ratio grows. We used paired t-tests to check for differences of mean end-points between the three penalties to reward ratio conditions: 0:1, -1:1, and -5:1. All differences reported as significant have p-values that were Bonferroni corrected for multiple comparisons.

Additionally, if our participants were sensitive to their own movement variability, their aim-point (relative to the penalty region) should be proportional to their movement variability.

We used a spearman's rank-order correlations to test for a positive relationship between movement variability and mean end-point across all non-zero penalty conditions.

Loss aversion task. After the participant finished the last block of the test session, they completed a gambling task which measured the extent to which they weighted losses more than objectively equivalent gains—that is, their loss aversion[10,11]. A behavioral loss aversion coefficient (λ) was computed via logistic regression (see Supplementary Methods for complete task and analysis details).

Eight participants (four in each experiment) had performance that resulted in an implausible λ value that was less than one (loss-seeking) or greater than 20. These participants appeared to be following a simple decision-making rule rather than evaluating each gamble independently. For these participants, we let $\lambda = 1$ when fitting the Loss-averse Model, implying that they are not loss-averse. For these participants, the loss-averse and ideal models make identical predictions about where they will aim their reaching movements.

Experiment 2

Participants. 16 students (12 female) from the University of Michigan. All participants gave written consent to participate in the study. The University of Michigan Institutional Review Board of Health Sciences and Behavioral Sciences approved all procedures used in this experiment. All participants were right-handed and had normal or corrected-to-normal vision. At the end of the experiment, participants were compensated at a rate of $10/hour plus performance bonuses and their outcome from the loss aversion task described below, if applicable. A participant's base hourly compensation was not affected by their performance in any task.

Reaching task. Participants performed the same incentivized reaching task described in Experiment 1. All task parameters were identical except those described below. Experiment 2 similarly consisted of training and test sessions completed on two separate days and lasting approximately one hour each. The training session was composed of 10 blocks of 30 trials each. In training, participants always saw the overlapping reward and penalty regions, but the penalty value was zero. The time-limit for each reaching movement decreased from unlimited in block 1 to 850 ms in blocks 2 – 4 to 700 ms in blocks 5 – 10. The only penalties in the training session were due to reaches completed after the time-limit. The separation distance between the reward and penalty circles was fixed at 1R (8.5 mm).

The test session was composed of 12 blocks of 30 trials each. In the test session, participants accrued points throughout the experiment that were translated into cash earnings at a rate of 25 cents per 1000 points at the conclusion of the experiment. The reward value was fixed at +100 points and the penalty value was fixed within each block at either zero, -100, or -500 points (4 blocks each, in random order). The time-limit was 700 ms and the penalty for completing a movement after the time-limit was -700 points. The separation distance between the reward and penalty circles was again fixed at 1R (8.5 mm).

Data analyses and modeling. All analyses and models were identical to those used in Experiment 1.

References

1. Trommershäuser, J., Maloney, L. T., & Landy, M. S. Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision* **16,** 255–275 (2003).
2. Trommershäuser, J., Maloney, L. T., & Landy, M. S. Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. A.* **20,** 1419–1433 (2003).
3. Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. Optimal compensation for changes in task-relevant movement variability. *J. Neurosci.* **25,** 7169–78 (2005).
4. Trommershäuser, J., Landy, M. S., & Maloney, L. T. Humans rapidly estimate expected gain in movement planning. *Psych. Science* **17,** 981–988 (2006).
5. Trommershäuser, J., Maloney, L. T., & Landy, M. S. Decision making, movement planning and statistical decision theory. *Trends in Cognitive Sciences*, **12,** 291–297 (2008).
6. Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases in *Utility, probability, and human decision making* 141-162 (Springer, 1975)
7. Tversky, A., & Kahneman, D. The framing of decisions and the psychology of choice in *Environmental Impact assessment, technology assessment, and risk analysis* 107-129 (Springer, 1985).
8. Hutchinson, J. M. C., & Gigerenzer, G. Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, **69,** 97-124 (2005).

9. Gigerenzer, G., & Gaissmaier, W. Heuristic decision making. *Annual Review of Psychology*, **62,** 451–482 (2011).
10. Tom, S. M., Fox, C.R., Trepel, C., Poldrack, R.A. The neural basis of loss aversion in decision-making under risk, *Science,* **315,** 515-518 (2007).
11. Kahneman, D., & Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica*, **47,** 263–291 (1979).
12. Shah, A. K., & Oppenheimer, D. M. Heuristics made easy: an effort-reduction framework. *Psych. Bulletin,* **134,** 207-222 (2008)
13. Simon, H. A. A behavioral model of rational choice. *Quarterly J. Econ.* **69,** 99–118 (1955).
14. Howes, A., Warren, P.A., Farmer, G., El-Deredy, W., Lewis, R.L. Why contextual preference reversals maximize expected value. *Psych. Rev.* **123,** 368-391 (2016).
15. Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. Large stakes and big mistakes. *Rev. of Econ. Studies*, **76,** 451–469 (2009).
16. Chib, V.S., De Martino, B., Shimojo, S., O'Doherty, J.P. Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron*, **74,** 582-594 (2012).
17. Lee, T. G., & Grafton, S. T. Out of Control: Diminished prefrontal activity coincides with impaired motor performance due to choking under pressure. *Neuroimage*, **105,** 154-155 (2015).
18. Mathôt, S., Schreij, D., & Theeuwes, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods*, **44,** 314–324 (2012).

19. Vuong, Q. H. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, **57,** 307-333 (1989).

20. Jarvstad, A., Hahn, U., Warren, P.A., Rushton, S.K. Are perceptuo-motor decisions really more optimal than cognitive decisions? *Cognition*, **130,** 397-416 (2013).

21. Iyengar, S. S., & Lepper, M. R. When choice is demotivating: Can one desire too much of a good thing? *J. Person. and Soc. Psych.*, ***79,*** 995–1006 (2000).

22. Bach, D. R., & Dolan, R. J. Knowing how much you don't know: a neural organization of uncertainty estimates. *Nature Reviews Neuroscience*, **13,** 572-586 (2012).

23. Landy, M. S., Trommershäuser, J., & Daw, N. D. Dynamic estimation of task-relevant variance in movement under risk. *J. Neurosci.*, **32,** 12702–12711 (2012).

24. Orbán, G., & Wolpert, D. M. Representations of uncertainty in sensorimotor control. *Curr. Op. in Neurobio.*, **21,** 629–635 (2011).

25. Zhang, H., Daw, N. D., & Maloney, L. T. Testing whether humans have an accurate model of their own motor uncertainty in a speeded reaching task. *PLoS Comp. Bio.*, 9, https://doi.org/10.1371/journal.pcbi.1003080 (2013).