Research Article

Psychological Science
2021, Vol. 32(5) 743–754

DOI: 10.1177/0956797620977516
www.psychologicalscience.org/PS
SAGE

Encoding Context Determines Risky Choice

Christopher R. Madan[1], Marcia L. Spetch[2], Fernanda M. D. S. Machado[3], Alice Mason[3], and Elliot A. Ludvig[3]
[1]School of Psychology, University of Nottingham; [2]Department of Psychology, University of Alberta; and [3]Department of Psychology, University of Warwick

Abstract
Both memory and choice are influenced by context: Memory is enhanced when encoding and retrieval contexts match, and choice is swayed by available options. Here, we assessed how context influences risky choice in an experience-based task in two main experiments (119 and 98 participants retained, respectively) and two additional experiments reported in the Supplemental Material available online (152 and 106 participants retained, respectively). Within a single session, we created two separate contexts by presenting blocks of trials in distinct backgrounds. Risky choices were context dependent; given the same choice, people chose differently depending on other outcomes experienced in that context. Choices reflected an overweighting of the most extreme outcomes within each local context rather than the global context of all outcomes. When tested in the nontrained context, people chose according to the context at encoding and not retrieval. In subsequent memory tests, people displayed biases specific to distinct contexts: Extreme outcomes from each context were more accessible and judged as more frequent. These results pose a challenge for theories of choice that rely on retrieval as guiding choice.

Keywords
risky decision-making, memory, decisions from experience, memory biases, behavioral economics, context, encoding, open data, open materials, preregistered

Received 1/31/20; Revision accepted 10/7/20

People's decisions are often informed by prior experiences, reflecting the influence of memory on decision-making (e.g., Hertwig et al., 2004; Ludvig et al., 2015; Murty et al., 2016; Shohamy & Daw, 2015). Context has a large impact on memory (see Stark et al., 2018, for a review), leading, for example, to reduced recall when the location changes between study and test (Hupbach et al., 2007; Smith et al., 1978) and playing a prominent role in computational models of memory recall (Howard & Kahana, 2002). Context also significantly influences choice: Other available options in a context can lead to range adaptation (Bavard et al., 2018) or even preference reversal in multiattribute choice (Huber et al., 1982). Some researchers have posited that choice is determined by context-dependent samples drawn from memory (e.g., Stewart et al., 2006). Here, we show that people choose differently between functionally identical pairs of risky options and remember them differently depending on the context. Moreover, we show that choice is determined by the set of available options present during encoding rather than at retrieval.

Contextual information from the local environment can influence choices. For example, when French music is playing in a supermarket, people buy more French than German wine, and the opposite is true when German music is played (North et al., 1997). Similarly, locating polling stations in schools nudges people toward support of school funding (Berger et al., 2008;

Corresponding Author:
Christopher R. Madan, University of Nottingham, School of Psychology
E-mail: christopher.madan@nottingham.ac.uk

Pryor et al., 2014). The local context provided by other available options can also influence choice (Huber et al., 1982; Simonson, 1989; Simonson & Tversky, 1992; Spektor et al., 2019). Consumer preference between two multidimensional products can reverse when a third, decoy option is introduced that is inferior along one dimension (e.g., cost or quality). Nonhuman animals also show similar local-context effects in their choices (e.g., Shafir et al., 2002).

Experience-based risky choices are also influenced by the set of available values in a decision context. When making decisions on the basis of experience, people tend to show more risk-seeking behavior for the possibility of relative gains than relative losses—but the opposite pattern appears for decisions made from explicit descriptions (e.g., Hertwig & Erev, 2009; Kahneman & Tversky, 1979; Konstantinidis et al., 2018; Ludvig & Spetch, 2011; Wulff et al., 2018). This pattern of experienced-based risky choice appears to be driven by overweighting of the most extreme (i.e., best and worst) outcomes in the decision context (Ludvig et al., 2014, 2018). This effect of extremes was confirmed by including other options in the decision context that potentially led to higher (or lower) outcomes, thereby eliminating the bias in risky choice. Moreover, these biases in choice correlate with biases in memory for the extreme outcomes (Madan et al., 2014, 2019).

People will sometimes even choose differently for identical decisions across experiments that have different ranges of possible outcomes, suggesting session-level context dependence (Ludvig et al., 2014; Stewart et al., 2015). For example, one decision in Ludvig et al.'s study was between a fixed gain of 20 points and a risky option associated with a 50/50 chance of winning 40 points or nothing. People showed more risk aversion toward this decision in an experiment that included other, larger wins (so that winning nothing was the worst possible outcome) than in an experiment that also included losses (so that winning 40 points was the best possible outcome). Thus, across experiments involving different decision sets, these differences in risky choice for identical decisions imply that context is an important determinant of risky choice.

Here, we tested whether people's choices shift with context changes even within a single experimental session and whether context-dependent effects on choice are based on the decision set present at encoding or retrieval. The main text reports two experiments, and the Supplemental Material available online contains two additional experiments that replicated the main findings and refined what determines the decision context. Stimuli and data from all experiments are available on OSF at https://osf.io/3mbwu.

Statement of Relevance

People make risky choices in a variety of contexts, whether gambling at a casino, selecting a stock portfolio, or deciding which traffic-prone route to drive on the way home. The context determines the range of available options and outcomes, influencing what people choose. Context, such as location or time of day, also influences what people remember. Here, in a series of experiments, we assessed how people make risky choices when they learn about the odds and outcomes from their own experience. We show that people select differently even between identical options when those options appear in different contexts. Moreover, we show that people's memories and risky choices depend on the context in which options are initially encountered rather than the context at decision time. These results provide a novel demonstration of how memory for past outcomes influences choice and have wide-reaching implications for theories of memory and choice.

Experiment 1: Local Decision Contexts

This experiment tested the stability of choice behavior by evoking distinct decision contexts that alternated within a session. In memory research, discrete contexts are often evoked through distinct background images (e.g., Anderson & Bower, 1974; Ezzyat & Davachi, 2014). Inspired by this approach, we designed the current experiment to provide different contexts by alternating between blocks of decisions with distinct background images and choice options (Fig. 1). One choice (between a fixed gain of 20 points and a risky gain of 10 or 30 points) was common to both contexts and served as the target choice. In the gain/loss context, other values were a fixed loss of 20 points and a risky loss of 10 or 30 points. In the high/low context, other values were a fixed gain of 60 points and a risky gain of 50 or 70 points. Thus, the target risky option provided the best possible outcome (+30) in the gain/loss context but the worst possible outcome (+10) in the high/low context.

