Research

Decision uncertainty during hypothesis testing enhances memory accuracy for incidental information

Xinxu Shen,[1] Ian C. Ballard,[2] David V. Smith,[1] and Vishnu P. Murty[1]

[1]*Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122, USA;* [2]*Helen Wills Neuroscience Institute, University of California at Berkeley, Berkeley, California 94720, USA*

Humans actively seek information to reduce uncertainty, providing insight on how our decisions causally affect the world. While we know that episodic memories can help support future goal-oriented behaviors, little is known about how hypothesis testing during exploration influences episodic memory. To investigate this question, we designed a hypothesis testing paradigm, in which participants figured out rules to unlock treasure chests. Using this paradigm, we characterized how hypothesis testing during exploration influenced memory for the contents of the treasure chests. We found that there was an inverted U-shaped relationship between decision uncertainty and memory, such that memory was best when decision uncertainty was moderate. An exploratory analysis also showed that surprising outcomes lead to lower memory confidence independent of accuracy. These findings support a model in which moderate decision uncertainty during hypothesis testing enhances incidental information encoding.

[Supplemental material is available for this article.]

Individuals explore the uncertain world by seeking information and updating hypotheses garnered from prior experience. For example, when purchasing milk in a new grocery store, a customer will typically go to the back corner of the store with the hypothesis that milk is usually stored there. If a customer fails to find milk in the back corner of a new grocery store, they will continue looking for milk with another hypothesis of where milk might be. This type of exploration facilitates hypothesis testing to resolve uncertainty, but open questions remain as to how this type of hypothesis testing influences episodic memory. For example, does the uncertainty about the grocery store's layout affect the strength of this event in memory? We designed a novel hypothesis testing paradigm to uncover how active hypothesis testing influences learning and memory during exploration.

Prior research on exploration and memory suggested that active exploration enhances memory and several underlying mechanisms, such as elaborative encoding, metacognitive monitoring, and selective attention (Voss et al. 2011; Markant et al. 2016; Ruggeri et al. 2019), have been proposed. However, little is known about how hypothesis testing during exploration of complex environment influences memory. Niv et al. (2015) suggested that individuals test hypotheses about higher-order concepts to resolve uncertainty to explore in multidimensional environments, but this prior work did not investigate whether and how this type of hypothesis testing has a downstream influence on episodic memory. We propose that hypothesis testing during exploration requires an assessment of the relative uncertainty in the environment, and these estimates of decision uncertainty will have a downstream effect on episodic memory by influencing the internal motivational state during encoding. Notably, this type of decision uncertainty may facilitate interrogative motivational states in which memory is not only enhanced for the targets of goal pursuit, but rather generalized to multiple features of the environment (Murty and Adcock 2014). In this way, decision uncertainty during encoding could lead to better memory for incidental features of the environment. Prior research has explored how motivational states can influence incidental information for goal-irrelevant information in the context of reward (Murty and Adcock 2014) and curiosity (Gruber et al. 2014), but these mechanisms have yet to be explored in the context of hypothesis testing.

A growing body of research has studied intrinsic motivational states during information seeking, and shown an inverted U-shaped relationship with task difficulty during problem-solving. Specifically, individuals are most motivated to work on problems with moderate levels of difficulty because information uptake was the highest (Metcalfe and Kornell 2003). Thus, resolving moderate levels of uncertainty should lead to most information uptake and decrease in decision uncertainty in uncertain environments. Given that individuals are motivated to reduce prediction errors during exploration (Friston and Kiebel 2009), intermediate uncertainty should be most motivationally relevant. Prior research also suggested that an individual's motivational state is a significant determinant of memory (Murty and Adcock 2014), including incidental yet salient events; therefore, we hypothesize that moderate decision uncertainty during hypothesis testing would yield the greatest memory enhancements. However, surprise and decision-uncertainty often co-occur during hypothesis testing. Given that prior research has shown that surprise enhances memory in the context of reward (Rouhani et al. 2018; Jang et al. 2019; Rouhani and Niv 2021), we are also interested in testing how surprise influences memory in the context of hypothesis testing (see the Supplemental Material).

