Title
Humans use episodic memory to access features of past experience for flexible decision making

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Nicholas, Jonathan
Mattar, Marcelo G

Publication Date
2024

Peer reviewed

Humans use episodic memory to access features of past experience for flexible decision making

Jonathan Nicholas
New York University
Department of Psychology
jonathan.nicholas@nyu.edu

Marcelo G. Mattar
New York University
Department of Psychology
marcelo.mattar@nyu.edu

Abstract

Our choices often require us to prioritize some features of our rich sensory experience over others. Past work suggests that humans solve this problem by focusing on relevant information while discarding that which is irrelevant. Yet learning which features to prioritize requires extensive experience. Moreover, features that are irrelevant now may become relevant in the future. One way to address these issues is by sampling individual richly encoded experiences from episodic memory. Here we hypothesize that episodic memory is used to guide decisions based on multiple features of past events. We test this hypothesis using an experiment in which participants made choices about the value of features that were present in multiple past experiences. We find evidence suggesting that participants used episodic memories to flexibly access features of past events during decision making. Overall, these results suggest that episodic memory promotes adaptive decisions when knowledge of multiple features is necessary.

Keywords: decision making; episodic memory; reinforcement learning

Introduction

In many daily tasks only a few features of our experience are relevant for the decisions we make, necessitating that some features be prioritized over others. For example, deciding where to go for lunch may generally depend more on a restaurant's quality of food rather than whether a particular colleague has joined you there in the past. Previous work has argued that people tend to discard irrelevant features when making these types of choices (Leong, Radulescu, Daniel, DeWoskin, & Niv, 2017; Niv et al., 2015; Wilson & Niv, 2012). Yet this approach can be problematic should features deemed previously irrelevant become relevant in the future– if your colleague proposes to have lunch at the restaurant you visited together last week but all you remember is the quality of the food, you will likely be dining alone. How do we preserve the ability to make flexible decisions based on many features of past events?

One way to solve this problem is by retaining information about all previously encountered features to reference whatever is needed for a choice in the moment. Formalizing such a strategy is possible using the computational framework of reinforcement learning (RL), which delineates how people gradually learn to estimate the value of choice options over many past experiences (Sutton & Barto, 2018). This framework provides a straightforward mechanism for summarizing the past without maintaining a memory of every individual experience. In this way, RL agents can learn to choose between options with multiple features by tracking the value of each feature separately. Yet this strategy does not scale; as the number of features to be learned grows, so do computational complexity and memory requirements. This well-known issue, typically called the "curse of dimensionality" (Sutton & Barto, 2018), limits the applicability of RL algorithms to decisions in the real-world.

This issue can be mitigated by allowing an agent to identify a small subset of relevant features to attend to while foregoing new information about the rest (Wilson & Niv, 2012). Taking an approach like this works well under circumstances in which it is possible to infer a feature's relevance, and it is likely that people deploy such a strategy in these kinds of environments (Leong et al., 2017; Niv et al., 2015). But augmenting RL with attention enforces rigid learning, ultimately harming future choices that may depend on features that were initially ignored. Separately, standard RL algorithms are most successful when experiences are certain to be repeated. Yet due to the sheer richness of our sensory input, actual experiences are unlikely to be encountered more than a single time.

Instead, another way to learn information about many features is to try and maintain the "raw data" of experienced events (Gershman & Daw, 2017; Lengyel & Dayan, 2008). Humans have developed a fast and dedicated memory system for precisely this purpose: episodic memory. Episodic memory has two primary properties that, together, distinguish it from other forms of memory: i) the ability to store individual events experienced in one-shot and ii) the ability to store the many spatial and temporal details of these events (Tulving, 1972). Importantly, these properties are each ideally suited to fill the gaps left by traditional RL approaches.

Although it has long been suggested that episodic memory may provide a number of advantages for adaptive behavior (Anderson & Milson, 1989; Bartlett, 1932; Schacter & Addis, 2007), it has only recently been found that episodic memory is frequently relied upon during decision making. In particular, past work has focused on the one-shot nature of episodic memory, finding that it allows humans to reference single past events to support fast and accurate decisions (Bornstein & Norman, 2017; Bornstein, Khaw, Shohamy, & Daw, 2017; Duncan, Semmler, & Shohamy, 2019; Mason, Madan, Simonsen, Spetch, & Ludvig, 2020; Nicholas, Daw, & Shohamy, 2022; Plonsky, Teodorescu, & Erev, 2015). De-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society*.