If decision contexts create discrete sets of memories, then the extreme-outcome rule predicts that the best and

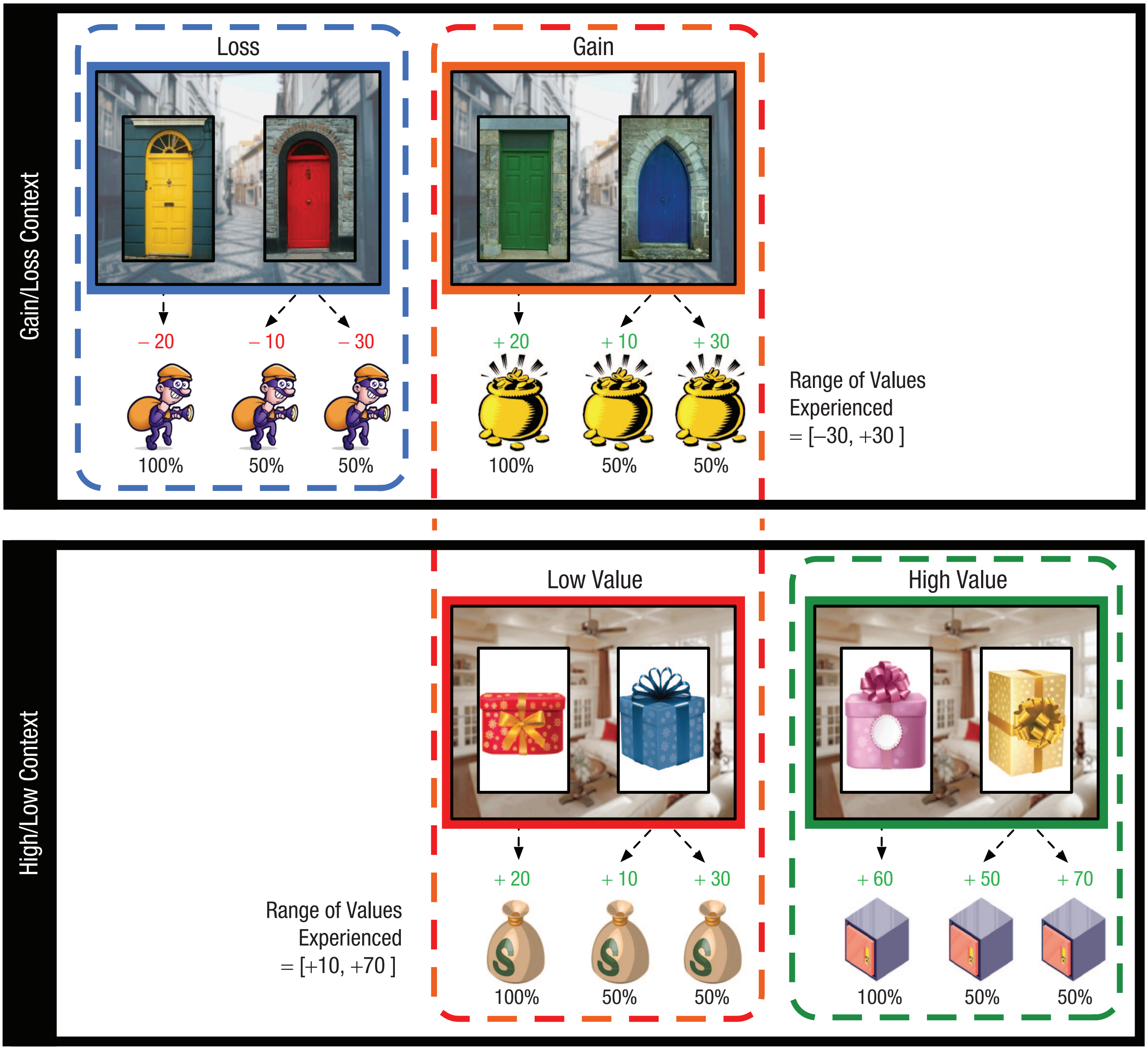

Fig. 1. Illustration of the options, outcomes, and context manipulations used in Experiment 1. Participants were first shown the choice options (e.g., two doors) along with a background image. After the participants made their choice, the chosen door was replaced with an outcome image (e.g., robber or pot of gold) indicating the number of points won or lost following the outcome contingencies shown here below each image; the unchosen door was no longer shown. To differentiate among the four option pairs (losses, gains, low value, high value), we used different option images (distinct doors or distinct gift boxes) and different outcome images (robber, pot of gold, bag of money, and safe, respectively). The target choices, outlined by the orange/red dashed line, had identical values in the two contexts indicated.

worst outcomes in each local context will be overweighted in memory and choice (Ludvig et al., 2014). This overweighting would produce more risk seeking for the target choice in the gain/loss context than in the high/low context (see the comparison outlined in orange/red dashes in Fig. 1). If people do not distinguish the contexts, risky choice should be identical in both cases, as the options yield the same values. In either case, we expected that people would show more risk seeking for the highest value decisions (+60 vs. +50 or +70) and more risk aversion for the lowest value decisions (−20 vs. −10 or −30).

Method

Participants. A total of 128 participants (99 women; age: M = 19.4 years, SD = 1.9) were recruited from the University of Alberta psychology participant pool. An additional 52 participants were recruited but were instructed and paid according to an incorrect payment scheme; consequently, their data were excluded and not analyzed. Informed consent was obtained, and participants received course credit and a cash bonus for participating. They were instructed in groups of up to 15 but performed the task in individual rooms. The number recruited exceeded the number needed (97) to detect a medium-size effect (Cohen's $d = 0.4$) with an α of .01, according to a power analysis for this within-subjects design. Procedures were approved by the University of Alberta Research Ethics Board.

Procedure. The experiment consisted of six blocks of trials. Blocks providing a gain/loss context, indicated by an outdoor background image, alternated with blocks providing a high/low context, indicated by an indoor background image (Fig. 1). Fixed options always led to the same outcome, whereas risky options provided two outcomes each with a 50% chance. In the gain/loss context, options were selected from four doors that each led to a different possibility: a fixed gain (+20), a risky gain (+10 or +30), a fixed loss (−20), or a risky loss (−10 or −30). In the high/low context, options were selected from four gifts that each led to a different possibility: a fixed high-value gain (+60), a risky high-value gain (+50 or +70), a fixed low-value gain (+20), or a risky low-value gain (+10 or +30). Thus, there were four different option pairs in the experiment: gain, loss, high value, and low value. Critically, as highlighted by the orange/red dashed box in Figure 1, the target choices—gain options in the gain/loss context and low-value options in the high/low context—led to identical outcome values, but their relative values within their respective contexts differed. Participants could learn about the odds and outcomes only by selecting the options.