To capture how hypothesis testing influences episodic memory, we designed a novel hypothesis testing paradigm combined with reinforcement learning models, which allows us to investigate the mechanism of how hypothesis testing influences memory by estimating trial-by-trial variation in decision uncertainty. Furthermore, these same reinforcement learning models allow us to simultaneously assess surprise to contrast which is a better predictor of differences in memory encoding (see the Supplemental

Corresponding authors: vishnu.murty@temple.edu, david.v.smith@temple.edu
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.053458.121.

Material). We hypothesized that (1) individuals would seek information about higher-order concepts during our task, (2) memory would be enhanced during active hypothesis testing, and (3) enhanced memory during active hypothesis testing is driven by trials with moderate decision uncertainty.

Results

We tested participants on a novel hypothesis testing paradigm to examine how decision uncertainty influences memory in the context of hypothesis testing, based on prior research examining reinforcement learning in a multifeatured environment (Niv et al. 2015). In brief, participants completed a deterministic hypothesis testing task (Fig. 1A), where participants were instructed to choose one of three keys to unlock a treasure chest. Participants could learn the target feature by testing the efficacy of different keys. If participants selected the correct key, the chest would open revealing a trial-unique object. To ensure participants continue exploring different target features over the entire block, the target feature changed after participants responded correctly for four consecutive trials. In the control condition (Fig. 1C), participants completed forced-choice trials in which they were instructed to choose a golden key from three keys given (one golden key and two gray keys). Outcomes of the control task were yoked to the outcomes of the previous hypothesis testing task to control for factors other than participants' motivational state, such as overall trial number and trial spacing.

Feature, rather than item, updating better captures participants' hypothesis testing behavior

Prior research has suggested that participants could use at least two approaches to learn in a multidimensional environment: item updating and feature updating (Niv et al. 2015). During item updating, individuals attempt to learn the associations between specific keys and their outcomes. During feature updating, individuals attempt to learn the value of individual features of the key (e.g., handle color) and pick the key with the most predictive features. This strategy integrates across multiple experiences to identify the most reward-predictive features of the environment. In the instructions, participants were told to figure out correct keys by testing out different features; therefore, we hypothesize that the feature updating strategy would better predict participants' behaviors. Cross-validated model prediction showed that both strategies significantly predicted behavior (item updating RL: mean learning rate: 0.99, mean probability: 0.40, chance: 0.33; $t = 21.09$, $P < 0.001$; feature updating: mean learning rate: 0.36, mean probability: 0.53, chance: 0.33; $t = 43.02$, $P < 0.001$) (Fig. 2B). As predicted, model comparison using a "leave one subject out" cross-validation approach showed that feature updating outperformed item updating in predicting choice behavior ($t = 22.73$, $P < 0.01$).

We next examined trial-level performance to confirm that within a block, feature learning outperforms item updating. The four trials before a feature switch are an important test of our models because subjects have identified a strategy that can successfully identify the correct key for four consecutive trials. We found increasing predictive accuracy from four trials before a feature change to the feature changing trial for both item updating RL ($\hat{y}-4 = 0.38$, $\hat{y}-3 = 0.41$, $\hat{y}-2 = 0.50$, $\hat{y}-1 = 0.48$, $\hat{y}0 = 0.49$) and feature updating RL ($\hat{y}-4 = 0.44$, $\hat{y}-3 = 0.60$, $\hat{y}-2 = 0.73$, $\hat{y}-1 = 0.77$, $\hat{y}0 = 0.79$) (Fig. 2C). Again, predictive accuracy was higher for feature updating RL than for item updating RL ($t = 24.86$, $P < 0.01$), further indicating that feature updating strategies outperformed an item updating strategy in predicting choice behavior, which is consistent with previous findings on decision-making in multidimensional environments (Niv et al. 2015). For the winning feature updating model, we found that although predictive accuracy increased across four trials before a feature change, there was no significant learning across the entire task (Fig. 3). Moreover, there was small variance of learning rate across participants (mean = 0.36; SD = 0.006) and learning rate did not predict memory performance on the second days.

