**Introduction**

What we prefer to eat, which products we decide to buy or even whether we take the metro or the bus to get to work today are all decisions that rely, in part, upon knowledge of outcomes of previous decisions without explicit knowledge of the outcomes and probabilities of each action. In contrast to risky decision-making from description, in which the probabilities of different outcomes associated with choices are explicitly known to decision-makers (Kahneman & Tversky, 1979), risky decision-making on the basis of trial-and-error experience involves drawing upon knowledge of past outcomes to guide decision-making (Hertwig & Erev, 2009). Indeed, a large body of research in the last decade has sought to understand how people make decisions under uncertainty using their previous experiences as a guide (\*citations to be added\*).

At the same time, mnemonic processes are critical to this sort of decision-making as our choices must draw upon previously stored memories of outcomes associated with actions in order (Hertwig & Erev, 2009; Madan, Ludvig, & Spetch, 2013; Murty, FeldmanHall, Hunter, Phelps, & Davachi, 2016). Underlining the consequences of biased memory representations, in risky decision-making from experience, Madan and colleagues (2013) posited an extreme-outcome rule where extreme outcomes in risky choice are overweighted in memory (Talarico & Rubin, 2003) and consequently are given more weight when choosing from experience. In a decision-from-experience task where the expected values between the certain and risky action were equal, the extreme outcome was most frequently reported to be the first “to come to mind." This memory salience for the extreme outcome was also associated with more risk-taking in decisions from experience (Madan et al., 2013).

More specifically, episodic memory, a form of declarative memory specific to the moment and place it was acquired (Tulving, 1983, 2002), has previously been shown to be involved in decision-making. The hippocampus has been shown to associate rewards to similar but previously unseen items (Wimmer & Shohamy, 2012), suggesting that episodic memory enhances the associations between events and rewards. Consistent with these findings, episodic memory has been shown to play an adaptive role in decision-making (Murty et al., 2016; Duncan & Shohamy, 2016).

Intriguingly, a recent body of memory research suggests that the extent to which individual focuses on specific details of a past experience can alter behavior on subsequent tasks believed to be dependent on episodic memory, such as divergent thinking (Madore et al., 2015) and solving of hypothetical social problems (Madore and Schacter, 2014). The episodic specificity induction has been shown to increase the amount of recalled internal (episodic) details on the Autobiographical Interview, but not the amount of external (semantic) details recalled and imagined (Madore, Gaesser & Schacter, 2014). To this end, the episodic specificity induction procedure developed by these researchers allows affords direct experimental manipulation of an individual’s tendency to focus on specific details, taken in comparison to a control induction procedure.

Here we leverage the episodic specificity induction procedure to shed light upon the extreme-outcome effects observed by Madan et al. (2014) and probe the extent to which apparent risk preferences in decision-making depend on episodic memory. Since the hippocampus has been shown to spread positive value of rewards across associated memories (Wimmer & Shohamy, 2012), episodic memory might be a potential cause of the extreme-outcome effect. To investigate this, we designed an experiment where participants made risky choices from experience (following the procedure of Madan, Ludvig & Spetch, 2014) after having been exposed to the episodic specificity induction, an experimental procedure where participants are briefly trained in recollecting details of recent experiences.

In short, we hypothesized that the episodic specificity induction could potentiate the observed overweighting of extreme outcomes, in turn increasing apparent preference for the risky (as opposed to sure-thing) action. To further examine the effects of the episodic specificity induction upon learning, we fit a simple RL model that quantifies the extent to which an individual participant weighs positive versus negative prediction errors (PEs) in learning the values of the two actions.

**Methods**

In Experiment 1, participants were either briefly trained in recollecting details of a video or had to describe general information about the same videos. (Jing, Madore, & Schacter, 2016; Madore, Szpunar, Addis, & Schacter, 2016; Madore, Addis & Schacter, 2015; Madore et Schacter, 2014). Experiment 2 served as a control study, in which participants performed the gambling task without any prior video or induction.