After a choice, the options disappeared, and feedback for the chosen option appeared for 1.2 s. Feedback consisted of the points earned or lost along with an outcome image. The order of the two contexts was counterbalanced across participants, as was the assignment of options to particular outcomes.

For each context, prior to the first block of choice trials, participants completed 24 single-option training trials so they would be experienced with the experimental procedure. For these trials, the outcomes associated with each risky option occurred equally often, which prevented differences in initial experiences from influencing later choice (e.g., hot-stove or primacy effects; Denrell & March, 2001). Within this block, the gain or high-value options each appeared eight times, whereas the loss or low-value options each appeared four times; consequently, participants ended the training phase with a positive number of points in both contexts.

Each block of choices consisted of 56 trials and included a mixture of trial types: There were 32 decision trials, which required a choice between fixed and risky options from the same option pairs (16 of each), and 16 catch trials, which required a choice between options from different option pairs with substantially different expected values (e.g., fixed gain vs. fixed loss). On eight single-option trials, there was only one option that had to be selected to continue; these trials guaranteed that all reward contingencies continued to be experienced, even if the options were initially unlucky. This further limited any hot-stove effects.

In all blocks, trial order was randomized, and each option appeared equally often on either side of the screen. Performance lower than 60% on catch trials in either context, across the whole experiment, was used as an exclusion criterion, following established protocol from previous experiments (Ludvig et al., 2014; Ludvig & Spetch, 2011; Madan et al., 2014). Participants won or lost points on all trials and were paid $1 for every 2,000 points to a maximum of $5 (Canadian).

After the choice task, memory for the outcomes associated with each option was tested in two ways. First, participants were shown the eight options in random order, and, for each option, were asked to report the first outcome that came to mind. Second, participants were again shown the eight options in random order and asked to judge the frequency in percentages of each possible outcome (−30, −20, −10, +10, +20, +30, +50, +60, +70). For each option, these nine possible outcomes were displayed simultaneously, and participants typed a number from 0 to 100 below each respective outcome. For both memory tests, each option was presented against a uniform gray background on all trials.

Analysis. Data from 9 of the 128 participants were excluded from the analyses because these participants scored below 60% on the catch trials, leaving 119 participants for the main analyses. The primary dependent measure was the proportion of risky choices in the final choice blocks. Two specific hypotheses were tested: the *decision-context hypothesis* and the *contextual-memory hypothesis*.

The decision-context hypothesis posits that the extreme outcomes in each context will be overweighted. As a result, risky choice should be higher for the gain/high-value options (with a high extreme) than for the loss/low-value options (with a low extreme) in the corresponding context. In addition, the target choice that has identical outcomes (i.e., low value or gain, depending on the context) should differ between the two contexts, with more risk seeking for that choice in the gain/loss context than in the high/low context. These directional predictions were assessed through three one-tailed paired-samples t tests.

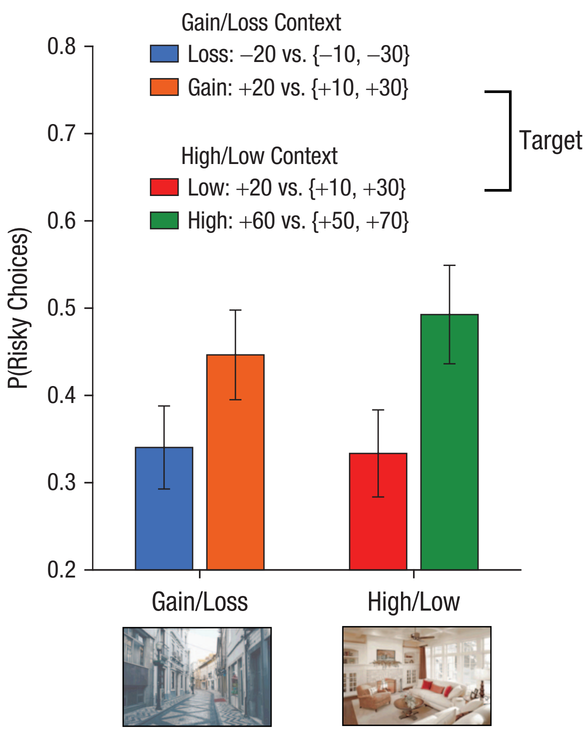

Fig. 2. Proportion of risky choices for each of the four option pairs, separated by their respective decision contexts and averaged across the last block in each context for Experiment 1. Error bars represent 95% confidence intervals.

The contextual-memory hypothesis posits that by the last block in each context, the extreme outcomes in each context will be more salient in memory. For the first-outcome-reported test, this hypothesis was assessed using four χ^2 tests—one for each risky option. For the frequency-judgment tests, this hypothesis was assessed using four one-tailed paired-samples *t* tests, again one for each risky option. Data from a participant were excluded for a given door if the summed responses for that door fell below 75% or rose above 150%. On the basis of prior work, we expected a robust effect for the low-value risky-loss option but a milder effect for the high-value gain options because we have previously found memory biases to be weaker for gain/high-value outcomes than for the loss/low-value outcomes (e.g., Madan et al., 2014, 2017, 2019).

Results

Risky choice. Figure 2 shows the mean proportion of risky choice for each context and option pair. In the gain/loss context, participants were 10.6% more risk seeking for gains than losses (95% confidence interval [CI] = [4.0%, 17.2%]), $t(118) = 3.15$, $p = .001$, Cohen's $d = 0.39$. In the high/low context, participants were 15.9% more risk seeking for high-value than low-value options (95% CI = [9.3%, 22.5%]), $t(118) = 4.73$, $p < .001$, $d = 0.54$. These results qualitatively replicated our previous findings of an extreme-outcome effect, providing evidence for greater differences in risky choice for high- versus low-value gains than for gains versus losses (Ludvig et al., 2014; Madan et al., 2014).

Critically, when comparing choice in the two contexts, participants were 11.3% more risk seeking for the target choices in the gain/loss context than in the high/low context (95% CI = [5.0%, 17.6%]; see the comparison highlighted in Fig. 1 and the orange and red bars in Fig. 2), despite these options leading to identical outcome values, $t(118) = 3.52$, $p < .001$, $d = 0.40$. Interestingly, the magnitude of the extreme-outcome effect in the final block of each context was uncorrelated between the two contexts, $r(117) = −.04$, $p = .69$, indicating that the two contexts had been learned relatively independently. Overall risk seeking collapsed across gains and losses, however, was correlated between the two contexts, $r(117) = .45$, $p < .001$.

Thus, participants' biases in risky choice shifted as the visually distinct contexts alternated between blocks. The effect was sufficiently pronounced that even for the identical target choice (between +20 and a 50/50 chance of +10 or +30), risky choice shifted by more than 10% even within the same participants within the same session, depending on the decision context.