Influences of hypothesis testing on 24-h memory

Next, we compared memory performance during the hypothesis testing condition and the control condition. Participants showed significantly better memory for objects presented during the hypothesis testing condition than control condition ($\beta = 0.24$, SE = 0.10, $P = 0.02$) (Fig. 3A). There were no significant effects of counterbalance order ($\beta = -1.83$, SE = 0.30, $P = 0.54$) or a condition* counterbalance order interaction ($\beta = -0.24$, SE = 0.14, $P = 0.10$) (Fig. 3B). However, multiple features of encoding could be driving memory differences across tasks. Thus, in our next set of analyses we explored how specific mechanisms underlying hypothesis testing, such as decision uncertainty, influence memory encoding.

Decision uncertainty was modeled as the difference between the highest chosen probability and the second highest chosen probability from the feature updating reinforcement learning model for a given trial. Decision uncertainty decreased across the four trials before a feature change, where trials at feature change mostly constitute low decision uncertainty, fourth trial before a feature change mostly constitute high decision uncertainty and the second and the third trial before a feature change mostly constitute intermediate decision uncertainty (Figs. 4, 5). We found an inverted U-shaped relationship between decision uncertainty and memory accuracy ($\beta = -1.04$, SE = 0.48, $P = 0.03$), such that memory was enhanced at middle levels of decision uncertainty, compared with high and very low levels of decision uncertainty.

To control for the effects of factors such as the distribution of decision uncertainty, overall trial number in our task and to rule out spacing effect, we performed a control analysis in which the decision uncertainty estimates were applied to the yoked control condition. This control analysis did not reveal a linear ($\beta = -0.54$, SE = 0.51, $P = 0.29$) or quadratic ($\beta = 0.57$, SE = 0.49, $P = 0.25$) relationship between decision uncertainty and memory accuracy for the control condition. Moreover, the inverted U-shaped relationship between decision uncertainty and memory accuracy remained significant after direct comparison of the experiment condition and the control condition ($\beta = -0.32$, SE = 0.15, $P = 0.03$). More specifically, relative to memory for objects presented during the control task, memory was only enhanced at moderate decision uncertainty ($\beta = 0.19$, SE = 0.08, $P = 0.02$). Memory at high ($\beta = 0.04$, SE = 0.13, $P = 0.72$) and low decision ($\beta = -0.03$, SE = 0.12, $P = 0.77$) difficulty was not different from the control condition. In addition to this quadratic effect, there was also a linear relationship between decision uncertainty and memory accuracy ($\beta = 1.12$, SE = 0.50, $P = 0.03$), such that memory accuracy decreased as decision uncertainty increased. Model comparison suggested that the quadratic model was significantly better than the linear model ($AIC_{quadratic} = 2455.1$, $AIC_{linear} = 2457.6$, $P = 0.03$) (Fig. 6). The same analyses were run to examine how surprise influenced memory accuracy (see the Supplemental Material). Moreover, an exploratory analysis has been run to examine how decision uncertainty and surprise influenced memory confidence (see the Supplemental Material; Supplemental Fig. S1).

Discussion

To understand the influence of hypothesis testing on episodic memory, we tested participants on a novel hypothesis testing task. Using reinforcement learning models, we found participants formed and tested hypotheses about target features and