*Experiment 1*

Experiment 1 was conducted as a within-subject design, with every participant undergoing both the control and episodic specificity induction procedures. Participants performed the episodic or control interview and then performed the gambling task. Due to apparent carryover effects between across the two sessions, below we report only the results from the first session of the experiment.

*Participants*

We collected data from 47 participants who were recruited through McGill’s classified ads system. This study was approved by McGill’s Research Ethics Office (REB). Five participants were excluded from the analysis for having insufficient levels of exploration; four of them chose the risky option 3 times or less during the first 30 trials, and one chose the risky option 7 times overall. One participant from the control condition was excluded for being approximately 4 standard deviations (estimated without the outlier) away from the mean overall level of risk-taking in the control group. Of the remaining 41 participants, 21 participants were randomly assigned to the Episodic condition, and 20 participants were randomly assigned to the Control condition. 22 participants were randomly assigned to the first-win condition, and 19 participants to the second-win condition. Participants were compensated $10 CAD for one hour, and received an average of $1.25 CAD, (*SD*= 0.069) for each of the two sessions. We administered the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) both at the beginning of the experimental session and at the end, and the Offer Self-Image Questionnaire (OSIQ; Patton & Noller, 1991) at the end of the experimental session.

*Episodic specificity induction*

The experimental procedure began with an episodic specificity or control (“general”) induction, following the procedure of Madore et al. (2014). The episodic specificity induction is an experimental manipulation that has been inspired by the Cognitive interview, which has been shown to enhance the number of accurate details eyewitnesses can recall about events (Memon, Meissner, & Fraser, 2010). The participants first watched a 4-minutes long videos of “Mr. Bean” and were told to pay close attention to it since questions would be asked afterward. These videos are similar to those used in the original episodic specificity induction insofar as they contain a sequence of actions between characters, with plenty of episodic details. These videos were previously shown to be effective for the episodic specificity induction in our lab (results not yet published).

At the end of the video, participants were interviewed and asked different questions depending on the memory condition (Episodic or Control), following past work (Jing, Madore, & Schacter, 2016; Madore, Szpunar, Addis, & Schacter, 2016; Madore, Addis & Schacter, 2015; Madore et Schacter, 2014). In the Episodic condition, participants were first asked to describe as many specific details as they could remember about the surroundings. They were then asked to do the same about the physical appearances of the participants in the scene. Finally, they were asked to describe the actions in the video in chronological and in as much detail as they could remember. In the Control (i.e., “General Impressions”) condition, participants were instructed to use adjectives to describe the setting/people/actions as well as general questions about the video to broad questions about the video (e.g. What adjectives would you use to describe the actions in the video?).

*Gambling task*

Immediately after the induction procedure, participants performed a gambling task based upon the gains task of Madan et al. (2014) and Ludvig et al. (2015). On each of 100 trials, participants chose between two doors which both yielded real-monetary rewards. One of the doors was safe and always yielded a reward of 1.25 cents, while the other door was designated as the “risky” door and had a 50% chance to give 2.5 cents and a 50% chance to yield nothing. After choosing a door, participants were shown the reward they received from that door. Participants were not told beforehand the possible outcomes associated with each door; they had to learn the task from their own experience. Participants performed 100 trials of the gambling task per session.

Previous studies using similar gambling tasks reveal that initial outcomes carry a disproportionately large influence on further decisions (Shteingart et al., 2013). In this experiment, each participant received an equal amount of good and bad outcomes when choosing the gamble for the first four times. Accordingly, we controlled the first four outcomes participants obtained when choosing the risky door: They received one of the following patterns: win-loss-win-loss (the “first-win” condition) or loss-win-loss-win (the “first-loss” condition). The first outcome manipulation allowed us to systematically evaluate the impact of the very first outcome on risk preferences while controlling for the overall probability of wins and losses in the first four trials.