Memory tests. Figure 3 shows how both memory tests suggested some overweighting of the extreme outcomes, consistent with prior findings (Madan et al., 2014), as well as some context dependence in overweighting. The memory biases were more robust for the loss and low-value options, which is also consistent with prior work.

In the first-outcome-reported test, for both the loss and low-value options, participants were significantly more likely to report the worse value (−30 and +10, respectively)—loss: $\chi^2(1, N = 88) = 35.64$, $p < .001$; low value: $\chi^2(1, N = 92) = 31.70$, $p < .001$. Participants did not exhibit a bias in their reported outcomes for the gain option, $\chi^2(1, N = 85) = 0.11$, $p = .74$, and there was only a weak trend toward responding with the better outcome for the high-value option, $\chi^2(1, N = 99) = 2.92$, $p = .088$. Results were similar in the frequency-judgment test, in which people reported a significantly larger percentage for the worse outcome for the loss and low-value options—loss: $t(102) = 6.16$, $p < .001$, $d = 1.06$; low value: $t(102) = 7.02$, $p < .001$, $d = 1.19$—but did not report a reliable difference in judged percentage for the outcomes of the gain and high-value options—gain: $t(102) = 0.29$, $p = .39$, $d = 0.05$; high value: $t(102) = 0.82$, $p = .21$, $d = 0.14$. Thus, for both measures, the worst outcome in each context seemed to be particularly salient in memory. The context dependence of this salience is highlighted by the +10 outcome, which was reported more often and

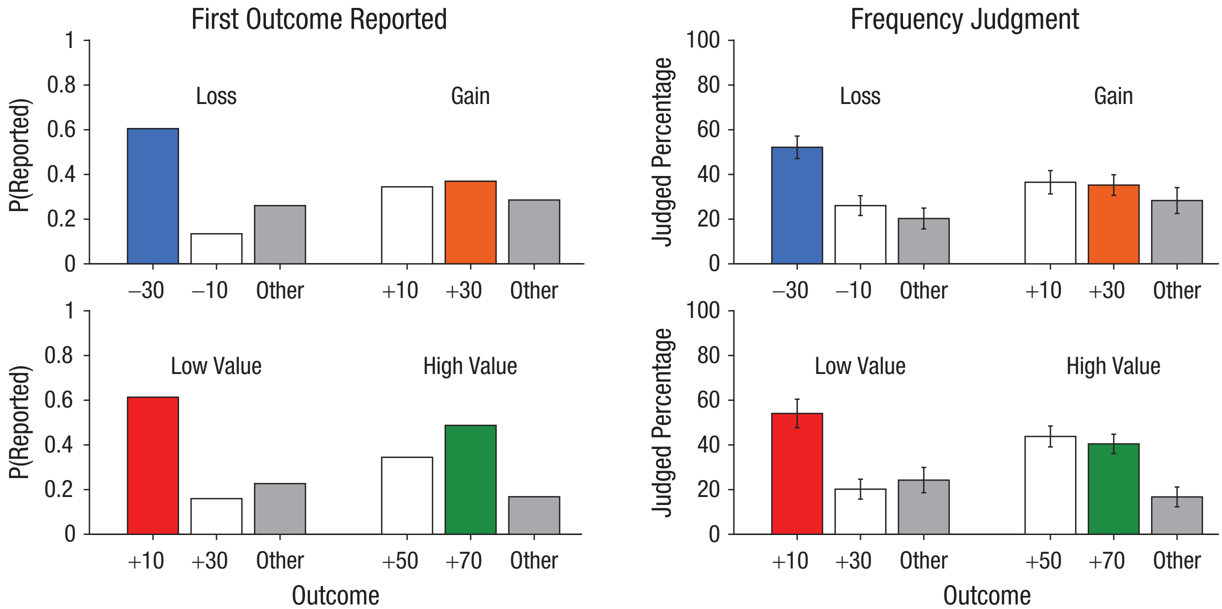

Fig. 3. Results of the two memory tests in Experiment 1. The graphs on the left show the proportion of participants who reported each outcome as the first that came to mind, and the graphs on the right show participants' judgments of the frequency of each outcome. For both tests, results are shown separately for the gain/loss context (top) and the high/low context (bottom). Colored bars indicate local extreme outcomes, and white bars indicate nonextreme outcomes. The color code matches the conditions in Figure 1 (blue = loss, orange = gain, red = low, green = high). Error bars represent 95% confidence intervals.

judged as having a higher frequency in the high/low context than in the gain/loss context.

Discussion

The context manipulation in this experiment successfully established distinct decision contexts. Participants made different risky choices even for option pairs that led to identical values; choices depended on the other values present in the same context (i.e., choices in the gain and low-value decisions, as highlighted in Fig. 2). The memory tests also showed context dependence: People were more likely to report the extreme outcomes in each context as the first to come to mind and judged the worst outcome in each context as more frequent (see Fig. 3). Though we have previously demonstrated different risky choice for options leading to the same outcomes across experiments (e.g., Ludvig et al., 2014; Madan et al., 2014), this experiment is the first to demonstrate that risk preference for a given decision and related memory biases can differ across blocks of trials within a single session on the basis of the local context.

Experiment 2: Encoding or Retrieval of Contextual Cues

Here, we sought to extend the findings of Experiment 1 by testing whether the context of encoding or retrieval is crucial for determining which outcomes are overweighted in memory and choice. The results of Experiment 1 could be due to processes operating at either encoding or retrieval. From an encoding perspective, outcome values might be encoded relative to the other values present in the context during learning (Rangel & Clithero, 2012). Values at the extremes of that set may be given more weight during encoding, causing them to be retrieved or sampled more readily when the option is later reexperienced. An encoding account is also congruent with a selective-attention mechanism whereby goal-congruent items influence value integration (e.g., Kunar et al., 2017; Usher et al., 2019).

Alternatively, context-dependent biases could be due to retrieval processes during choice. For example, if outcome values are encoded together with an association to their learning context, then the context present during choice may retrieve a memory of other values associated with that context. This retrieved set of values may determine the comparison set for evaluating values during choice (as in decision by sampling; Stewart et al., 2006), with extreme values being given most weight. A retrieval-based interpretation is consistent with findings that risky choice can be altered by presenting reminders of previous outcomes (Bornstein et al., 2017; Ludvig et al., 2015).

To distinguish between encoding and retrieval hypotheses, we used the same design as in Experiment

1 but with two modifications. First, choice stimuli and background images were altered to make the target options more interchangeable. Specifically, we used eight distinct doors (rather than four doors and four gifts) as choice stimuli and two distinct street scenes as background images for the two decision contexts. Second, after the six choice blocks, we presented two blocks of probe tests without feedback, in which the doors providing the target choice were presented in either their training context (same) or the untrained context (reversed).