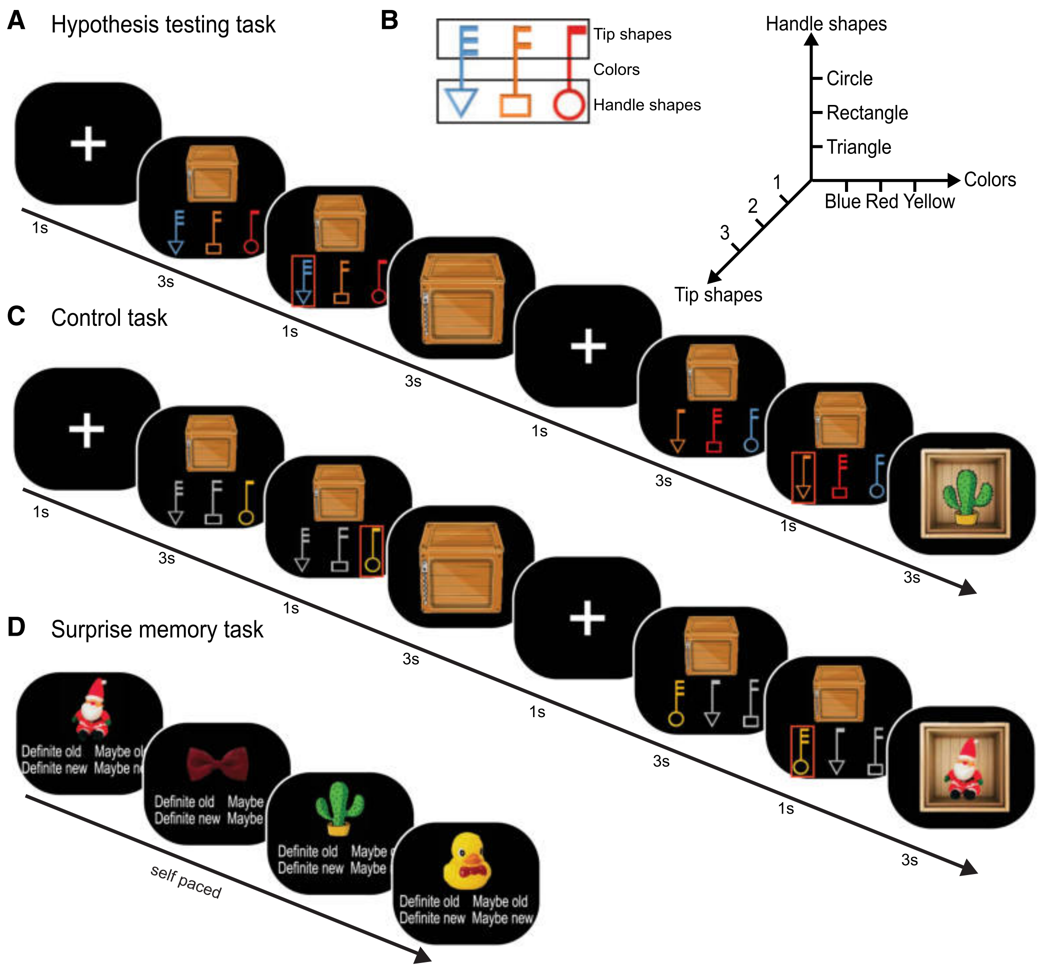

Figure 1. Overview of the task. (*A*) Hypothesis testing task. Participants were instructed to choose a key from three keys given, with the goal to open treasure chests. Choosing a key with target feature would open the treasure chest and a trial-unique object would be shown inside the treasure chest. Otherwise, the treasure chest stayed closed. (*B*) Three different dimensions for a key and three different features for a dimension. (*C*) Control task. Participants were instructed to choose the golden key from three keys given. Outcomes of the control task were yoked to outcomes of the hypothesis testing task. (*D*) Surprise memory task. Participants were presented with all the images from the hypothesis testing task and control task and the same number of new images. Participants need to indicate whether they saw the image during encoding.

generalized this learning to new stimuli. Furthermore, we found that hypothesis testing enhanced incidental memory for the targets of the tested hypothesis compared with a visuo–motor matched control condition. Critically, memory enhancements only emerged when individuals had moderate amounts of decision uncertainty, which theoretically represents when individuals should be most motivated to seek information. Together, these findings support a new model in which decision-uncertainty during hypothesis testing enhances memory encoding for incidental information encountered during exploration.