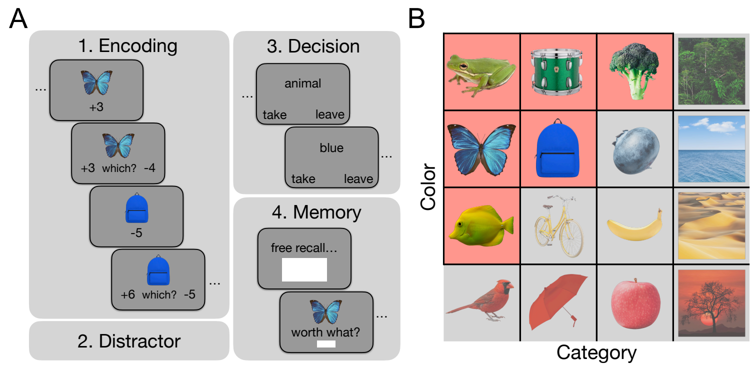

Figure 1: Task Design. A) The four phases completed by participants in each round of the experiment. B) The full set of images shown to participants. Six images were sampled to be shown in each round. Here, an example six images are highlighted in red. Images were sampled such that a differing number (between 1 and 3) were required to compute the value of each feature.

spite this progress, no studies have investigated whether the ability of episodic memory to encode experiential details provides its own advantages for choice.

By storing rich details of experienced events, episodic memory should theoretically provide substantial benefits over traditional learning algorithms. Specifically, episodic memory may grant us the ability to make flexible decisions about stimuli with multiple feature dimensions, such as those that are commonly encountered outside of the laboratory. Rather than focusing only on some currently prioritized features when encoding an event, an agent using episodic memory can instead encode the full event with all features, deferring feature selection to a later time, such as when a choice is required. This proposal is based on previous suggestions that an attentional filter may be applied to multidimensional stimuli at choice time rather than at the time of encoding (Dayan, Kakade, & Montague, 2000; Gershman & Daw, 2017). The flexibility afforded by this approach may be one reason for our ability to remember so many details of the past, but this idea has yet to be tested experimentally.

Here we ask whether episodic memory is indeed used to guide decisions based on multiple features of past events. We test this hypothesis using a novel behavioral experiment in which participants were required to make choices about features that were present in multiple past experiences. Critically, participants were free to adopt either an RL-style strategy or an episodic memory-based strategy when completing the experiment. After first validating this task across two independent samples, we find evidence suggesting that participants tend to reference episodes during these types of decisions. Overall, our results suggest that humans use their episodic memory to access features of past experience during decision making.

Methods

Participants

We recruited undergraduate students from the New York University subject pool. Participants were compensated with course credit. 83 participants with normal or corrected-to-normal vision were recruited to participate in the primary experiment, and another 58 participants were recruited to participate in a replication sample. Participants were excluded if they indicated on a post-task questionnaire that they either i) wrote any information down during the task or ii) did not try their best or reported being distracted throughout the task. Participants were further excluded if they provided nonsense responses to this questionnaire. Based on these criteria, 16 subjects were excluded from the primary sample and 11 subjects were excluded from the replication sample, leading to final Ns of 67 and 47 in each sample, respectively.

Experimental Procedure

Participants completed a four-part experiment over the course of a single online session designed to measure whether people access individual episodes to make decisions based on multiple features of past experiences (**Figure 1A**). Completing all four parts (a ”round”) took approximately five minutes, and participants completed five rounds in total. Unless otherwise noted, all procedures were identical between the main and replication samples.

Stimuli We selected stimuli that varied across two features: color (red, yellow, blue, or green) and category (animal, object, food, scene). A total of sixteen possible items were used throughout the task, with a subset of six pseudo-randomly sampled to be used in each round. All items and an example subset (highlighted in red) are shown in **Figure 1B**.

