



The composition of the choice set modulates probability weighting in risky decisions

M. A. Grubb^{1,2} · Y. Li^{2,3} · R. Larisch² · J. Hartmann⁴ · J. Gottlieb^{5,6,7} · I. Levy²

Accepted: 3 January 2023 / Published online: 26 January 2023
© The Psychonomic Society, Inc. 2023

Abstract

Probability distortion—the tendency to underweight larger probabilities and overweight smaller ones—is a robust empirical phenomenon and an important driver of suboptimal choices. We reveal a novel contextual effect on probability distortion that depends on the composition of the choice set. Probability distortion was larger in a *magnitude-diverse* choice set (in which participants encountered more unique magnitudes than probabilities) but declined, resulting in more veridical weighting, in a *probability-diverse* choice set (more unique probabilities than magnitudes). This effect was consistent in two, large, independent datasets (N = 481, N = 100) and held for a subset of lotteries that were identical in the two contexts. It also developed gradually as a function of exposure to the choice set, was independent of attentional biases to probability versus magnitude information, and was specific to probability weighting, leaving risk attitudes unaffected. The results highlight the importance of context when processing probabilistic information.

Keywords Decision-making · Uncertainty · Attention

Introduction

Humans routinely face consequential decisions involving uncertain outcomes with known probabilities, that is, decisions involving risk. Monetary lotteries have been extensively used to study risky decision making in the laboratory (Holt & Laury, 2002), and empirical choice data have provided indispensable insights into human preferences for uncertainty. In

a common experimental setup, participants are faced with a series of choice scenarios that require a decision between two options, each of which offers a chance (the probability) of some reward (the magnitude). Lottery probabilities and magnitudes typically vary from trial to trial so that participants encounter diverse amounts of risk and rewards. Choice data can then be used to estimate parameters in theoretical models of decision making, with which theoretical predictions can be tested and refined. Converging evidence from empirical and theoretical work has shown that, when making decisions under risk for explicitly stated outcomes with precise probabilities (i.e., decisions from description; Hertwig & Erev, 2009), humans tend to exhibit risk aversion (i.e., preferring smaller, safer rewards to larger, riskier ones; Holt & Laury, 2002; but see also Barron & Erev, 2003 for feedback as a potential modulating factor) and probability distortion (i.e., the tendency to underweight larger probabilities and overweight smaller ones; Kahneman & Tversky, 1979).

Investigations into choice in the absence of risk have revealed that participants evaluate available options differently depending on the choice context—specifically, the composition of the choice set. One contextual influence is range adaptation, whereby sensitivity to value varies inversely with the range of the values available in the local context (Padoa-Schioppa, 2009). Another, potentially

✉ M. A. Grubb
michael.grubb@trincoll.edu

¹ Department of Psychology, Trinity College, Hartford, CT, USA

² Comparative Medicine, Yale School of Medicine, New Haven, CT, USA

³ Department of Psychological & Brain Sciences, Boston University, MA, Boston, USA

⁴ Columbia University, New York, NY, USA

⁵ Department of Neuroscience, Columbia University, New York, NY, USA

⁶ Kavli Institute for Brain Science, Columbia University, New York, NY, USA

⁷ Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA

related, contextual influence in trinary, risk-less choice involves normalized value coding in which the value of irrelevant options can impact decision making (Khaw et al., 2017; but see also, Gluth et al., 2020). These contextual influences have been well investigated and may play critical roles in allowing the brain to efficiently allocate limited coding resources to the immediately relevant value range.

However, an additional contextual regularity that often arises in risky choice but has yet to be systematically studied concerns the frequency of encountering a given magnitude or probability. In a typical crossed experimental design, such as the one described in the opening paragraph, participants encounter several probabilities and magnitudes presented in all possible combinations, so that the frequency with which the participant encounters each value depends on the number of unique magnitudes and probabilities in the choice set. For example, in a study that crosses 3 probabilities (e.g., 0.25, 0.5, and 0.75) with 10 magnitudes (e.g., \$5–\$50 in \$5 increments), participants confront each individual probability much more frequently than each individual magnitude (specifically, in a 10:3 ratio). Ample evidence suggests that human observers can implicitly learn contextual regularities while performing a task (Sherman et al., 2020) and use these regularities to modulate expectations (Summerfield & de Lange, 2014). Moreover, variability influences attention (Jiang, 2018) which, in turn, can influence choice (Fiedler & Glockner, 2012; Glickman et al., 2019; Harrison & Swarthout, 2019; Hunt et al., 2018; Johnson & Busemeyer, 2016; Kim et al., 2012; Smith & Krajbich, 2018; Spitmaan et al., 2019; Stewart et al., 2016). Thus, contexts that differ in the numbers of unique probabilities versus magnitudes they present may induce differential attention to probability versus magnitude information, which may influence choice. However, it is unknown whether or how the relative frequencies of unique magnitudes and probabilities in the choice set influence decision making under uncertainty.

We examined this question by comparing choice behavior in two distinct contexts: a *magnitude-diverse* context in which participants were exposed, across trials, to more unique magnitudes than unique probabilities (respectively, 21 and 5), and a *probability-diverse* context in which participants were exposed to more unique probabilities than unique magnitudes (the inverse ratio). A subset of identical choice trials, the *Overlapping Choice Set*, appeared in both contexts, allowing for direct between-group comparisons. We found robust evidence in two independent samples that the probability-magnitude asymmetry selectively modulated probability distortion with no effect on risk attitudes. Participants showed lower probability distortion (i.e., their probability weighting was closer to optimal) in the *probability-diverse* context relative to the *magnitude-diverse* context, and this difference emerged gradually while participants were performing the task. Moreover, the contextual effect

on probability weighting was independent of overt attention to choice-relevant information, suggesting that this contextual manipulation affects the subjective evaluation of probabilities for decision formation independently of attention.

Methods

Participants A total of 581 adults were tested in two studies. Study 1 was conducted on Amazon Mechanical Turk (“web-based,” N = 481), and Study 2 was conducted in person at Yale University (“lab-based,” N = 100). All participants provided informed consent. Experimental procedures were approved by the institutional review boards at Columbia University (web-based study) and Yale School of Medicine (lab-based study), and all methods were performed in accordance with the relevant guidelines and regulations. Web-based study demographics: mean age: 34.08 years (standard deviation [SD] = 9.68); gender: female, 199; male, 234; other, 2; not reported or missing, 46. Lab-based study demographics: mean age: 26.6 years old; gender: female, 60; male, 40. For additional details on participant demographics, see Supplementary Material, *Additional Demographic Information*.