*Memory Recall*

Immediately after the choice task, we assessed participants by asking them what was the first outcome that came to their mind when thinking about the risky door, following the procedure from Madan et al. (2014). Participants were shown each of the two doors, in random order, and asked what was the first outcome that occurred to their mind when seeing them. This allowed to see whether the experimental manipulations influence participants’ memories of the outcomes they received in the task, and if the memory of the outcomes guided behavior. Immediately following the recall task, participants were instructed to draw the two doors to the best they could remember with a paper and pencil and label the colors.

*Data Analysis*

To compare risk preferences across conditions, we computed the mean level of risk from trial 40 onward for each participant (Madan et al., 2014). Excluding early trials allowed us to compare decisions that were made after having sufficient prior experience with the task. Risk preferences across groups were compared by conducting ANOVAs upon proportions of risky choices. Learning effects were tested using mixed-effects logistic regressions with random intercepts and slopes for each participant. This was done using the lme4 package (Pinheiro & Bates, 2000) for the R programming language.

*Reinforcement Learning (RL) Model*

Following basic formulations of RL models (Gershman, 2015; Sutton & Barto, 1998), this model operates by developing and updating expected reward values for each option, *aj,* on each trial, *t.* these *Q-values* aredenoted here and elsewhere as *Q(aj, t)*. The Q-values for each option (in the present task there are two options) are used to determine the model’s probability for selecting each option via a softmax decision rule:

Here is an exploitation parameter that determines the degree to which the option with the highest Q-value is chosen. As  approaches infinity the highest valued option is chosen more often, and as approaches 0 all options are chosen equally often.

On each trial the option that is chosen () is updated for the next trial (*t*+1) based on a simple incremental updating rule:

In Equation 2 above and are learning rate parameters for positive and negative prediction errors, and *r*(*t*) is the reward received from the chosen option on trial *t*. As these learning rate parameters approach 1, greater weight is given to the most recent rewards in updating Q-values indicative of more active updating of Q-values on each trial, and as the learning rate parameters approach 0, recent rewards are given less weight.

Our model fitting procedure used the Nelder-Mead optimization algorithm to find parameter values that maximized the likelihood of participants’ choices given their previous rewards and choices. To avoid estimates at parameter range boundaries, we imposed a ‘pseudo-prior’ over parameters, which for the learning rates, took the form of a beta distribution with *a* = *b* = 2, and for the inverse temperature parameter ()*,*a gamma distribution with *k*=1 and *θ*=3.

*Experiment 2*

Experiment 2 was comprised of a single “control” condition to assess people’s baseline levels of risk-taking. In this condition, 24 participants completed the gambling task without any prior induction or video. One participant was excluded from the analysis for having insufficient levels of exploration and not choosing the risk option at all until trial. Of the remaining 23 participants, 12 were randomly assigned to the first-win condition and 11 were assigned to the second-win condition. Participants were paid $8 CAD for approximately 20 minutes of their time, plus a bonus averaging $1.242 CAD (*SD* = 0.076).. The gambling task and memory recall procedure was identical to that of Experiment 1 except that they were performed in the absence of an episodic or control impressions induction. The same data analysis and modeling procedure from Experiment 2 was used.

**Results**

*Choice Task Behavior*

*Experiment 1*

We first sought to determine if apparent risk preferences differed across memory conditions and the first risky outcome. Upon examining the mean level of risky choices for each participant after trial 40 (Figure 2A), we found that risk-taking in the episodic condition (*M* = 0.485, *SD*=0.179) was significantly higher than in the control condition (M = 0.304, SD = 0.123; F(1, 39) = 14, p = 0.00058).