Encoding Phase In the first part of a round, participants completed a task designed to allow them to encode individual items and their associated value (which we refer to throughout as an ”episode”). Each item was presented on the screen for 1 second, after which its value appeared alongside it for another 6 seconds. An item’s value was a pseudo-randomly sampled integer between -9 and 9 (excluding 0). Immediately after viewing the episode, participants completed an attention check consisting of the item alongside two options, either the value that was just shown or another randomly selected value. They had 3 seconds to respond. Each episode was viewed once, for a total of six trials per round.

Distractor Phase Immediately following the encoding task, participants completed a 90 second distractor task to prevent active rehearsal of the episodes. This distractor consisted of a 2-back working memory task in which participants were shown one of several letters in sequence. Participants were asked to identify whether the current letter matched the one presented two steps earlier.

Decision Phase Immediately following the distractor task, participants then made up to six decisions based on the features of each item (all six possible decisions were made in the main sample, while only a subset of three out of six decisions were made in the replication sample.). Each decision consisted of an offer in which a single feature (e.g. ”animal”) was displayed on the screen, and participants were asked to either take or leave this offer. Participants were informed that the value of each offer consisted of the sum of each episode that was described by the offer (e.g. the the value of the ”animal” offer would be the sum of all animals seen during encoding), and that they should take positive offers and leave negative offers. Participants had 7.5 seconds to make each decision.

Importantly, there are at least two possible strategies that can be used to make good decisions in this task. The first (which we refer to as an *incremental* strategy) is an RL-style strategy that relies on extracting useful information from episodes at encoding time (in this case, a running sum for each of the six features). This strategy does not require that traces of each episode be maintained following encoding because the value of each offer is effectively computed at encoding time. In contrast, the second (which we refer to as an *episodic* strategy) relies on carrying the memory of each episode through to the decision phase, and then using these memories to compute an offer's value on-the-fly. This strategy offers the advantage of being more flexible, as each episode can be used and re-used according to the demands of the present decision.

Memory Phase Finally, immediately after the decision task, we assessed participants' memory for the episodes in two ways. First, participants were asked to freely recall the items that they saw in each round. They were provided with six empty text boxes and were told to enter the items in the order in which they remembered them. Participants were further told to halt their recall and move on to the next task if they could no longer remember any items. Following the free recall portion, participants were shown each item and were asked to provide their memory for the value of each item.

Analytical Approach

Behavioral Models All data was analyzed with regression models estimated using hierarchical Bayesian inference such that group-level priors were used to regularize subject-level estimates. All predictors were specified as fixed effects alongside random slopes and intercepts that were allowed to vary across subjects. Logistic regression was used to model choice data, linear regression with a shifted log-normal response distribution was used to model reaction time data, and linear regression with a normal response distribution was used otherwise. The joint posterior was approximated using No-U-Turn Sampling as implemented in stan (Hoffman & Gelman, 2014). Four chains with 2000 samples (1000 discarded as burn-in) were run for a total of 4000 posterior samples per model. Chain convergence was determined by ensuring that the Gelman-Rubin statistic R was close to 1. Default weakly-informative priors implemented in the brms package were used for each regression model (Bürkner, 2017). For all models, fixed effects are reported in the text as the mean of each parameter's marginal posterior distribution alongside 95% credible intervals (CIs), which indicate where that percentage of the posterior density falls. Parameter values outside of this range are unlikely given the model, data, and priors. Thus, if the range of likely values does not include zero, we conclude that a meaningful effect was observed.

Model Comparison Where applicable, model fit was assessed by separating the data into 10-folds and performing a cross-validation procedure by leaving out N/10 subjects per fold, where N is the number of subjects in each sample. The expected log pointwise predictive density (ELPD) was then computed and used as a measure of out-of-sample predictive fit for each model. Higher ELPD values suggest better model fit, as they indicate a higher likelihood of accurately predicting new data.