Data exclusions In the web-based study, an additional 96 people participated but were excluded due to excessive first-order stochastic dominance violations, choosing a chance of 5 points over a guaranteed 5 points on half or more of the trials; this exclusion criterion was set *a priori* and has been used in our previous work (Grubb et al., 2016). Ten such people were excluded from the lab-based study. In the web-based study, an additional 44 datasets were collected but not included in the analysis for the following reasons: 18 datasets contained trial data for more than the programmed number of trials, which indicates a writing error or potential modification by the user; 18 participants completed the study more than once (as judged by the mTurk ID), and only the first session was included in the analysis; 8 datasets had no mTurk ID, and we were thus unable to verify that the data was not from a repeat session.

Experimental task In each study, participants made 210 decisions between a fixed, certain gain (web-based: 5 points; lab-based: \$5) and a lottery whose magnitude and probability of payout varied from trial to trial (Fig. 1). To examine the relationship between choices and inspection of the lottery attributes before the choice, we included icons to the left and right of the fixation point indicating locations from which participants could obtain information about magnitude or probability (Fig. 1, “Feature location information”). In the web-based study, participants pressed a button to indicate which information

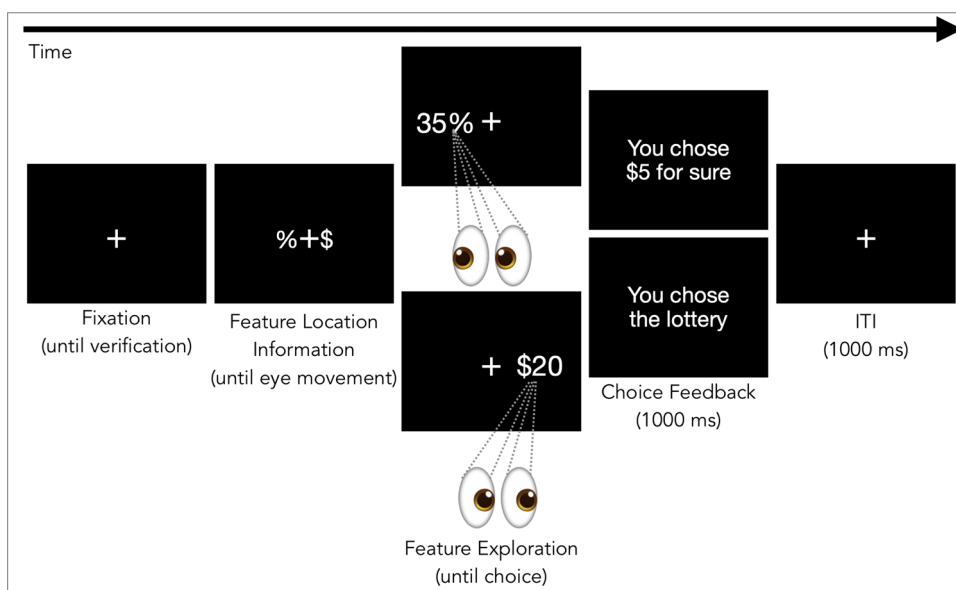


Fig. 1 Example trial sequence from the lab-based study for a lottery offering a 35% chance of \$20. \$ and % icons at fixation indicated where magnitude and probability information for that trial’s lottery could be found; visual fixations at those locations were required to

reveal and keep visible the information. In the web-based study, participants were required to press and hold the left or right arrow key to reveal magnitude and probability information

they wished to reveal. In the lab-based study, we used eye-tracking and provided lottery feature information in a gaze-contingent manner if a participant maintained fixation on the side corresponding to the icon in the display.

Experimental manipulation Participants were randomly assigned to one of two choice sets: a *magnitude-diverse* choice context, in which the variable lottery could be comprised of 1 of 5 unique probabilities but 1 of 21 unique magnitudes, and a *probability-diverse* choice context, in which the variable lottery could be comprised of 1 of 21 unique probabilities but only 1 of 5 unique magnitudes. The range of magnitudes and probabilities was identical in both choice sets (respectively, magnitudes of 5–80 points/dollars and probabilities of 0.2–0.8; see below for full choice sets). Each lottery was presented twice (with magnitude and probability information spatially counterbalanced), and an identical subset of 25 lotteries, the *Overlapping Choice Set*, was present in both choice contexts. Values that constitute the *Overlapping Choice Set* are shown in bold.

Choice model To estimate risk attitudes and assess probability weighting, we fit a subjective utility model and a logistic function to the choice data. We modeled subjective utility (*SU*) using the functional form:

$$SU(p, m) = e^{-(-lnp)^\gamma} m^\alpha \tag{1}$$

where *p* and *m* denote, respectively, the probability and magnitude offered by the lottery; α controls the curvature of the utility function, and γ indicates the degree of distortion in the probability weighting function (Prelec, 1998). We used a maximum likelihood fitting procedure to fit the probability of choosing the lottery on each trial using a logistic choice function:

$$P_{lotto} = \frac{1}{1 + e^{-(SU_{lotto} - SU_{ref})/\sigma}} \tag{2}$$

For additional details on parameter recovery, parameter identifiability, and the issue of interdependencies, see Supplementary Material, *Parameter Recovery and Identifiability*.

Magnitude-diverse choice context

%	20	35	50	65	80																
*/\$	5	6	7	8	9	10	12	14	16	18	20	23	26	30	35	40	48	56	64	72	80

Probability-diverse choice context

%	20	23	26	29	32	35	38	41	44	47	50	53	56	59	62	65	68	71	74	77	80
*/\$	5	10	20	40	80																

Payment At the end of the experiment, one trial was randomly selected, and its outcome was realized for a real bonus payment. In the web-based study, the magnitudes referred to points, with *5 being equivalent to \$0.05 (explained in advance); in the lab-based study, the magnitudes indicated amounts in U.S. dollars. In addition to this bonus payment, participants received a flat fee of \$1 (web-based study) or \$10 (lab-based study) for participating.

Eye-tracking (lab-based study) Eye position was monitored using a Tobii Eye-Tracking system recording at 60Hz, and a 9-point calibration routine was used to calibrate the eye tracker for each participant. During the Feature Exploration phase of each trial (Fig. 1), real-time eye position was continuously obtained. Magnitude and probability information was presented at ± 480 pixels from the center of the screen and was only, and selectively, revealed when horizontal eye position was greater than 300 or less than -300 pixels from the center of the screen.

Feature-based attention index (FBAI) During a given trial, participants could overtly attend to magnitude information, probability information, or neither. Overt attention to neither feature presumably reflects some degree of internal attention to the nonvisible, static reference (\$5 for sure), or to previously overtly attended information (e.g., mental consideration of the magnitude and probability of the lottery currently on offer). The proportion of time spent revealing either feature, through a button press (web-based study) or an eye movement (lab-based study), was calculated for each trial, and the mean across trials was computed for each participant.