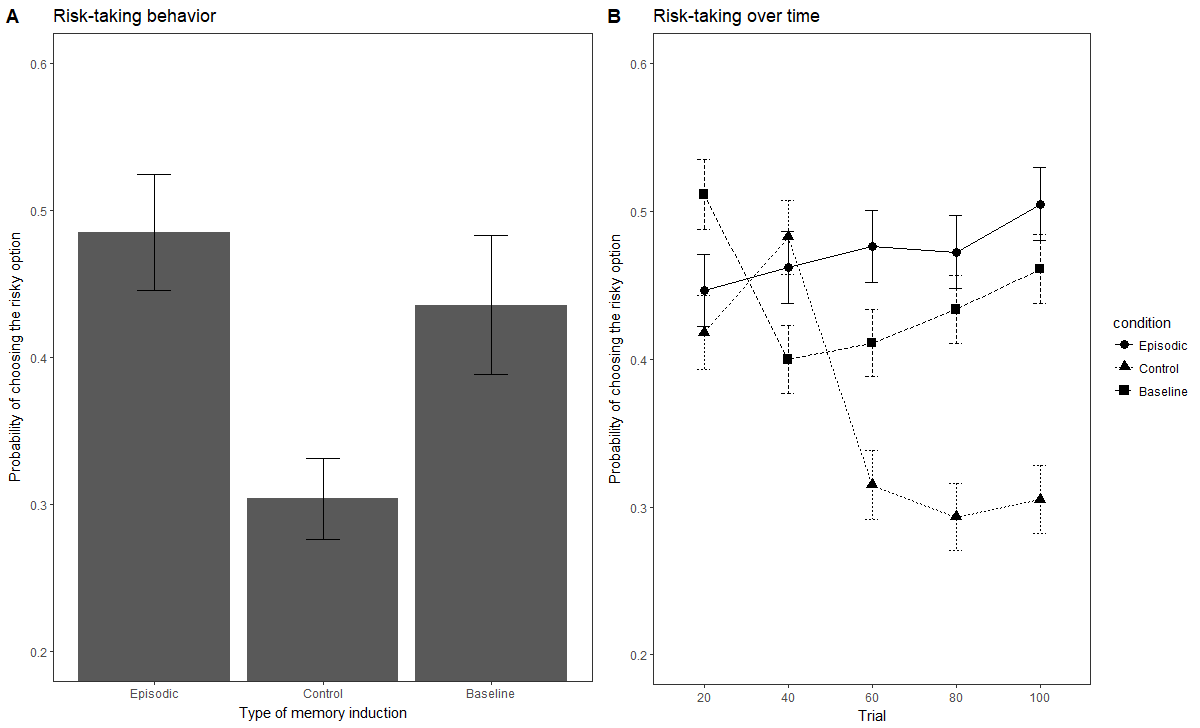


Figure 2: Figure 2A represents the proportion of risky choices for the three induction conditions (Episodic, Control and Baseline) from trial 40 to 100. Figure 2B represents changes in risky preferences over time in the three conditions (Episodic, Control and Baseline)

Examining choice behavior across blocks, we found that risk-taking developed differently over time in the episodic specificity and control conditions (Figure 2B). A mixed-effects logistics regression revealed that participants in the control condition became significantly more risk-averse over time than participants in the episodic condition (condition X trial interaction; *β* = 1.25, SE = 0.349, *p*= 0.000346). Thus, the two groups exhibited apparent differences in their time courses of apparent risk preference. More precisely, risk-taking tended to decrease over time in the control condition (*β* = -1, SE = 0.253, p = 0.0000767) but did not significantly change over time in the episodic condition (*β =* 0.222, SE = 0.241, p = 0.36).

When asked which outcome first comes to their mind, participants in the episodic condition were significantly more likely to report the positive outcome, χ2(1, N = 21) = 3.86, p= 0.0495 (see figure 4). This was not the case for the control condition χ2 (1, N = 20) = 0.2, p= 0.655. However, the episodic condition did not significantly report the positive outcome more than the control condition χ2(1, N = 21) = 1.96, p = 0.16. The first outcome that came to mind was not significantly correlated with risky behavior across the episodic and control condition F(1,39) = 0.881, p = 0.35.