Results

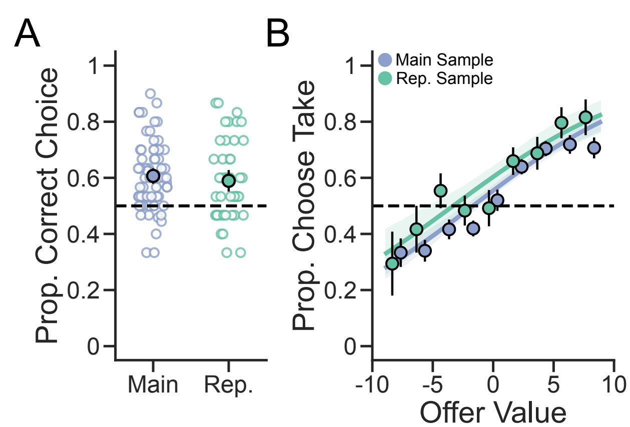

Figure 2: Overall choice performance. A) The average proportion of correct choices (choosing take when an offer is objectively positive or choosing leave when an offer is objectively negative) made by each subject across rounds (individual points) and at the group-level (filled points). B) The proportion of take choices that were made as a function of objective summed offer value. Points are binned value and the line represents a logistic regression fit. All error bars and bands represent 95% confidence intervals.

Overall choice performance

Before turning to answer our primary question, we first investigated whether participants learned to make effective decisions in the task. Participants' choices reflected their ability to compute the value of each offer by summing over individual experiences. At the group-level, they tended to correctly choose to take positive offers and leave negative offers (Main sample: β_0 = 0.61, 95% CI = [0.58, 0.64]; Replication sample: β_0 = 0.59, 95% CI = [0.55, 0.63]; **Figure 2A**). Furthermore, participants were sensitive to the summed value of each offer, as they were more likely to choose to take an offer that was more positive, and to leave an offer that was more negative (Main sample: β_{value} = 0.17, 95% CI = [0.13, 0.21]; Replication sample: β_{value} = 0.14, 95% CI = [0.08, 0.20]; **Figure 2B**). Overall, these results indicate that participants successfully performed the task.

Overall memory performance

Our next goal was to examine whether participants maintained memory traces for the individual episodes encoded in the first part of each round. To accomplish this, we analyzed participants' explicit responses on the memory phase, which consisted of free recall and value memory questions.

When asked to freely recall the six items they had seen in each round, participants in both samples tended to forget only one item on average and were highly accurate (Main Sample: β_0 = 0.78, 95% CI = [0.73, 0.82]; Replication Sample: β_0 = 0.82, 95% CI = [0.78, 0.86]; **Figure 3A**). To further characterize whether this recall adhered to typical properties of episodic memory, we asked whether items that were presented close together in time at encoding were more likely to be recalled after one another (commonly referred to as the contiguity effect (Kahana, 2020)). To do so, we computed the conditional response probability, which quantifies the likelihood of recalling a specific item based on its position in the initial encoding order relative to a previously recalled item. Replicating classic findings in episodic memory, we found that items immediately following or preceding a recalled item in the initial encoding order were more likely to be recalled next (**Figure 3B**).

Finally, we assessed participants' ability to accurately recall the value of each item presented in a round. Overall, there was a strong positive relationship between participants' remembered value of each item and the item's actual value (Main sample: $\beta_{recalledValue}$ = 0.51, 95% CI = [0.43, 0.58]; Replication sample: $\beta_{recalledValue}$ = 0.44, 95% CI = [0.34, 0.54]; **Figure 3C**).

Together, these results demonstrate that participants retained strong memories for each episode beyond the decision making phase. Critically, while these findings indicate that individual memories were available for potential recall at choice time, they do not provide direct evidence for any particular decision making strategy. Thus, we next sought to determine whether participants' episodic memories were recalled during choice.

Impact of episodes on choice

Having established that participants generally made good decisions in the task and exhibited strong memory for both the items and values of individual episodes, we were next positioned to investigate whether participants used either an incremental or episodic strategy for their decisions. Only an episodic strategy relies on information contained in single episodes. Accordingly, we reasoned that any impact of individual episodes on choice should provide evidence for this strategy. To test this idea, we used participants' responses during the memory phase to guide a re-analysis of their choice behavior.