In both studies, participants who took longer to make a decision spent more time viewing magnitude and probability information (Spearman correlation: web-based, $r = 0.3927$, $p < 0.0001$; lab-based, $r = 0.7745$, $p < 0.0001$), and the time spent viewing magnitude and probability information were highly correlated (Spearman correlation: web-based, $r = 0.8662$, $p < 0.0001$; lab-based, $r = 0.7799$, $p < 0.0001$). To control for these correlations, we normalized attention to each feature by the trial's response time and calculated for each participant a *Feature Based Attention Index* (FBAI) as the average difference in the proportion of each trial (PT) spent attending to the two features:

$$FBAI = \frac{PT_{mag} - PT_{prob}}{PT_{mag} + PT_{prob}} \quad (3)$$

Thus, an FBAI score of zero indicates an equivalent amount of attention to magnitude and probability information, increasingly positive FBAI scores indicate an increasingly greater attentional bias toward magnitude information, and increasingly negative FBAI scores indicate an increasingly greater attentional bias toward probability information, all after controlling for RT and total viewing time.

Bootstrapped confidence intervals To create bootstrapped confidence intervals (CI), we randomly sampled, with replacement, 226 or 255 observers in the web-based study (51 or 49 observers, lab-based study) from the *magnitude-diverse* and *probability-diverse* choice context groups, respectively. We then computed the median of each new, random sample and repeated the process 10,000 times. The inner 95% of the resulting distribution was extracted to determine the bounds of the CI. This process was conducted separately for each group and for each study.

Randomization tests on group medians To evaluate between-group differences, we conducted the following randomization procedure: 1) group labels were randomly shuffled, 2) group medians were computed and the difference was recorded after each shuffle, 3) this process was repeated 10,000 times to generate a null distribution, 4) the p -value of the randomization test was designated as the proportion of the null distribution greater than or equal to the actual, empirically observed group difference. Absolute values were used to make these two-tailed tests.

Results

In this experiment, we measured the impact of probability-magnitude asymmetries in the choice set on risk tolerance and probability weighting. Participants were randomly assigned to either a *magnitude-diverse* choice context in which they were exposed across trials to more unique magnitudes than unique probabilities, in a 21:5 ratio, or to a *probability-diverse* context that contained the converse. An identical subset of 25 lotteries, the *Overlapping Choice Set* (OCS), was present in both choice contexts. On each trial, participants chose between a safe, certain option (\$5 for sure) and a risky lottery. To isolate and measure the amount of goal-directed attention to each distinct lottery feature, we took inspiration from process-tracing approaches that rely on mouse position to selectively reveal information (Johnson et al., 1989). We presented magnitude and probability on opposite sides of the screen (counterbalanced) and used button-press-contingent (Study 1) and gaze-contingent presentation (Study 2) to ensure that only one lottery feature was overtly attended at a time (Fig. 1).

Participants' choices evidenced lawful behavior, alleviating potential concerns about the spatial separation of lottery information. To facilitate a direct, between-group comparison of model-free data, choices from the *Overlapping Choice Set* trials are shown in Fig. 2. The proportion of lottery choices increased as the magnitude (two left panels) and probability (two right panels) of winning the lottery increased, confirming sensitivity to these experimental features. The proportion of lottery choices, in each

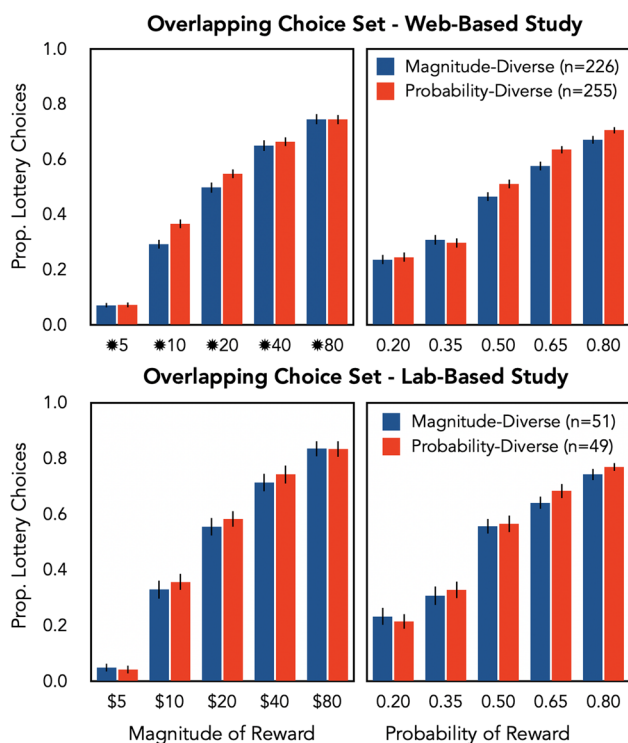


Fig. 2 Proportion of lottery choices confirms lawful behavior. Preference for the lottery increases as magnitude and probability of reward increases

experiment and in each group, was substantially less than what an expected value framework predicts (0.64 for the *Overlapping Choice Set* lotteries), confirming the presence of risk aversion with this model-free metric of risk tolerance. Web-based, *magnitude-diverse* choice context: 0.44 (median), [0.42, 0.49] (bootstrapped 95% CI); web-based, *probability-diverse* choice context: 0.48, [0.44, 0.50]; lab-based, *magnitude-diverse* choice context: 0.46, [0.44, 0.52]; lab-based, *probability-diverse* choice context: 0.54, [0.48, 0.58].

In reporting the results of the remaining analyses, we first present the outcome when all trials in each choice context are considered, which maximizes statistical power. We then present the outcome when only trials from the *Overlapping Choice Set* are considered, which rules out the possibility that our results are an artifact of the precise magnitude-probability pairs used in the analysis.

Experimentally manipulating the choice set had a consistent impact on probability weighting, without affecting other parameters of the choice function. We fit the *SU* model and logistic function to individual choice data from all trials and compared the resultant distributions of parameters independently in each study (Fig. 3A, Table 1, Table 2). For participants in the *magnitude-diverse* choice context, median γ estimates and their corresponding 95% confidence