We found an inverted U-shaped relationship between decision uncertainty and memory, such that memory is enhanced at moderate decision uncertainty. Previous research suggested that individuals are most motivated to learn items that are neither too difficult nor too easy (Metcalfe and Kornell 2003). For easy items, learning is fast and information uptake is the largest initially, but with little subsequent increase. For items of intermediate difficulty, even though the initial information uptake is smaller, gains are more sustained throughout the task (Metcalfe and Kornell 2003). During hypothesis testing, we propose that decision uncertainty reflects trial difficulty and influences individuals' judgment of learning. Trials at moderate decision uncertainty are of intermediate difficulty. Given that individuals are motivated to reduce uncertainty during exploration (Friston and Kiebel 2009), trials at moderate decision uncertainty are most likely to trigger motivation to resolve uncertainty (Monosov 2020). Motivational states during encoding have previously been shown to enhance hippocampus-dependent memory for incidental events in the context of reward (Murty and Adcock 2014), and our findings extend these mechanisms to the domain of hypothesis testing. Our findings also dovetail well with a growing body of research on curiosity, which shows that individuals are most curious when there is intermediate decision uncertainty (Gottlieb et al. 2016), which may then enhance incidental encoding of the environment (Gruber et al. 2014; Gruber and Ranganath 2019; Murphy et al. 2021). However, we do not have a direct measure of motivational states and curiosity. Therefore, the interpretations could be speculative. Future studies are needed to directly test the underlying mechanisms.

Our hypothesis testing paradigm showed that individuals learn and encode high-order features during hypothesis testing. In turn, hypothesis testing enhances memory at moderate decision uncertainty. However, hypothesis testing in the real world is more complex because the outcomes not only depend on an individual's own actions, but also on uncertainty in the environment (Wilson

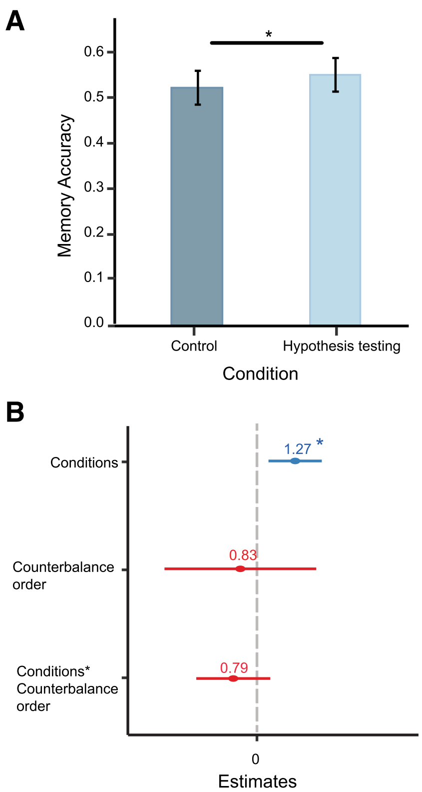

Figure 2. Hypothesis testing enhances memory encoding. (*A*) After 24-h of delay, memory was better for objects presented during the hypothesis testing task than the control task. Error bars are 95% confidence interval. (*B*) Model estimate from mixed effect model, where task condition and counterbalance order and the interaction were submitted as fixed effects and objects and subjects were submitted as random effects. Only the main effect of conditions was significant.

and Niv 2012; Choung et al. 2017). In the current study, we could not fully capture this aspect of uncertainty in the environment because we implemented a deterministic design. For example, in our design, if a subject picked a key and found the treasure chest did not open, the subject could be confident that the chosen key did not include a target feature, could rule out the three features of the chosen key, and test the other features. However, in many situations, we cannot attribute an outcome to a single cause or completely rule out options during hypothesis testing. Therefore, a probabilistic design, where an unopened treasure chest could be due to an incorrect key, but can also be due to the probabilistic nature of the context, is needed in the future to unpack how uncertainty influences memory.

In summary, our study uncovers the learning and memory mechanisms underlying hypothesis testing, suggesting that individuals learn information about higher-order features to update their beliefs about an uncertain environment, which then enhances memory encoding for incidental information. Our study also goes beyond previous research on decision-making and memory and suggests that the underlying mechanism of decision-making enhanced memory is moderate decision uncertainty.