If the context of encoding is crucial, choices should be independent of the testing context. Participants should be more risk seeking for the target choices initially encountered in the gain/loss context than for those initially encountered in the high/low context, regardless of the test context. If the context of retrieval determines choice, however, then people should choose differently between the same pairs of doors in the two testing contexts. Specifically, participants should be more risk seeking for both target choices when tested in the gain/loss context than in the high/low context. The design, hypotheses, analysis plan, and expected choice results for this experiment were preregistered on OSF (https://osf.io/kv458).

Method

Participants. A total of 103 participants (72 women; age: M = 20.8 years, SD = 3.4) were drawn from the same participant pool, and recruitment and consent procedures were the same as in Experiment 1. Participants were paid $1 (Canadian) for every 200 points after the first 8,000 earned up to a maximum of $5.

Procedure. General procedures were the same as in Experiment 1, with the following exceptions. The task consisted of eight blocks. The first six blocks alternated between two contexts in which four possible doors appeared alone or in pairs against a background outdoor scene that was unique to each context; these will be referred to as the training blocks. The last two blocks were test blocks, one for each context. In these blocks, choices were not followed by feedback. Prior to these two blocks, participants were informed by an instruction screen that they would not receive feedback for their choices but that points would still be won or lost in the same way as before.

Trials during the training blocks were identical to those in Experiment 1 except that all choice stimuli were doors and the two backgrounds were distinct street scenes rather than an outdoor and indoor scene. In the test blocks, only the doors that led to the target choice of either +20 or a 50/50 chance of +10 or +30 appeared. These were tested in both contexts (order randomized across participants) without feedback. There were two test blocks of 16 trials each, providing a total of eight trials with each target choice in each context.

Following the test blocks, participants were given the same two types of memory tests (first outcome reported and frequency judgment) described in Experiment 1.

Analysis. Five participants were excluded from the analysis for scoring less than 60% on the catch trials, leaving 98 participants. Following our preregistration, we evaluated comparisons with an α of .01. The primary dependent measure was the proportion of risky choices in the final training blocks and in the test blocks. Four specific preregistered hypotheses were tested: the *context-replication hypothesis*, the *encoding hypothesis*, the *retrieval hypothesis*, and the *noise hypothesis.*

The *context-replication hypothesis* posits that by the end of training, the extreme outcomes in each context will be overweighted. This hypothesis was assessed through three one-tailed paired-samples t tests. First, we tested the prediction that risky choice would be higher for the higher-value option (high or gain) than for the lower-value option (low or loss) in both contexts in the final block of the training phase. Second, we compared risky choice for the target choice in the two contexts. We predicted more risk seeking for that choice in the gain/loss context than in the high/low context.

The *encoding hypothesis* posits that the context effects are due to the way the doors were initially encoded in the training contexts. To test this, we predicted that, regardless of the test context, there would be more risk seeking for the target choice learned in the gain/loss context than for the target choice learned in the high/low context. This was assessed with two one-tailed paired-samples t-tests examining risky choice for the target in the two contexts during testing.

The *retrieval hypothesis* posits that the context effects are due to the context in which outcomes are retrieved at the time of choice. This means that the test context should matter and that for the target choice, people should be more risk seeking when tested in the gain/loss context (in which the other options were worse) than in the high/low context (in which the other options were better). This was assessed through a two-way (Training Context × Test Context) repeated measures analysis of variance (ANOVA). This hypothesis predicted a main effect of test context.

The *noise hypothesis* posits that the context shift in the test blocks makes people behave more randomly because the discrepant context makes them rely less

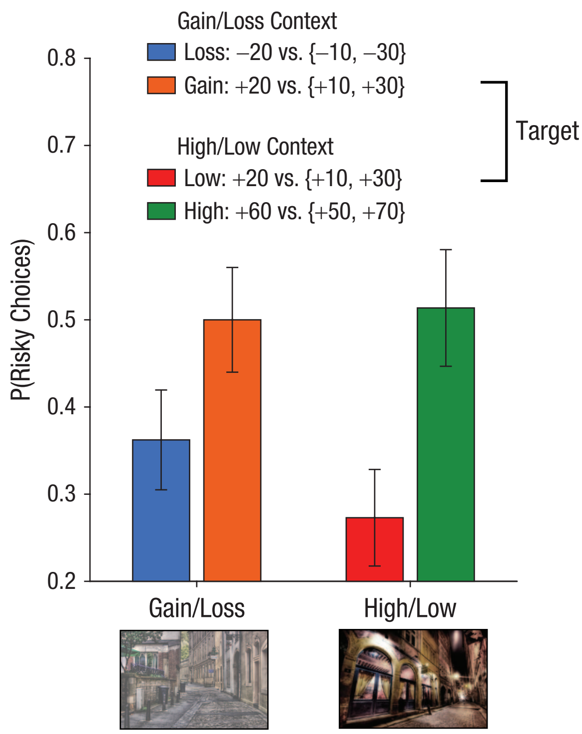

Fig. 4. Proportion of risky choices for each of the four option pairs, separated by their respective decision contexts and averaged across the last block in each context for Experiment 2. Error bars represent 95% confidence intervals.

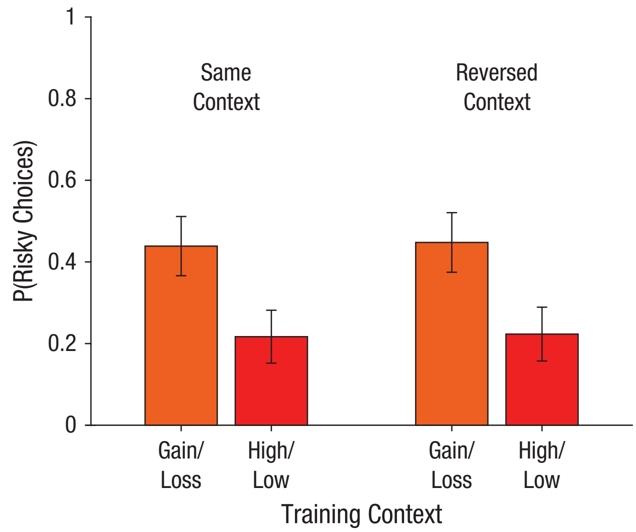

Fig. 5. Results of the probe-choice tests in Experiment 2. The graph shows the proportion of risky choices for the target choice (+20 vs. +10 or +30) trained in the gain/loss context and the high/low context when tested in the same and reversed context without feedback. Error bars represent 95% confidence intervals.

on their prior feedback. As a result, choice should shift toward indifference whenever participants tested doors outside their training context. This hypothesis was tested by calculating the difference between each individual's average absolute deviation from 50% in their risky choices in the two test contexts; a shift toward indifference with a context change should result in lower absolute-deviation scores in the reversed context. A one-tailed one-sample *t* test was used to test for reliable differences from 0 across the two contexts.