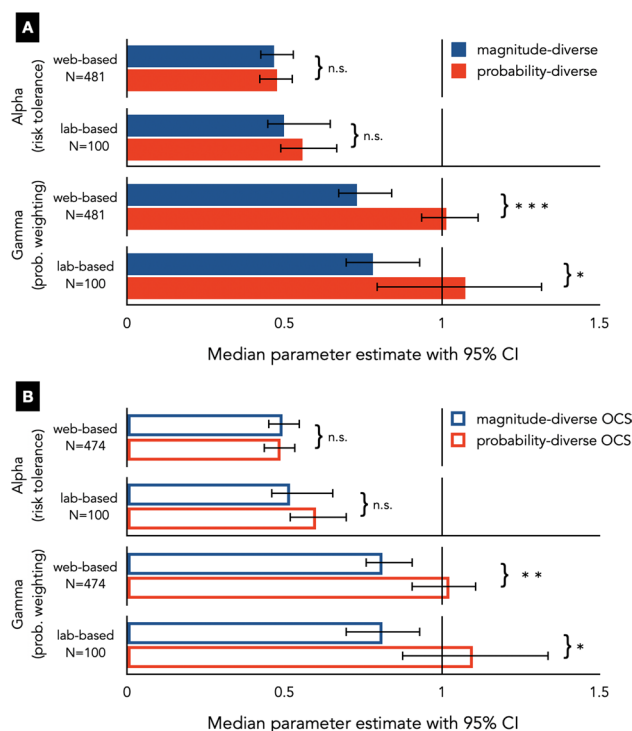


Fig. 3 Median parameter estimates from both studies. **A**) All trials. Error bars, bootstrapped 95% confidence intervals. *** $p < 0.0001$; * $p < 0.05$; n.s., not significantly different. **B**) Trials from the overlapping choice set (OCS), ** $p < 0.005$. Seven participants in the web-based study are excluded because the MLE fitting procedure failed to converge with the limited number of trials

intervals were less than one—indicating a nonlinear probability weighting function. In contrast, median γ estimates for participants in the *probability-diverse* choice context were near one—indicating a linear probability weighting function. In both studies, γ estimates differed significantly between groups (web-based: $p < 0.0001$; lab-based: $p = 0.0134$), but the degree of risk aversion was statistically indistinguishable between groups (web-based: $p = 0.8526$; lab-based: $p = 0.2995$), as were estimates of σ , the slope of the logistic choice function (web-based: $p = 0.4887$; lab-based: $p = 0.5020$).

Crucially, the effect of the global choice set on probability distortion held when we restricted our analysis to trials in the *Overlapping Choice Set* (23.8% of the data; Fig. 3B). When we examined this subset of choice scenarios that were identical in each group, we again found that in both studies, γ estimates differed significantly between groups (web-based: $p = 0.0021$; lab-based: $p = 0.0390$), with statistically indistinguishable estimates of risk tolerance (web-based: $p = 0.8000$; lab-based: $p = 0.2263$) and the slope of the logistic choice function (web-based: $p = 0.4793$; lab-based: $p = 0.5718$). Given that the lotteries in the *Overlapping Choice Set* were identical, and given

Table 1 Web-based study with key tracing (N = 481)

	Alpha (risk preference)		Gamma (probability weighting)	
	Magnitude-Diverse	Probability-Diverse	Magnitude-Diverse	Probability-Diverse
Median estimate	0.4676	0.4760	0.7302	1.0146
Bootstrapped 95% CI	[0.4220, 0.5297]	[0.4188, 0.5266]	[0.6699, 0.8431]	[0.9336, 1.1165]
CI includes 1?	N	N	N	Y
Between-group randomization test	$p = 0.8526$		$p < 0.0001$	

Median parameter estimates and statistical results from individual fits. Significant group differences are shown in bold

Table 2 Lab-based study with eye-tracking (N = 100)

	Alpha (risk preference)		Gamma (probability weighting)	
	Magnitude-Diverse	Probability-Diverse	Magnitude-Diverse	Probability-Diverse
Median estimate	0.4984	0.5571	0.7813	1.0745
Bootstrapped 95% CI	[0.4442, 0.6481]	[0.4865, 0.6676]	[0.6929, 0.9312]	[0.7921, 1.3187]
CI includes 1?	N	N	N	Y
Between-group randomization test	$p = 0.2995$		$p = 0.0134$	

Median parameter estimates and statistical results from individual fits. Significant group differences are shown in bold

that group assignment was random, the only thing that can account for this between-group difference in probability weighting is the wider context in which these choices were embedded (i.e., the global choice set).

Converging evidence that the choice set exerts an impact on probability distortion, but not on risk tolerance or choice stochasticity, was obtained using a different modeling approach. We simultaneously fit the choice model to all participants' data, clustering the standard errors (SE) on participants and allowing each parameter to vary as a linear function of the choice context and experimental setting (Gilaie-Dotan et al., 2014; Grubb et al., 2016; Harrison, 2008; Harrison & Rutstrom, 2008). This produced coefficients indicating the extent to which risk tolerance (α), distortions in the probability weighting function (γ), and the slope of the logistic choice function (σ) depended on the composition of the choice context and on the experimental setting (Model 1: $\alpha = \beta_0 + \beta_1 \times isMagnitudeDiverse + \beta_2 \times isLabBased$; $\gamma = \beta_0 + \beta_1 \times isMagnitudeDiverse + \beta_2 \times isLabBased$; $\sigma = \beta_0 + \beta_1 \times isMagnitudeDiverse + \beta_2 \times isLabBased$, where *isMagnitudeDiverse* and *isLabBased* are dummy predictors representing, respectively, the choice context and experimental setting). Consistent with the results presented above, we observed a significant negative coefficient for the *isMagnitudeDiverse* predictor in Model 1 (z-test: $n = 122,010$, SE clustered on 581 participants, $z = -3.25$, $p = 0.001$),

indicating greater distortion in the probability weighting function (i.e., more probability weighting) for participants in the *magnitude-diverse* choice environment, compared with those in the *probability-diverse* choice environment (Table 3).¹ The *isMagnitudeDiverse* coefficients for α and σ were not significant ($ps < 0.373$), indicating that the choice environment manipulation specifically affected probability weighting. Repeating this modeling approach for *Overlapping Choice Set* trials produced consistent results: we again found a significant, negative coefficient for our dummy predictor *isMagnitudeDiverse* on γ only (z-test: $n = 29,050$, SE clustered on 581 participants, $z = -2.98$, $p = 0.003$; Table 4).

If our experimentally induced manipulation of probability distortion relies on exposure to the choice environment, we might expect the effect to grow as the experiment progresses. To assess this possibility, we fit the *SU* model to trial-number-based deciles of choice data from all participants in each choice context (10 bins of 21 trials each) and clustered the standard

¹ While it is true that participants in the lab-based study were found to be significantly more risk tolerant than those in the web-based study, we are agnostic as to whether or not this result is causally related to study location or some other methodological difference between the two studies (e.g., differences in the payout scheme). Unlike the choice set assignment, we did not randomly assign participants to participate in the lab or via the web.