We first asked whether participants' choices were better predicted by their reported memories rather than each offer's veridical value. To do so, we computed each participant's recalled value of an offer by summing over the remembered value of each offer-relevant item that was also recalled during the free recall phase. Participants' choices were sensitive to recalled offer value (Main sample: $\beta_{recalledValue}$ = 0.15, 95% CI = [0.11, 0.18]; Replication sample: $\beta_{recalledValue}$ = 0.12, 95% CI = [0.08, 0.17]; **Figure 4A**), and direct comparisons revealed that recalled offer value better predicted choices relative to the true value of an offer (Main sample: $ELPD_{true}$ = -875.12, SE = 13.13, $ELPD_{recalled}$ = -850.02, SE = 14.05; Replication sample: $ELPD_{true}$ = -871.18, SE = 13.07, $ELPD_{recalled}$ = -854.71, SE = 14.31). This result indicates that information contained in individual episodes, namely the identity and value of items, affected participants' choices.

To further examine which strategy was used during the task, we next considered the amount of time it took for participants to make their choices. We reasoned that recalling an individual episode should take time, and that referencing more episodes during choice should lead to longer decisions. Importantly, an incremental strategy makes no such prediction, as all choices can be based on a single summed value regardless of the number of episodes that were present at encoding time. Thus, the presence of a positive relationship between the number of memories recalled and response times during choice is evidence for an episodic strategy.

In addition to using this analysis to rule out the presence of either strategy, we tested two specific hypotheses about how episodes may be used to compute the value of each offer. The first was that participants may have recalled only memories that matched the feature presented during an offer (e.g. animal), while ignoring those that did not (e.g. recalling only animals). We tested this idea by examining whether the number of recalled memories that matched each offer predicted participants' response times. Our second hypothesis was that participants may have instead approached their decisions by determining whether each item they remembered in a round matched an offer, or not (e.g. recalling all episodes). We tested this second idea by examining whether participants took more time to respond to offers during rounds in which they recalled a greater number of items.

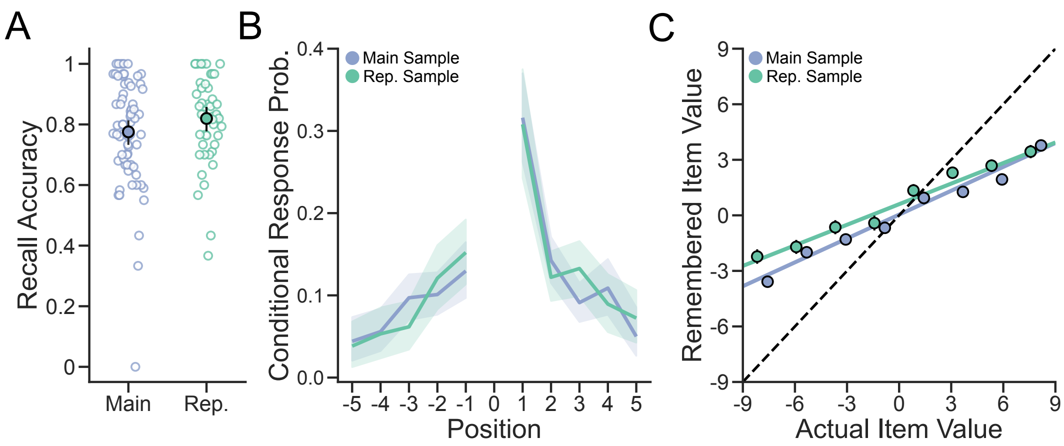

Figure 3: Overall memory performance. A) Average accuracy on the free recall task, computed as the proportion of correctly recalled items by each subject across rounds (individual points) and at the group-level (filled points). B) Lag-conditional reponse probability (CRP) curve demonstrating the classic contiguity effect in free recall data. The line represents the group-level average. C) The relationship between remembered item value and actual item value. Points are binned value and the line represents a linear regression fit. All error bars and bands represent 95% confidence intervals.

In line with this second possibility, participants took longer to make choices on rounds in which they remembered more items during the subsequent free recall phase (Main sample: $\beta_{nMemories}$ = 0.05, 95% CI = [0.02, 0.09]; Replication sample: $\beta_{nMemories}$ = 0.08, 95% CI = [0.03, 0.13]; **Figure 4B**). This finding provides evidence that an episodic strategy was indeed used for decisions throughout the task. Interestingly, there was no consistent effect of the number of matching memories on response time across our samples (Main sample: $\beta_{nMemories}$ = -0.05, 95% CI = [-0.08, -0.02]; Replication sample: $\beta_{nMemories}$ = 0.01, 95% CI = [-0.06, 0.10]).