Materials and Methods

Participants

Thirty-five participants (ages 18–26, seven males) were recruited from Temple University via SONA. This sample size was calculated using a power analysis with 0.8 power and $d = 0.5$ effect size with an alpha of 0.05 (Murty et al. 2015). Temple University's Institutional Review Board approved study materials and procedures. All participants provided informed consent and were compensated for their time with course credits. Four participants were excluded because they did not complete more than one rule in the hypothesis testing condition (see below for details).

Stimuli

Our task involved participants selecting keys to unlock a treasure chest, which revealed an object image (detailed below). Object images were drawn from the Snodgrass and Vanderwart (1980) image set. Forty object images were selected for the practice run, 120 object images were selected for the encoding task on the first day, and another 120 object images were selected for the memory task on the second day. Images were randomized as to which portion of the task they would appear. Each key consisted of three features selected from three dimensions: key handle shape (circle, triangle, or rectangle), color (red, blue, or yellow), and key tip shape (1, 2, or 3), which resulted in 27 different keys (Fig. 1B).

Procedure

In the hypothesis testing condition (Fig. 1A), participants were presented with a treasure chest and three different keys on each trial. Participants were instructed to choose a key within 3 sec, with the goal of opening as many treasure chests as possible. The level of decision uncertainty varied during hypothesis testing, allowing us to investigate how decision uncertainty influenced trial-by-trial memory later (see below). If participants did not make a choice within 3 sec, "too slow" was presented on screen and the task automatically moved to the next trial. After a key was chosen, the chosen key was highlighted for 1 sec, followed by the outcome of the choice (either closed or open treasure chest). The treasure chest opened if the participant selected the key that included the target feature. If the treasure chest opened, participants viewed a trial-unique object. Because the object inside a treasure chest did not provide any useful information regarding updating their beliefs about target features, we considered the objects as incidental information and conducted a surprise memory test for the objects at a 24-h delay. The outcome remained on screen for 3 sec. The outcomes of the task were deterministic. Treasure chests opened as long as the key with the correct feature was chosen. In the control

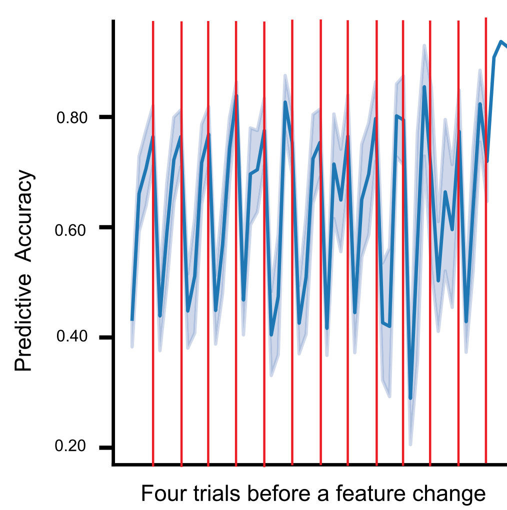

Figure 3. Predictive accuracy for the feature updating reinforcement learning model across the entire task. Because participants completed different number of trials for the task, we only show learning across the four trials before a feature change. We found that predictive accuracy increased across four trials before a feature change, but no learning across rules. Red vertical lines indicate the fourth trial before a feature change.