Table 3 Web- and lab-based studies combined (N = 581), all trials

Model 1	Coef.	Robust SE	z	p	95% CI	
Risk tolerance (α)						
isMagnitudeDiverse	0.0164	0.0184	0.89	0.373	-0.0197	0.0525
isLabBased	0.0551	0.0220	2.51	0.012	0.0120	0.0982
constant	0.4497	0.0133	33.74	0.000	0.4236	0.4759
Probability weighting (γ)						
isMagnitudeDiverse	-0.1639	0.0505	-3.25	0.001	-0.2628	-0.0650
isLabBased	0.1124	0.0582	1.93	0.053	-0.0016	0.2265
constant	0.8350	0.0403	20.72	0.000	0.7560	0.9140
Logistic slope (σ)						
isMagnitudeDiverse	-0.0010	0.0539	-0.02	0.986	-0.1065	0.1046
isLabBased	-0.0971	0.0547	-1.77	0.076	-0.2044	0.0101
constant	0.7202	0.0401	17.96	0.000	0.6416	0.7988

Parameter estimates and statistical results. Robust SE, standard error clustered on participant. CI, confidence interval. Coefficients significantly different from zero in bold

Table 4 Web- and lab-based studies combined (N = 581), Overlapping Choice Set

Model 1 - OCS	Coef.	Robust SE	Z	p	95% CI	
Risk tolerance (α)						
isMagnitudeDiverse	-0.0103	0.0179	-0.58	0.565	-0.0455	0.0249
isLabBased	0.0600	0.0213	2.81	0.005	0.0181	0.1018
constant	0.4647	0.0135	34.31	0.000	0.4382	0.4913
Probability weighting (γ)						
isMagnitudeDiverse	-0.1458	0.0490	-2.98	0.003	-0.2418	-0.0498
isLabBased	0.1228	0.0563	2.18	0.029	0.0125	0.2330
constant	0.7757	0.0393	19.74	0.000	0.6987	0.8527
Logistic slope (σ)						
isMagnitudeDiverse	-0.0587	0.0501	-1.17	0.241	-0.1568	0.0394
isLabBased	-0.0949	0.0529	-1.80	0.073	-0.1985	0.0087
constant	0.7360	0.0397	18.56	0.000	0.6583	0.8138

Parameter estimates and statistical results. Robust SE, standard error clustered on participant. CI, confidence interval. Coefficients significantly different from zero in bold

error at the participant level. Figure 4 shows the resultant parameter estimates for each decile and for each choice context. Confidence intervals indicate that the two groups exhibited minimal, and highly similar, probability weighting at the start of the experiment, but by the 85th trial, the two groups reliably diverged. Risk tolerance appears to be relatively stable throughout the duration of the experiment and, consistent with previous results, does not differ between groups. Given the limited number of choices in the *Overlapping Choice Set*, this analysis focuses only on the trials from the full choice set.

Our experimental paradigm also was designed to measure overt attention to magnitude and probability information,

allowing us to test whether differences in attention can explain differences in choice. Because our contexts differed in the relative numbers of unique values of magnitude versus probability, the most relevant attention metric is the extent to which participants are biased to attend to one or the other dimension. Thus, we devised a feature-based attentional index (FBAI; Methods) such that FBAI scores greater than zero indicate an attentional preference for magnitude information, whereas FBAI scores less than zero indicate a bias toward probability information. We found that an attentional bias for magnitude information is associated with less risk aversion and an attentional bias toward probability information is associated with

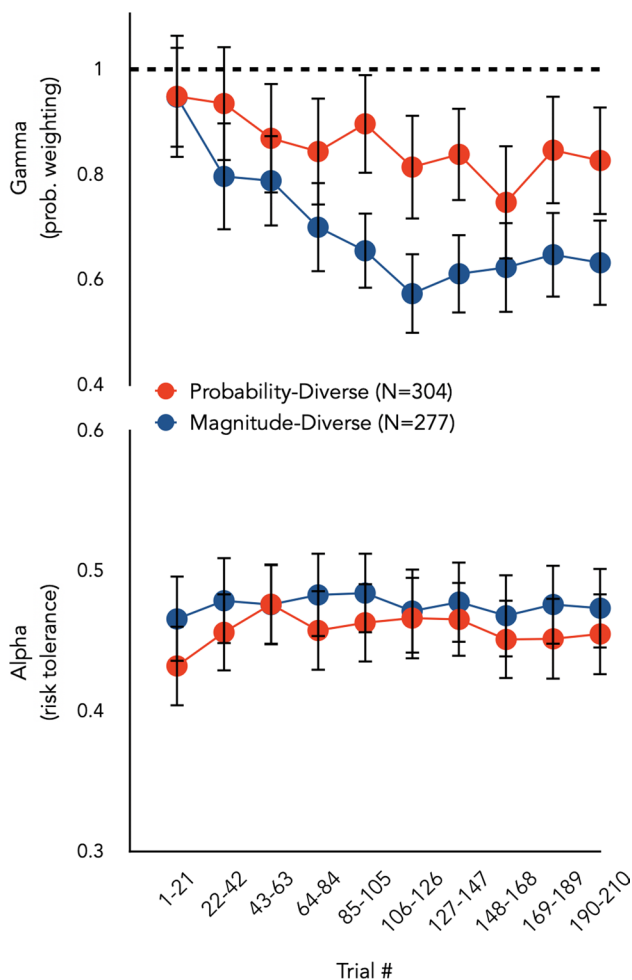


Fig. 4 Probability weighting and risk tolerance as a function of task exposure. See main text for details. Error bars, 95% confidence intervals

lower probability distortion. To show the consistency and reliability of these correlations, Fig. 5 depicts Spearman rank-order correlations for all trials, separately for each group and separately for each study. Restricting this analysis to the *Overlapping Choice Set* trials only, we observed consistent patterns of relationships between FBAI, risk tolerance, and probability weighting (Table 5), but given the much smaller number of trials in the subset, these correlations were statistically robust only in the web-based study, where the sample size was more than four times as large. We observed similar results with linear mixed-effects models in which we regressed trial-level FBAI on trial-level probability, trial-level magnitude, trial-level EV, subject-level gamma parameter rank, subject-level alpha parameter rank, and a dummy variable indicating group (see Supplementary Material, *Linear Mixed-effect Models with Trial-level FBAI*). In short, across all four LME models (web-based and lab-based datasets, assessing all trials and overlapping choice set trials), individual risk tolerance was positively correlated with trial-level FBAI, after controlling

for the aforementioned variables (Tables S10–S13). For three of the four models, individual probability weighting was negatively correlated with trial-level FBAI, after controlling for the aforementioned variables (Tables S10–S13); in the remaining regression (lab-based, overlapping choice set), this p -value was 0.0622 (Table S13).