Taken together, these findings support the hypothesis that people use episodic memories to make decisions that require knowledge of the value of multiple features of past events. Further, they suggest that such an episodic strategy relies upon referencing an overall pool of memories during each choice rather than isolating recall to only those memories that are most relevant for a choice.

Discussion

Research on learning about and choosing between stimuli with multiple features has focused primarily on how repeated experiences with these stimuli allow people to filter out information about task-irrelevant features (Leong et al., 2017; Niv et al., 2015; Wilson & Niv, 2012). This approach has successfully described choice behavior under scenarios in which stimuli are both highly familiar and when it is clear which features are most valuable in the present. Yet neither of these assumptions are commonly met outside of laboratory experiments, and past explanations fail to capture situations in which they are weakened.

Here we demonstrate that augmenting decision making with episodic memory can fill the gaps left by this work. We used a task that asked participants to compute the overall value of different features from the value of experiences seen only once, and found evidence suggesting that they referenced these episodes at the time a decision was made. In particular, we found that participants' subsequent memory for individual episodes explained multiple aspects of their choices, including the time it took to generate a response as well as the response that was ultimately given.

While these findings lend initial support to the idea that episodic memory is used for decisions requiring flexible knowledge about multiple stimulus features, the present work cannot speak to whether memories were directly accessed during choice. This is because our memory measures were collected immediately following choice rather than during the choice itself. We did not ask participants to directly recall items during their decisions because our aim was to assess the strategy participants relied upon without instructing them to use any strategy in particular, and we reasoned that this approach may bias them toward using their episodic memories for choice.

One way to circumvent this limitation would be to record neural activity during the decision phase. For example, past approaches using magnetoencephalography (MEG) have successfully decoded both the recall of individual episodes during standard memory tasks (Michelmann, Staresina, Bowman, & Hanslmayr, 2019; Wimmer, Liu, Vehar, Behrens, & Dolan, 2020) and the rapid replay of sequences of memories during decision making (Liu, Mattar, Behrens, Daw, & Dolan, 2021; McFadyen, Liu, & Dolan, 2023; Wimmer, Liu, McNamee, & Dolan, 2023). One possible future direction would be to similarly attempt to decode memory access in the present task. In addition to providing direct evidence for the recall of individual memories during choice, taking such an approach may provide multiple insights into the ways in which value is computed from memories. For example, the

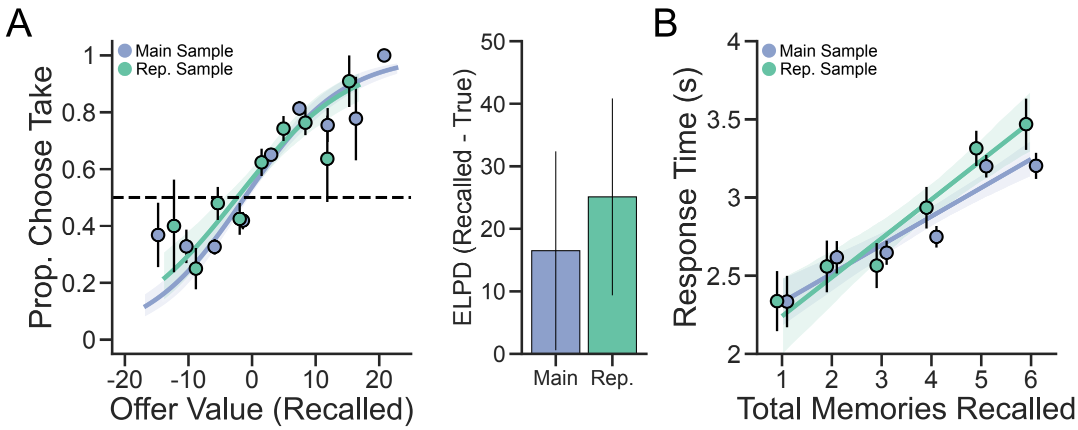

Figure 4: Impact of episodes on choice. A) Left: The proportion of take choices that were made as a function of summed recalled offer value. Points are binned value and the line represents a logistic regression fit. Right: The difference in expected log pointwise predictive density (ELPD) from choices predicted by either an offer's true value or sum of the value of offer-specific items recalled by participants. Error bars represent standard error around ELPD estimates. B) Response time during choices as a function of the total number of memories that were recalled in each round. Points are the group-average response time for different numbers of recalled memories and the line represents a linear regression fit. All error bars and bands represent 95% confidence intervals.

present findings suggest that participants likely recalled all possible memories for each choice, and participants also recalled items according to the temporal context in which they were encoded. A neural decoding study could directly test hypotheses about both the number and order of memories that are recalled during decision making.