*Experiment 2*

We replicated every analysis conducted previously to compare the baseline condition with both the episodic and control conditions. Upon examining the mean level of risky choices for each participant after trial 40 (Figure 2A), we found that risk-taking in the episodic condition (*M* = 0.485, *SD*=0.179) was not significantly different than in the baseline condition (M = 0.435, SD = 0.226; F(1,42) = 0.64, p = 0.43) and that risk-taking in the baseline condition was significantly higher than in the control condition (M = 0.304, SD = 0.123; F(1, 41) = 5.37, p = 0.0255) A mixed model effects logistics regression (see Figure 2B) revealed that learning of risk preferences over time (condition X trial interaction) in the episodic and baseline condition were not significantly different from each other(*β* = 0.41, SE = 0.531, p = 0.42). This interaction was not significant either between the control and baseline conditions (*β* = 0.86, SE = 0.53, p = 0.11).

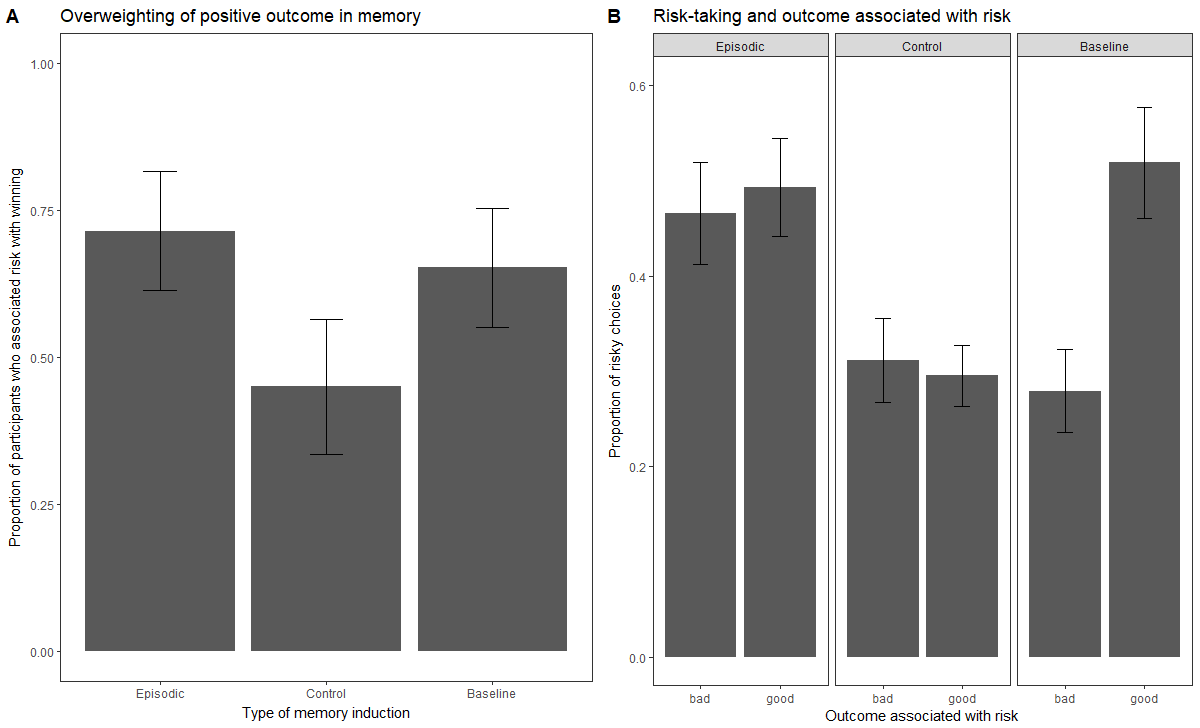
Participants in the baseline condition were not significantly more likely to report the positive outcome as the first one to come to mind χ2(1, N = 23) = 2.13, p= 0.144 (see Figure 3A). Within the baseline condition, participants with positive recall were more likely to choose the risky option (F(1,21) = 7.64, p = 0.011; see figure 3B). 

Figure 3: Figure 3A represents the first outcome that comes to mind in the episodic, control and baseline conditions. Figure 3B represents risk-taking as a function of the first outcome that comes to mind when thinking of the risk option, in different groups.