Given the correlations between overt attention and choice within each individual context, we next asked if the attentional biases could explain the contextual differences in probability weighting. The evidence did not support this hypothesis. We reasoned that, if attentional biases explained the contextual modulations, FBAI should differ across the two contexts. However, while a difference was found in the lab study using eye-tracking (with median FBAI scores of 0.0121, 95% CI = [0.0049, 0.0208] in the *magnitude-diverse* context versus -0.0126 , 95% CI = [-0.0246 , 0.0072] in the *probability-diverse* context; between-group randomization test: $p = 0.0033$), this difference was not replicated in the Overlapping Choice Set (lab study, randomization test: $p > 0.8$; *magnitude-diverse*, median = 0.0004, 95% CI = [-0.0110 , 0.0192]; *probability-diverse*, median = 0.0025, 95% CI = [-0.0145 , 0.0205]) or in the web-based study (All trials: randomization test: $p = 0.1076$; *magnitude-diverse*, median = 0.0005, 95% CI = [-0.0003 , 0.0013]; *probability-diverse*, median = 0.0015, 95% CI = [0.0006, 0.0025]). OCS trials: randomization test: $p = 0.1203$; *magnitude-diverse*, median = 0.0004, 95% CI = [-0.0008 , 0.0025]; *probability-diverse*, median = 0.0025, 95% CI = [0.0005, 0.0047]). Thus, the contextual effects on FBAI were inconsistently replicated, making it improbable that they mediated between-group changes in probability weighting (which, as we noted above, were highly reliable in both participant samples and in the full and Overlapping Choice Set).

The linear mixed-effect models mentioned above also support this conclusion, as none of the models produced a significant effect of context (Tables S10–S13). A possible concern, however, is that these trial-level analyses may have missed contextual effects expressed at the individual subject level. To address this concern, we conducted additional subject-level analyses in which we fit the participants' probability distortion parameter (gamma) as a function of context and FBAI. Using the ranked gamma coefficients in the Overlapping Choice Set, we fit them with 3 alternative models in which the predictors were the ranked FBAI scores from the Overlapping Choice Set (Table 6, Model 1), a dummy predictor of context (isMagnitudeDiverse; Table 6, Model 3) or both the FBAI ranks and the dummy predictor (Table 6, Model 2). We reasoned that, if the FBAI accounted for the context effect, the latter effect should diminish or disappear in the simultaneous model. In contrast to this hypothesis, the coefficients for each term were approximately the same whether the predictors were entered individually or together, suggesting that the FBAI/gamma and context/gamma relationships were similar if we did or did not control for the other factor.

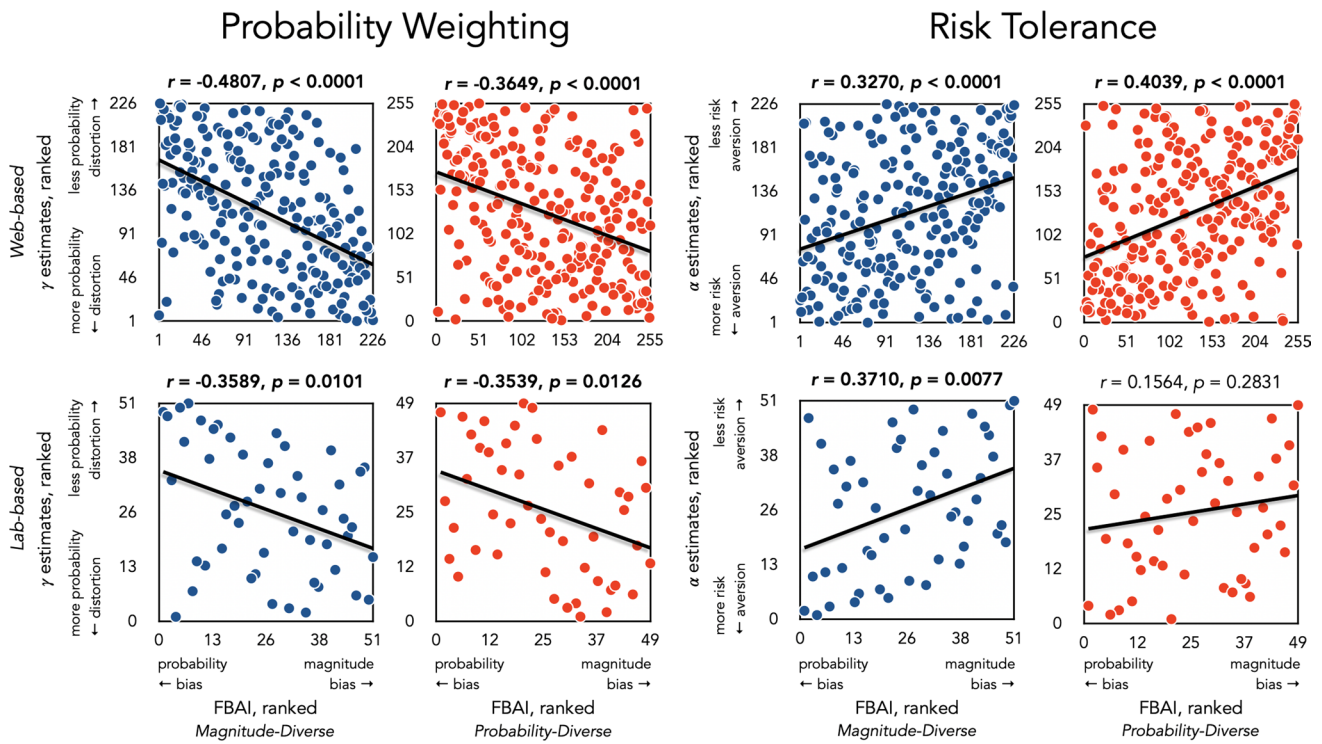


Fig. 5 Attention–choice correlations using individually estimated model parameters. Top row: web-based study. Bottom row: lab-based study. Leftmost columns: probability weighting. Rightmost columns: risk tolerance. Blue circles: participants in the *magnitude-diverse*

choice set. Red circles: participants in the *probability-diverse* choice set. Samples sizes for each Spearman correlation given by largest rank in each panel

Table 5 Attention–choice correlations using individually estimated model parameters from the OCS trials only

	Web-Based				Lab-Based			
	Alpha		Gamma		Alpha		Gamma	
OCS trials	MD	PD	MD	PD	MD	PD	MD	PB
Correlation with FBAI	0.2722	0.1360	−0.3107	−0.3057	0.1866	0.1236	−0.2393	−0.1495
p-value	<0.001	0.0309	<0.001	<0.001	0.1899	0.3973	0.0908	0.3054

MD magnitude-diverse, PD probability-diverse. Coefficients significantly different from zero in bold

Table 6 Attention–choice correlations and between-group differences in probability weighting

	Web-based		Lab-based	
	Coefficient	p-value	Coefficient	p-value
Model 1: gamma ~ FBAI				
FBAI	−0.29504	<0.001	−0.2007	0.0453
Model 2: gamma ~ FBAI + isMagnitudeDiverse	↕		↕	
FBAI	−0.3065	<0.001	−0.1998	0.0439
isMagnitudeDiverse	−38.708	0.00132	−10.164	0.0752
Model 3: gamma ~ isMagnitudeDiverse	↕		↕	
isMagnitudeDiverse	−31.883	0.01130	−10.224	0.0780

See text for details. Double-headed arrows indicate relevant comparisons

Discussion

We report a novel effect of the choice context on probability weighting during risky choice: a dependence on the ratio of unique magnitudes and probabilities in the choice set. Participants who experienced a choice set in which five unique probabilities were paired with 21 unique magnitudes made choices consistent with distorted probability weighting: overweighting of low probabilities and underweighting of high ones. In contrast, participants who experienced the converse (21 unique probabilities paired with five unique magnitudes) made choices consistent with a more veridical (i.e., optimal) assessment of probability information. Several lines of evidence attest to the robustness of this context effect. First, the effect replicated in two independent samples of participants who completed the task on Amazon Mechanical Turk and in the laboratory. Second, the effect held for a subset of lotteries that were identical in the two contexts, ruling out artifacts related to different choice options. Third, this context effect was specific to probability weighting without affecting risk aversion or the stochasticity of the choice, ruling out that it was due to nonspecific factors like arousal or general engagement in the task. Finally, the context effect emerged during the experiment, with between-group differences in probability weighting developing gradually as a function of exposure to the choice set, as would be expected if it were due to gradual learning of the context-specific probability/magnitude contingencies.