Another limitation of our design was that, at encoding time, participants were likely aware of the features that would be needed for future choices. This type of information is rarely available in the real-world. While our results demonstrate that episodic memory is used even under these circumstances, we expect that episodes should provide even more benefits when this information is unknown. We plan to explore this possibility in future studies.

Separately, other work has suggested that people tend to trade off between different systems for learning and memory depending on which is most dependable for the task at hand (Daw, Niv, & Dayan, 2005; Lee, Shimojo, & O'Doherty, 2014). A recent study in this vein found that episodic memory is recruited for decisions when incrementally constructed estimates are uncertain or difficult to track, and that individuals with worse episodic memory tend to rely more on incremental learning instead (Nicholas et al., 2022). Based on this work, it seems likely that, while episodic memory was used in the present task, there may exist situations in which incremental strategies are more dominant. Another potential direction may be to alter the relative ease in which either strategy can be used. Here we kept equivalent the number of individual values that must be tracked at encoding time across both strategies (each round had either the value of six possible episodes to remember or of six possible features to sum over). Yet these quantities could be altered relative to one another in future versions, such that increasing the number of memories may push people toward relying upon an incremental strategy at encoding, whereas increasing the number of features may instead give preference to episodic memory. Regardless, here we found that episodic memory dominated when the amount of information to be learned was matched across strategies.

Finally, the present work connects at least two established but largely separate literatures on memory and choice. First, a number of studies focused on decision making have explored the ways in which individual experiences may be recalled for choice (Biele, Erev, & Ert, 2009; Plonsky et al., 2015). This work, typically grouped under the heading "decision by sampling", proposes that decision variables may be constructed by drawing samples from memory, and explains a number of ways in which peoples' choice behavior differs when information is learned from experience rather than instructed descriptions (Hertwig & Erev, 2009). Second, other work has proposed that episodic memory plays a critical role in our ability to infer new information about the world by allowing the formation of new links between past experiences (Schlichting & Preston, 2017; Whittington et al., 2020; Biderman, Bakkour, & Shohamy, 2020). Critical to this proposed role is episodic memory's ability to store multiple features of experience, because each feature provides another opportunity to relate past events with one another. Here we connect these ideas by suggesting that features of episodes may allow for the formation of new decision variables on-the-fly, when they are needed for a choice.

In conclusion, our findings support the idea that one function of episodic memory is to allow adaptive decisions to be reached when knowledge of the value of multiple features is necessary.

Acknowledgments

The authors thank the two anonymous reviewers and members of the Mattar Lab for their helpful comments and feedback on the project.

References

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703–719. doi: 10.1037/0033-295X.96.4.703

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New York, NY, US: Cambridge University Press.

Biderman, N., Bakkour, A., & Shohamy, D. (2020, July). What Are Memories For? The Hippocampus Bridges Past Experience with Future Decisions. *Trends in Cognitive Sciences*, *24*(7), 542–556. doi: 10.1016/j.tics.2020.04.004

Biele, G., Erev, I., & Ert, E. (2009, June). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology*, *53*(3), 155–167. doi: 10.1016/j.jmp.2008.05.006

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

Bornstein, A. M., & Norman, K. A. (2017, June). Reinstated episodic context guides sampling-based decisions for reward. *Nature Neuroscience*, *20*(7), 997–1003. doi: 10.1038/nn.4573

Bürkner, P.-C. (2017, August). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1–28. doi: 10.18637/jss.v080.i01

Daw, N. D., Niv, Y., & Dayan, P. (2005, December). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. doi: 10.1038/nn1560