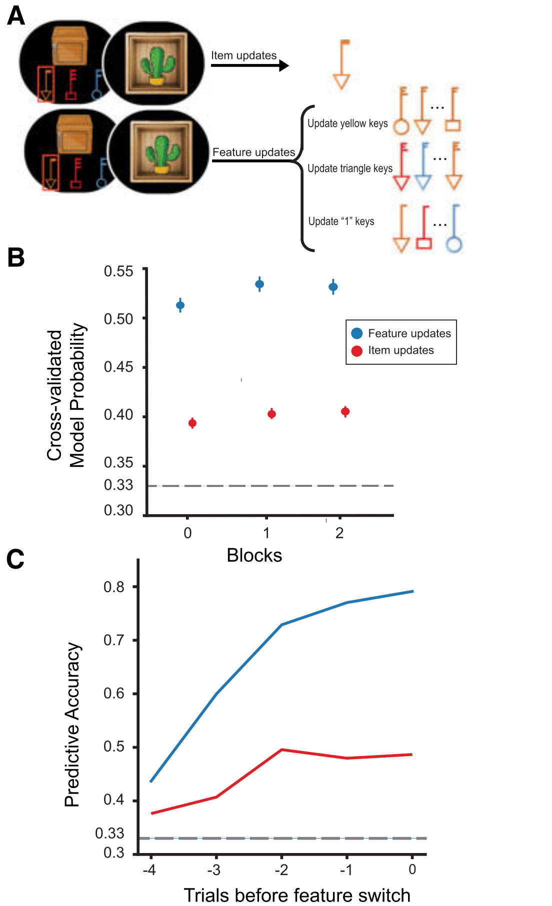

Figure 4. Feature RL outperformed item RL. (*A*) Item RL only updated value for the chosen key, whereas feature RL updated values for all keys that shared features with the chosen key. (*B*) Cross-validated model probability suggested that both feature RL and item RL predicted the data significantly better than chance (gray dash line), and feature RL outperformed item updating RL across all three blocks. (*C*) There is increasing predictive accuracy for the last four trials before a rule switch and the switching trial for both feature RL and item RL. Participants did not know the switching trial until they saw the feedback. At the trial level, both feature RL and item RL still predicted the data significantly better than chance (gray dashed line), and feature RL outperformed item RL in predicting at the trial level.

condition (Fig. 1C), individuals' choices had no effect on outcomes. On each trial, a treasure chest along with a golden key and two gray keys was presented. Participants were instructed to choose the golden key only and the treasure chest would open sometimes. Removing the color feature of the keys in the control condition precluded active hypothesis testing, but did not influence other measures, such as decision uncertainty. Then, as in the hypothesis testing task, the chosen key was highlighted before the outcome was shown. Because the number of trials for a hypothesis testing block can vary, the outcomes of the control task were yoked to the outcomes of the previous hypothesis testing task from that subject to ensure the equivalent outcome sequences and exposure to object image across conditions. Participants saw the same number of object stimuli during the hypothesis testing condition and the yoked control condition. Notably, treasure chests never opened when participants chose gray keys. Most subjects followed the instructions and only chose golden keys. Therefore, the control condition ensured factors other than motivational state to be the same as the hypothesis testing condition. The order of hypothesis testing condition and control condition was counterbalanced across subjects, such that one group of participants completed the hypothesis testing condition first, and another group of participants completed the control condition first.

Before performing the task, participants received on-screen instructions showing them the three different dimensions (color, key handle shape, and key tip shape) of a key and were instructed that in the hypothesis testing task, only one target feature would open the treasure chest. To ensure participants were not incentivized by monetary rewards, no monetary rewards were offered for opening the treasure chests. Participants practiced six trials in which the target feature was instructed, then they completed a quiz to ensure they understood the instructions. Finally, they completed a practice run in which target features were not instructed. The practice run was identical to a task run but used different stimuli. In the practice run, one group of participants completed one block of hypothesis testing task first, followed by one block of control task. The control task was yoked to the hypothesis testing task in the practice, where the outcome sequence for the control task was the same as the outcome sequence for the previous hypothesis testing task. Another group of participants completed one block of control task, followed by one block of hypothesis testing task. The control task in the practice run for the second group had 40 trials, with 50% of chance opening the treasure chest to see the object stimuli by choosing the golden key.

The task consisted of three runs, and each run consisted of a block of the hypothesis testing task and a block of the control task. The group that started with the hypothesis testing task in the practice also started with the hypothesis testing task in the real test, followed by a control task. The control task was yoked to the previous hypothesis testing task to ensure the outcome sequences for the control tasks were the same as the hypothesis testing task. The second group that started with the control condition in the practice run, started with the control task in the real test, followed by a hypothesis testing task. The first control task in the real test was yoked to the hypothesis testing task in the practice run. The control tasks in the second and third run were yoked to the hypothesis testing task in the previous runs. Pilot data suggested that four consecutive correct trials implied that participants successfully learned the target feature. Therefore, to motivate continuous hypothesis testing throughout the task, target features changed after participants responded correctly for four consecutive trials. A new target feature was randomly selected from a total of nine different features. Participants were not told that the target feature would change after certain trials, and they need to figure out the feature change by themselves. The block terminated after subjects opened twenty treasure chests. As a result, the number of trials