*Effects of Episodic Specificity Manipulation on RL Model Parameters*

The best-fitting RL model parameter estimates are reported in Table 1. Considering the entire sample, we found a significant main effect of PE valence (positive versus negative PEs) such that negative PE learning rates were significantly larger than positive PE learning rates (mixed-effects regression *β* = -0.15, *SE=*0.05, *p* = 0.026). Indeed, this observation corroborates previous observations of a ‘negativity bias’—a tendency to weigh negative PEs more strongly than positive PEs— in RL updating in similar tasks (Christakou et al., 2013; Gershman, 2015; Niv, Edlund, Dayan, & O’Doherty, 2012).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Condition** | **Learning Rate (Positive PE)** | **Learning Rate (Negative PE)** | **Inverse Temperature** | **Log Likelihood** |
| *Episodic* | 0.385 | 0.395 | 2.776 | -60.371 |
| *Control* | 0.242 | 0.401 | 3.759 | -56.223 |
| *Baseline (Experiment 2)* | 0.309 | 0.357 | 2.847 | -61.283 |

Table 1: Positive and negative learning rates for the episodic, control and baseline conditions.

However, examining the Episodic and Control conditions separately, we found that positive and negative PE learning rates were less asymmetric in the Episodic Specificity condition (Figure 4). In other words, participants who underwent the Episodic specificity induction appeared to weigh positive and negative more equally than participants who underwent the Control induction, who exhibited the typical negativity bias in learning rates (condition × PE type interaction *β* = .16, *SE=*0.06, *p* = 0.017).

../../../Downloads/pe_lr_v3.pdf

Figure 4: Best-fitting learning rate parameters for positive prediction errors and negative prediction errors, by condition. Error bars depict standard error of the mean.

**Discussion**

In experiment 1, we have found that the episodic specificity induction reduces risk-taking compared to a control induction (see figure 2A). This difference grows over time: While the episodic condition seems to be relatively consistent in risk-preferences over time, people in the control condition tend to progressively avoid risk more after making choices repeatedly (see figure 2B). The positive outcome was also more salient for the episodic condition (see figure 3A), even though it did not serve as a significant predictor of risky decisions (see figure 3B).

Risk-taking did not significantly change over time in the original study from Madan, Ludvig and Spetch (2013). Because of this, we wondered whether the difference between the two groups might have been due not to the episodic specificity induction but to the control induction (which still probes general impressions). In experiment 2, we found that a baseline induction with no induction lead to more risk-taking than the control induction (see figure 2A). This suggests that the differences between both condition of experiment 1 might have been due not to the episodic specificity induction enhancing risk-taking, but to the control condition lowering risk-taking.

Why would this happen? When learning a task purely from experience, the only information we have are previous outcomes and we are forced to rely on our memory of these outcomes. One possibility is that learning choices from experience already requires and induces episodic memory, and the episodic specificity induction could not enhance the use of episodic memory beyond that point. However, even though the control induction was meant as a control, it requires participants to recall information in a general manner. It is possible that doing so dampens the normal use of episodic retrieval or enhances semantic retrieval in learning from experience. This would be coherent with results by Madan et al. (2013), which were closer to the baseline and the episodic conditions than to the control condition.

Madore, Gaesser and Schacter (2014) have previously addressed this issue by showing that the episodic specificity induction increases the number of internal (episodic) details recalled but not the number of external (semantic) details even in a second control condition where participants had to solve mathematical problems. However, this second control group had a small sample size (N = 12) and there was no explicit comparison between the two control conditions to see whether these were significantly different from each other. The task we used measured something quite different (learning from experience) and it is also possible that the control induction interacted with our task but not theirs.

The mixed-model regression approach assumes learning to be linear. One limitation of this study is that linearity of risk-taking over time does not seem to hold for the control and baseline conditions (see figure 2B). Breaking this assumption might have prevented us to detect difference in learning across groups.

Regarding recall, we could only replicate overweighting of the positive outcome in memory for the episodic condition but not for the baseline condition (see figure 3A). Similarly, the tendency for participants who reported the positive outcome to be more risk-seeking could only be observed in the baseline condition (see figure 3B).