Memory tests were analyzed in the same way as in Experiment 1. We did not preregister specific predictions for these tests.

Results

Risky choice. Figure 4 shows the mean proportion of risky choices for each context and option pair during the last training block with each context. In the gain/loss context, participants were 13.8% more risk seeking for gains than losses (95% CI = [6.2%, 21.4%]), $t(97) = 3.62$, $p < .001$, $d = 0.37$. In the high/low context, participants were 24.1% more risk seeking for high-value than low-value choices (95% CI = [15.8%, 32.4%]), $t(97) = 5.72$, $p < .001$, $d = 0.58$. These results qualitatively replicate results from Experiment 1.

Critically, when comparing the two contexts, we found that participants were 22.7% more risk seeking for the target choice in the gain/loss context than in the high/low context ($SD = 7.9\%$), $t(97) = 5.68$, $p < .001$, $d = 0.57$. The magnitude of the extreme-outcome effect was again uncorrelated between the two contexts, $r(97) = .037$, $p = .72$, indicating that the two contexts were learned relatively independently. Overall risk seeking (collapsed across all risky decisions) was slightly, but not significantly, correlated between the two contexts, $r(97) = .191$, $p = .058$.

Test blocks. Figure 5 shows the mean risky choice for the target choices when they were presented without feedback during testing. The test context had no discernable effect. When tested in the same context, participants were 22.2% more risk seeking for the target choice trained in the gain/loss context than in the high/low context (95% CI = [12.4%, 32.0%]), $t(97) = 4.50$, $p < .001$, $d = 0.46$. Similarly, when tested in the reversed context, participants were 22.5% more risk seeking for the target choice trained in the gain/loss context than in the high/low context (95% CI = [12.5%, 32.5%]), $t(97) = 4.48$, $p < .001$, $d = 0.45$. A two-way ANOVA confirmed a main effect of choice, $F(1, 97) = 21.1$, $p < .001$, $\eta_p^2 = .18$, but no effect of test context, $F(1, 97) = 0.51$, $p = .48$, $\eta_p^2 = .005$, and no interaction, $F(1, 97) = 0.015$, $p = .90$, $\eta_p^2 = .00$.

There was no evidence in support of the noise hypothesis: The average deviation from indifference (0.5) did not differ for risky choices conducted in the same context ($M = 35.7\%$, 95% CI = [32.7%, 38.7%]) from the risky choices conducted in the reversed context

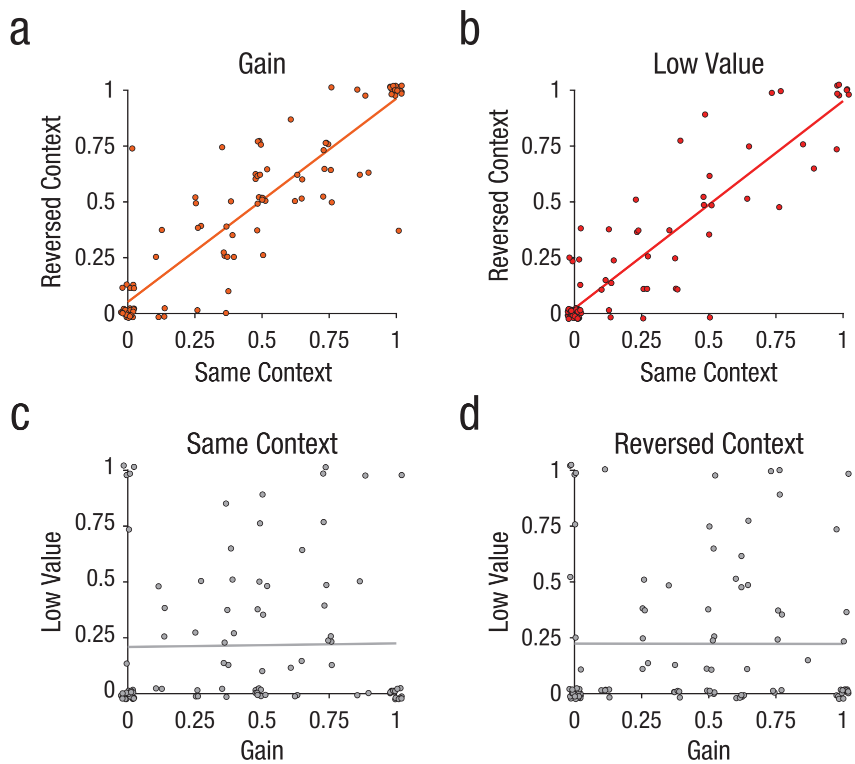

Fig. 6. Proportion of risky choices made in the test blocks in Experiment 2. The scatterplots in the upper row show risky choices in the reversed context as a function of risky choices in the same context, separately for (a) gain decisions and (b) low-value decisions. The scatterplots in the lower row show risky choices in the low-value decisions as a function of risky choices in the gain decisions, separately for the (c) same context and (d) reversed context. Dot locations are jittered to reduce overlap; lines represent correlation slopes.

(M = 36.0%, 95% CI = [33.2%, 38.8%]), $t(97) = 0.28$, $p = .78$, $d = 0.03$. These data support the notion that the encoding context is more important than the retrieval context in determining later choice.

In an additional exploratory analysis, we sought to solidify the argument for and against the encoding and retrieval hypotheses, respectively. Here, we tested whether risky choices in different conditions of the test blocks were independent. The encoding hypothesis predicts that the proportion of risky choices for gain and low-value decisions should be highly correlated between the same and reversed contexts because the choices should be invariant to test context. In addition, the encoding hypothesis predicts low correlations between risky choices for gain and low-value option pairs within each test context, as these would have been encountered independently in training. In contrast, the retrieval hypothesis predicts the opposite: low correlations for each option pair across test contexts but high correlations between the gain and low-value decisions within a context.

Figure 6 shows how these results strongly support the encoding hypothesis: Correlations were very strong when we compared the proportion of risky choices made for the gain decisions in the same or reversed test context, $r(97) = .901$, $p < .001$, and similarly high for the low-value decisions, $r(97) = .920$, $p < .001$. In contrast, correlations between risky choices for gain and low-value decisions within each context were very low, suggesting that these decisions were independent of each other despite having identical outcome values—same context: $r(97) = .014$, $p = .89$; reversed context: $r(97) = .002$, $p = .99$.