The context effect on probability weighting could not be accounted for by effects commonly discussed in the literature in the context of risk-less choice. In our design, the probabilities and magnitudes that participants encountered spanned identical ranges in both choice sets, so between-group differences in probability weighting could not be explained by range adaptation (Padoa-Schioppa, 2009). Moreover, the values in the two contexts had nearly identical means (probability, 0.5 in both contexts; magnitude, 28.5 vs. 31), and we did not include an irrelevant option in either choice context, ruling out potential explanations based on normalized value coding (Khaw et al., 2017) and value-based attention to irrelevant options (Gluth et al., 2020). Thus, our findings suggest that probability weighting is specifically sensitive to the number of discrete probabilities (vs. magnitudes), or to the precision of the probabilities, in the global choice context.

One plausible mechanism that may have explained our result involves changes in attention to the choice-relevant features. Many previous studies have shown that differences in attention—directed either to the lottery as a whole (Hunt et al., 2018; Smith & Krajbich, 2018; Stewart et al., 2016) or to its constituent parts (Fiedler & Glockner, 2012; Glickman et al., 2019; Harrison & Swarthout, 2019; Johnson &

Busemeyer, 2016; Kim et al., 2012; Spitmaan et al., 2019)—can explain differences in risky choice. Other authors have proposed that variability in a decision-relevant dimension influences attention (Horan et al., 2019; Pearce & Mackintosh, 2010) and saliency (Bordalo et al., 2012). Thus, participants may have differentially attended to probability versus magnitude information in the two contexts, potentially enabling them to estimate more veridically the probability in the *probability-diverse* relative to the *magnitude-diverse* context. In contrast to this prediction, we found that the robust group differences in probability weighting were not accompanied by analogous differences in attention. Consistent with an earlier study by Pachur and colleagues (2018), who looked at attention to outcome and probability information for gains and losses, we found that, in the gain domain, an attentional bias toward magnitude information was predictive of reduced risk aversion while a bias toward probability information was predictive of reduced probability distortion. However, these correlations were only found *within* each context and could not explain the differences in choice across contexts. The attention-choice correlations were approximately equivalent in the *magnitude-diverse* and *probability-diverse* contexts and controlling for attention did not alter the magnitude of the context effect on probability weighting. Thus, our findings suggest that the mechanisms by which context affects probability distortion are independent of our measure of selective attention—the time that participants allocated to inspecting magnitude versus probability information. An important question for future research, which may require a direct experimental manipulation of attention to resolve, is whether the effects are mediated by other aspects of selective attention, such as attention-induced changes in perceptual sensitivity and/or internal prioritization of (e.g., time spent thinking about) each type of information.

In the absence of an explanation based on overt attention, a plausible mechanism for our contextual effect involves differences in expectation for probability information. Predictive coding theories (Friston, 2010) posit that brains use a “simple but remarkable powerful trick or stratagem” to “get to know the world and act in it. [...] That trick is trying to guess at the incoming sensory stimulations as they arrive, using what [is known] about the world” (p.1, Clark, 2016). Psychophysical and neuroscientific evidence that expectation plays a fundamental role in perceptual decision making is quickly amassing (Egner et al., 2010; Kok et al., 2013; Rao & Ballard, 1999; Richter et al., 2018; for a review, Summerfield & de Lange, 2014), with extensive empirical work confirming that “[p]erception and perceptual decision-making are strongly facilitated by prior knowledge about the probabilistic structure of the world” (p. 764, de Lange et al., 2018). Thus, in our experiment, participants may have

implicitly learned the set of values available in a context, creating expectations that then influenced their treatment of probability information.

The effects of expectations may be mediated by several mechanisms. One possibility is related to the fact that the *probability-diverse* set included some “atypical” probabilities (e.g., 0.23, 0.62, 0.77), whereas the *magnitude-diverse* context contained more “standard” probabilities (e.g., 0.2, 0.5, 0.8). If participants bring to the task expectations, derived from long-term experience, about the probabilities and magnitude values that they are more likely to encounter in an experiment (i.e., expectations about common probabilities encountered in everyday life), this may influence their use of probability information. It is possible that participants in the *probability-diverse* environment find their expectations violated by the presence of many “atypical” probabilities, which causes a change in how they use probability information; such an implicit violation would not occur for the *magnitude-diverse* group, because they only encounter “standard” probabilities. A second possible explanation may lie in the fact that participants experienced more repeated exposure to the same unique probability in the *magnitude-diverse* relative to the *probability-diverse* context. The greater probability distortion in the *magnitude-diverse* set may come from short-term adaptation or repetition suppression driven by repeated and frequent exposure to the same probabilities (Grill-Spector et al., 2006; Krekelberg et al., 2006). A third hypothesis is based on the fact that participants could better anticipate the unique probability of each trial in the *magnitude-diverse* relative to the *probability-diverse* set (i.e., each unique probability had, respectively, a 1/5 chance vs. only 1/21 in the two contexts), suggesting that probability distortion may be enhanced by a better prior expectation of the specific probability. Dissecting the possible contributions of nonstandard probabilities, repetition suppression and prior expectations to our context effect will be important topics for future research.

A salient feature of our data was that, while between-group differences in probability weighting emerged gradually as a function of exposure to the choice context, risk tolerance remained constant and was insensitive to context. The stability of risk attitudes we found in this task is consistent with previous findings that risk tolerance is related to gray matter thickness in the parietal lobe—a structural parameter that is unlikely to change rapidly during an experimental session (Gilaie-Dotan et al., 2014; Grubb et al., 2016; Jung et al., 2018; Quan et al., 2022). Thus, although risk attitudes show marked differences in different domains and with different methods of elicitation (Hertwig et al., 2019), our results suggest that, for monetary lotteries, these attitudes are robust to differences in the composition of the choice set and are modulated by contextual factors distinct from those affecting probability weighting.