Dayan, P., Kakade, S., & Montague, P. R. (2000, November). Learning and selective attention. *Nature Neuroscience*, *3*(11), 1218–1223. doi: 10.1038/81504

Duncan, K., Semmler, A., & Shohamy, D. (2019, July). Modulating the Use of Multiple Memory Systems in Value-based Decisions with Contextual Novelty. *Journal of Cognitive Neuroscience*, 1–13. doi: 10.1162/jocn$_{a0}$1447

Gershman, S. J., & Daw, N. D. (2017, January). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. doi: 10.1146/annurev-psych-122414-033625

Hertwig, R., & Erev, I. (2009, December). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. doi: 10.1016/j.tics.2009.09.004

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. , 31.

Kahana, M. J. (2020, January). Computational Models of Memory Search. *Annual Review of Psychology*, *71*(1), 107–138. doi: 10.1146/annurev-psych-010418-103358

Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014, February). Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. *Neuron*, *81*(3), 687–699. doi: 10.1016/j.neuron.2013.11.028

Lengyel, M., & Dayan, P. (2008). Hippocampal Contributions to Control: The Third Way. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *Advances in Neural Information Processing Systems 20* (pp. 889–896). Curran Associates, Inc.

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017, January). Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. *Neuron*, *93*(2), 451–463. doi: 10.1016/j.neuron.2016.12.040

Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D., & Dolan, R. J. (2021, May). Experience replay is associated with efficient nonlocal learning. *Science*, *372*(6544), eabf1357. doi: 10.1126/science.abf1357

Mason, A., Madan, C., Simonsen, N., Spetch, M., & Ludvig, E. (2020, December). *Biased confabulation in risky choice.* PsyArXiv. doi: 10.31234/osf.io/vphgc

McFadyen, J., Liu, Y., & Dolan, R. J. (2023, April). Differential replay of reward and punishment paths predicts approach and avoidance. *Nature Neuroscience*, *26*(4), 627–637. doi: 10.1038/s41593-023-01287-7

Michelmann, S., Staresina, B. P., Bowman, H., & Hanslmayr, S. (2019, February). Speed of time-compressed forward replay flexibly changes in human episodic memory. *Nature Human Behaviour*, *3*(2), 143–154. doi: 10.1038/s41562-018-0491-4

Nicholas, J., Daw, N. D., & Shohamy, D. (2022, December). Uncertainty alters the balance between incremental learning and episodic memory. *eLife*, *11*, e81679. doi: 10.7554/eLife.81679

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015, May). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157. doi: 10.1523/JNEUROSCI.2978-14.2015

Plonsky, O., Teodorescu, K., & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological Review*, *122*(4), 621–647. doi: 10.1037/a0039413

Schacter, D. L., & Addis, D. R. (2007, May). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 773–786. doi: 10.1098/rstb.2007.2087

Schlichting, M. L., & Preston, A. R. (2017). The Hippocampus and Memory Integration: Building Knowledge to Navigate Future Decisions. In D. E. Hannula & M. C. Duff

(Eds.), *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* (pp. 405–437). Cham: Springer International Publishing. doi: 10.1007/978-3-319-50406-3$_1$3

Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. , 352.

Tulving, E. (1972). Episodic and semantic memory. In *Organization of memory* (pp. xiii, 423-xiii, 423). Oxford, England: Academic Press.

Whittington, J. C., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. (2020, November). The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell*, *183*(5), 1249-1263.e23. doi: 10.1016/j.cell.2020.10.024

Wilson, R. C., & Niv, Y. (2012, January). Inferring Relevance in a Changing World. *Frontiers in Human Neuroscience*, *5*, 189. doi: 10.3389/fnhum.2011.00189

Wimmer, G. E., Liu, Y., McNamee, D. C., & Dolan, R. J. (2023, February). Distinct replay signatures for prospective decision-making and memory preservation. *Proceedings of the National Academy of Sciences*, *120*(6), e2205211120. doi: 10.1073/pnas.2205211120

Wimmer, G. E., Liu, Y., Vehar, N., Behrens, T. E. J., & Dolan, R. J. (2020, August). Episodic memory retrieval success is associated with rapid replay of episode content. *Nature Neuroscience*, *23*(8), 1025–1033. doi: 10.1038/s41593-020-0649-z