Memory tests. Figure 7 shows that the results of the memory tests were similar to those seen in Experiment 1, with context-dependent overweighting of the extreme-loss and low-value outcomes. In the first-outcome-reported test, for both the loss and low-value options, participants were significantly more likely to report the worse value (−30 and +10, respectively)—loss: $\chi^2(1, N = 71) = 8.80$, $p = .003$; low value: $\chi^2(1, N = 76) = 23.21$, $p < .001$. Differences in reporting of outcomes were not significant for the risky gain option, $\chi^2(1, N = 76) = 1.90$, $p = .17$, or the risky high-value option, $\chi^2(1, N = 78) = 2.51$, $p = .11$. The frequency-judgment test also showed a

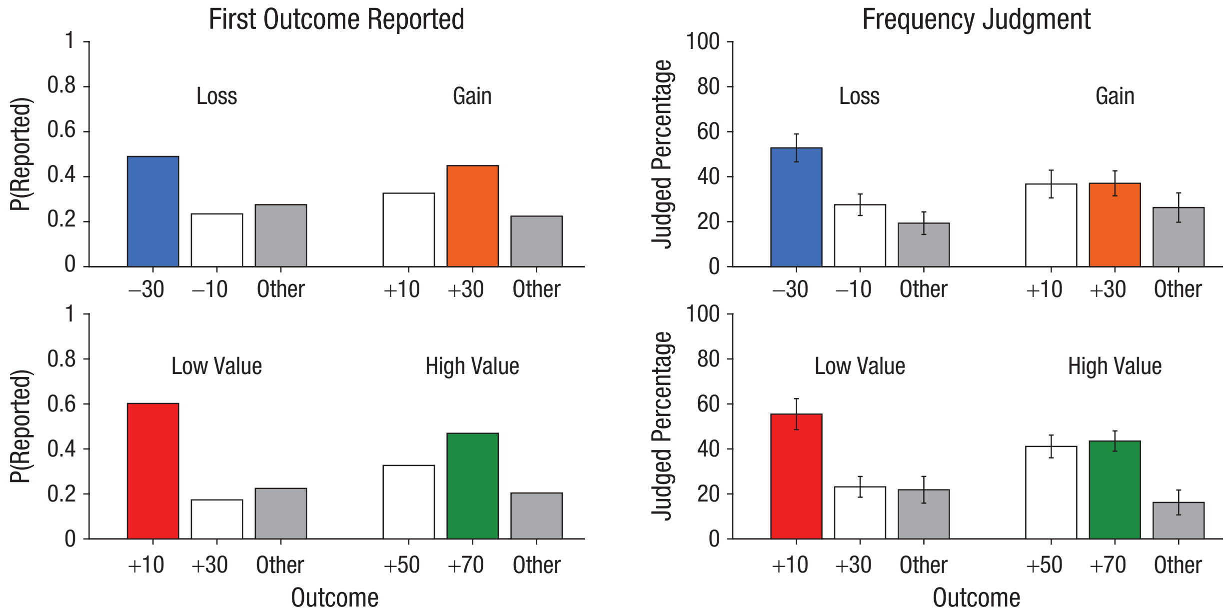

Fig. 7. Results of the two memory tests in Experiment 2. The graphs on the left show the proportion of participants who reported each outcome as the first that came to mind, and the graphs on the right show participants' judgments of the frequency of each outcome. For both tests, results are shown separately for the gain/loss context (top) and the high/low context (bottom). Colored bars indicate local extreme outcomes, and white bars indicate nonextreme outcomes. The color code matches the conditions in previous figures (blue = loss, orange = gain, red = low, green = high). Error bars represent 95% confidence intervals.

context-dependent bias in which people reported higher percentages for the worse outcome for the loss and low-value options—loss: $t(93) = 5.10$, $p < .001$, $d = 0.526$; low value: $t(90) = 6.19$, $p < .001$, $d = 0.65$—but no reliable difference in judged percentage for the outcomes of the gain and high-value options—gain: $t(92) = 0.07$, $p = .948$, $d = 0.01$; high value: $t(91) = 0.58$, $p = .56$, $d = 0.06$. Thus, in both measures, the worst outcome in each context was particularly salient in memory. The context dependence of this salience is highlighted by the +10 outcome, which was reported more often, $\chi^2(1, N = 91) = 8.01$, $p = .005$, and judged as having a higher frequency, $t(88) = 4.07$, $p < .001$, $d = 0.43$, in the high/low context (in which it was the worst outcome) than in the gain/loss context (in which it was an intermediate outcome).

Supplemental experiments. Two additional experiments are reported in the Supplemental Material that address alternative explanations related to the necessary and sufficient conditions for creating distinct decision contexts (see Table S1 in the Supplemental Material). The results show that distinct background images are not necessary for establishing a local decision context, but temporal grouping of the choices is not sufficient to establish discrete contexts. The distinct visual cues from the choice stimuli, however, are sufficient—and may even be necessary—to distinguish the contexts (see Experiment S2 in the Supplemental Material). These distinct visual cues may also serve as retrieval cues for the decision context in which they were encoded. Together with Experiment 2's findings, these results clearly show that choice is determined by the decision context during encoding and not by the decision context at retrieval.

General Discussion

In two experiments, we demonstrated that people's risky choices are not stable, even within a single experimental session, but rather depend on the other outcomes experienced during the context of encoding. Risky choice was biased by the most extreme outcomes in a particular decision context rather than by the global context of the whole experiment, and people also remembered those outcomes more strongly. Even for identical decisions (between +20 and a 50/50 chance of +10 or +30), changes in context substantially shifted both risky choice (> 10% in Experiment 1 and > 20% in Experiment 2) and memory for extremes, even for the same participants within a single session. Moreover, when tested in the opposite context, people's choices were in line with the initial training context, suggesting that the context of encoding is critical for this memory-based choice.

These findings have theoretical implications for memory-based theories of experience-based decision-making (e.g., Shohamy & Daw, 2015; Weber & Johnson, 2006). For example, according to the decision-by-sampling theory (Stewart et al., 2006), the values of options presented for people to choose from are compared with a small sample in working memory; the sample comes both from other values in the immediate context and from values stored in long-term memory. Our results suggest that such samples would have to come from values presented in the encoding context rather than in the context at the time of choice. Thus, our results pose significant challenges for retrieval-based models of how memory affects choice but are more consistent with a recent reinforcement-learning model that assumes that the influence of context on value operates during the learning process (Spektor et al., 2019).

The current results show how unstable choices can be and add to the growing evidence that choices depend on properties of the decision context (e.g., Huber et al., 1982; Simonson & Tversky, 1992). An important open question is how to pull the various context effects into a single process model of risky choice. One possibility is inspired by recent reinforcement-learning models in which researchers have attempted to integrate aspects of episodic memory (e.g., Gershman & Daw, 2017). Exactly how to incorporate other context effects from the decision-making literature is not clear, but it may require real-time integration mechanisms as in decision-field theory or the drift-diffusion model (Ratcliff & McKoon, 2008; Roe et al., 2001). Our work, however, suggests how important it will be to incorporate context effects into a reliable model of people's decision-making when learning from experience.