Conclusions

We documented a novel effect on probability weighting related to the number of discrete probabilities and magnitudes in the choice set, which provides a novel experimental manipulation for investigating the mechanisms of probability weighting and its implications for decision formation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13415-023-01062-y>.

Acknowledgments Study supported by NSF-2141860 CAREER Award to MAG, an NSF Research Opportunity Award to MG and IL, supplementing NSF grant BCS-1829439 to IL, an NIH/NIMH grant R01MH118215 to IL and by seed grants from Columbia University’s Data Science Institute and Research Initiative for Science and Engineering to JG.

Open practices statement Upon publication, the data upon which these results are based can be downloaded from www.AttentionPerceptionDecision.com/GLLHGL_CABN.

References

- Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, 16(3), 215–233.
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *Quarterly Journal of Economics*, 127(3), 1243–1285. <https://doi.org/10.1093/qje/qjs018>
- Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford ; New York: Oxford University Press.
- de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences*, 22(9), 764–779. <https://doi.org/10.1016/j.tics.2018.06.002>
- Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience*, 30(49), 16601–16608. <https://doi.org/10.1523/JNEUROSCI.2770-10.2010>
- Fiedler, S., & Glockner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, 3.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews. Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Gilaie-Dotan, S., Tymula, A., Cooper, N., Kable, J. W., Glimcher, P. W., & Levy, I. (2014). Neuroanatomy predicts individual risk attitudes. *The Journal of Neuroscience*, 34(37), 12394–12401. <https://doi.org/10.1523/JNEUROSCI.1600-14.2014>
- Glickman, M., Sharoni, O., Levy, D. J., Niebur, E., Stuphorn, V., & Usher, M. (2019). The formation of preference in risky choice. *PLoS Computational Biology*, 15(8), e1007201. <https://doi.org/10.1371/journal.pcbi.1007201>
- Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*, 4(6), 634–645. <https://doi.org/10.1038/s41562-020-0822-0>
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23. <https://doi.org/10.1016/j.tics.2005.11.006>
- Grubb, M. A., Tymula, A., Gilaie-Dotan, S., Glimcher, P. W., & Levy, I. (2016). Neuroanatomy accounts for age-related changes in risk

- preferences. *Nature Communications*, 7, 13822. <https://doi.org/10.1038/ncomms13822>
- Harrison, G. W. (2008). *Maximum likelihood estimation of utility functions using Stata* Working Paper 6–12. University of Central Florida.
- Harrison, G. W., & Rutstrom, E. E. (2008). Risk aversion in the laboratory. *Research in Experimental Economics*, 12, 41–196.
- Harrison, G. W., & Swarthout, J. T. (2019). Eye-tracking and economic theories of choice under risk. *Journal of the Economic Science Association-Jesa*, 5(1), 26–37. <https://doi.org/10.1007/s40881-019-00063-3>
- Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12), 517–523. <https://doi.org/10.1016/j.tics.2009.09.004>
- Hertwig, R., Wulff, D. U., & Mata, R. (2019). Three gaps and what they may mean for risk preference. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 374(1766), 20180140. <https://doi.org/10.1098/rstb.2018.0140>
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655. <https://doi.org/10.1257/000282802762024700>
- Horan, M., Daddaoua, N., & Gottlieb, J. (2019). Parietal neurons encode information sampling based on decision uncertainty. *Nature Neuroscience*, 22(8), 1327–1335. <https://doi.org/10.1038/s41593-019-0440-1>
- Hunt, L. T., Malalasekera, W. M. N., de Berker, A. O., Miranda, B., Farmer, S. F., Behrens, T. E. J., & Kennerley, S. W. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. *Nature Neuroscience*, 21(10), 1471–1481. <https://doi.org/10.1038/s41593-018-0239-5>
- Jiang, Y. V. (2018). Habitual versus goal-driven attention. *Cortex*, 102, 107–120. <https://doi.org/10.1016/j.cortex.2017.06.018>
- Johnson, E. J., Payne, J. W., Bettman, J. R., & Schkade, D. A. (1989). *Monitoring information processing and decisions: The mouse-lab system* (Vol. No. TR-89-4-ONR). Duke University Center for Decision Studies.
- Johnson, J. G., & Busemeyer, J. (2016). A computational model of the attention process in risky choice. *Decision*, 3, 254–280.
- Jung, W. H., Lee, S., Lerman, C., & Kable, J. W. (2018). Amygdala functional and structural connectivity predicts individual risk tolerance. *Neuron*, 98(2), 394–404 e394. <https://doi.org/10.1016/j.neuron.2018.03.019>
- Kahneman, D., & Tversky, A. (1979). Prospect theory - analysis of decision under risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences of the United States of America*, 114(48), 12696–12701. <https://doi.org/10.1073/pnas.1715293114>
- Kim, B. E., Seligman, D., & Kable, J. W. (2012). Preference reversals in decision making under risk are accompanied by changes in attention to different attributes. *Frontiers in Neuroscience*, 6(109).
- Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *The Journal of Neuroscience*, 33(41), 16275–16284. <https://doi.org/10.1523/JNEUROSCI.0742-13.2013>
- Krekelberg, B., Boynton, G. M., & van Wezel, R. J. (2006). Adaptation: From single cells to BOLD signals. *Trends in Neurosciences*, 29(5), 250–256. <https://doi.org/10.1016/j.tics.2006.02.008>
- Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*, 29(44), 14004–14014. <https://doi.org/10.1523/Jneurosci.3751-09.2009>
- Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and possible integration. In C. J. Mitchell & M. E. LePelle (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 11–39). Oxford University Press.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527. <https://doi.org/10.2307/2998573>
- Quan, P., He, L., Mao, T., Fang, Z., Deng, Y., Pan, Y., & Rao, H. (2022). Cerebellum anatomy predicts individual risk-taking behavior and risk tolerance. *Neuroimage*, 254, 119148. <https://doi.org/10.1016/j.neuroimage.2022.119148>
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87. <https://doi.org/10.1038/4580>
- Richter, D., Ekman, M., & de Lange, F. P. (2018). Suppressed sensory response to predictable object stimuli throughout the ventral visual stream. *The Journal of Neuroscience*, 38(34), 7452–7461. <https://doi.org/10.1523/JNEUROSCI.3421-17.2018>
- Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, 32, 15–20. <https://doi.org/10.1016/j.cobeha.2020.01.015>
- Smith, S. M., & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology. General*, 147(12), 1810–1826. <https://doi.org/10.1037/xge0000482>
- Spitmaan, M., Chu, E., & Soltani, A. (2019). Salience-driven value construction for adaptive choice under risk. *The Journal of Neuroscience*, 39(26), 5195–5209.
- Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making*, 29(2-3), 116–136. <https://doi.org/10.1002/bdm.1854>
- Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11), 745–756. <https://doi.org/10.1038/nrn3838>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.