Transparency

Action Editor: M. Natasha Rajah
Editor: Patricia J. Bauer
Author Contributions
C. R. Madan, M. L. Spetch, and E. A. Ludvig conceived and planned the experiments. M. L. Spetch and F. M. D. S. Machado collected the data. All the authors conducted analyses. C. R. Madan took the lead in writing the manuscript, with extensive input from M. L. Spetch and E. A. Ludvig. All the authors discussed the results, commented on the manuscript, and approved the final version for submission.
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
Funding
This research was funded by grants from the Alberta Gambling Research Institute (to M. L. Spetch, E. A. Ludvig, and C. R. Madan), a grant from the Natural Sciences and Engineering Research Council of Canada (to M. L. Spetch), and a Leverhulme Early Career Fellowship (held by A. Mason).
Open Practices
Stimuli and data for Experiments 1, 2, S1, and S2 are available on OSF at https://osf.io/3mbwu. The design, hypotheses, analysis, and expected choice results for Experiments 2 and S2 were preregistered on OSF at https://osf.io/kv458. Experiments 1 and S1 were not preregistered. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iD

Christopher R. Madan https://orcid.org/0000-0003-3228-6501

Acknowledgments

We would like to acknowledge Veronica Bergstrom for assistance with data collection and Yang Liu for help with task programming. Door images were extracted from the "Irish Doors" collection on www.fineartamerica.com with permission from curator Joe Bonita.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620977516

References

Anderson, J. R., & Bower, G. H. (1974). Interference in memory for multiple contexts. *Memory & Cognition*, *2*, 509–514. https://doi.org/10.3758/BF03196913

Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature Communications*, *9*, Article 4503. https://doi.org/10.1038/s41467-018-06781-2

Berger, J., Meredith, M., & Wheeler, S. C. (2008). Contextual priming: Where people vote affects how they vote. *Proceedings of the National Academy of Sciences, USA*, *105*, 8846–8849. https://doi.org/10.1073/pnas.0711988105

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, *8*, Article 15958. https://doi.org/10.1038/ncomms15958

Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect. *Organization Science*, *12*, 523–538.

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*, 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology*, *68*, 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539. https://doi.org/10.1111/j.0956-7976.2004.00715.x

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*, 517–523.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269–299. https://doi.org/10.1006/jmps.2001.1388

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*, 90–98.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*, *14*, 47–53. https://doi.org/10.1101/lm.365707

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–292. https://doi.org/10.2307/1914185

Konstantinidis, E., Taylor, R. T., & Newell, B. R. (2018). Magnitude and incentives: Revisiting the overweighting of extreme events in risky decisions from experience. *Psychonomic Bulletin & Review*, *25*(5), 1925–1933. https://doi.org/10.3758/s13423-017-1383-8

Kunar, M. A., Watson, D. G., Tsetsos, K., & Chater, N. (2017). The influence of attention on value integration. *Attention, Perception & Psychophysics*, *79*, 1615–1627. https://doi.org/10.3758/s13414-017-1340-7

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2014). Extreme outcomes sway risky decisions from experience. *Journal of Behavioral Decision Making*, *27*, 146–156. https://doi.org/10.1002/bdm.1792

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk seeking. *Journal of Experimental Psychology: General*, *144*, 24–29. https://doi.org/10.1037/xge0000046

Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description-experience gap in risky choice limited to rare events? *PLOS ONE*, *6*(6), Article e20262. https://doi.org/10.1371/journal.pone.0020262

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2014). Remembering the best and worst of times: Memories for extreme outcomes bias risky decisions. *Psychonomic Bulletin & Review*, *21*, 629–636. https://doi.org/10.3758/s13423-013-0542-9

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2017). The role of memory in distinguishing risky decisions from experience and description. *Quarterly Journal of Experimental Psychology*, *70*(10), 2048–2059. https://doi.org/10.1080/17470218.2016.1220608

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2019). Comparative inspiration: From puzzles with pigeons to novel discoveries with humans in risky choice. *Behavioural Processes*, *160*, 10–19. https://doi.org/10.1016/j.beproc.2018.12.009

Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, *145*, 548–558. https://doi.org/10.1037/xge0000158

North, A. C., Hargreaves, D. J., & McKendrick, J. (1997). In-store music affects product choice. *Nature*, *390*, Article 132. https://doi.org/10.1038/36484

Pryor, B., Mendez, J. M., & Herrick, R. (2014). Let's be fair: Do polling locations prime votes? *Journal of Political Sciences & Public Affairs*, *2*, Article 3. https://doi.org/10.4172/2332-0761.1000126

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence. *Current Opinion in Neurobiology*, *22*, 970–981. https://doi.org/10.1016/j.conb.2012.07.011

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*(2), 370–392. https://doi.org/10.1037/0033-295x.108.2.370

Shafir, S., Waite, T. A., & Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, *51*, 180–187. https://doi.org/10.1007/s00265-001-0420-8

Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*, *5*, 85–90. https://doi.org/10.1016/j.cobeha.2015.08.010

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*, 158–174.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*, 281–295. https://doi.org/10.2307/3172740

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, *6*, 342–353. https://doi.org/10.3758/BF03197465

Spektor, M. S., Gluth, S., Fontanesi, L., & Rieskamp, J. (2019). How similarity between choice options affects decisions from experience: The accentuation-of-differences model. *Psychological Review*, *126*(1), 52–88. https://doi.org/10.1037/rev0000122

Stark, S. M., Reagh, Z. M., Yassa, M. A., & Stark, C. E. (2018). What's in a context? Cautions, limitations, and potential paths forward. *Neuroscience Letters*, *680*, 77–87. https://doi.org/10.1016/j.neulet.2017.05.022

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26. https://doi.org/10.1016/j.cogpsych.2005.10.003

Stewart, N., Reimers, S., & Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, *61*(3), 687–705. https://doi.org/10.1287/mnsc.2013.1853

Usher, M., Tsetsos, K., Glickman, M., & Chater, N. (2019). Selective integration: An attentional theory of choice biases and adaptive choice. *Current Directions in Psychological Science*, *28*, 552–559. https://doi.org/10.1177/0963721419862277

Weber, E. U., & Johnson, E. J. (2006). Constructing preferences from memory. In S. Lichtenstein & P. Slovic (Eds.), *The construction of preference* (pp. 397–410). Cambridge University Press.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*, 140–176. https://doi.org/10.1037/bul0000115