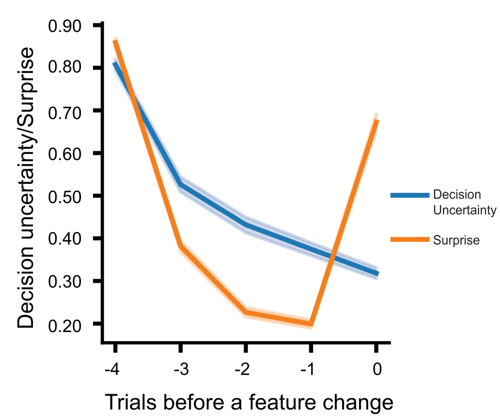

Figure 5. Changes of decision uncertainty and surprise across the four trials before a feature change, including the trial at switch. (Blue line) Decision uncertainty decreased across the four trials before a feature change, including the trial at switch. (Orange line) Surprise decreased across the four trials before a feature change, but increased at trial of switch (see the Supplemental Material for detailed results on surprise).

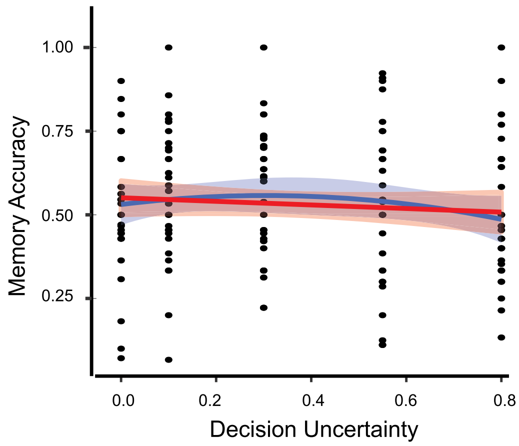

Figure 6. Relationship between memory accuracy and decision uncertainty. There was a significant linear relationship between memory accuracy and decision uncertainty (red). Memory accuracy decreased as decision uncertainty increased. On top of that, there was also a significant inverted U-shaped relationship between memory accuracy and decision uncertainty (blue). Memory was enhanced at moderate decision uncertainty. Black dots indicate individual subject data.

varied across runs. However, this ensured that each subject was exposed to the same number of object images.

At an ∼24-h delay, participants were instructed to perform a self-paced, recognition task testing their memory for the objects revealed by the treasure chests, allowing us to relate memory performance to signals, such as decision uncertainty and surprise, from the hypothesis testing task on the first day. Participants were shown an object image and had to indicate whether they had previously seen the object image and their confidence ("definitely old," "maybe old," "maybe new," and "definitely new"). Participants completed 240 recognition trials including 60 objects from the hypothesis testing condition, 60 objects from the control condition, and 120 novel objects.

Behavioral data analysis

All statistical analyses were conducted in R version 3.6.1. Mixed-effects models were run using the "lme4" package glmer function (Bates et al. 2015). Each model included random intercepts for each participant to control for individual differences in memory performance, each object image to control for different memorability of object images, total number of rules for each subject and overall trial number for each subject. Condition (Hypothesis testing, control), block order and their interaction were submitted as fixed effects. Reaction times (RTs) before feature changes were calculated by averaging across RTs of four trials before a feature change; RTs after feature changes were calculated by averaging across RTs of four trials after a feature change. RTs from the hypothesis testing task were submitted to a paired *t*-test with condition (RTs before feature changes, RTs after feature changes) as the independent variable. Memory accuracy was calculated as the hit rate. Memory confidence was calculated by grouping "definitely old" and "definitely new" responses as high memory confidence and grouping "maybe old" and "maybe new" responses as low memory confidence. We tested a model with memory accuracy and memory confidence as separate variables and tested another model including confidence in memory. We found that the model that treated memory accuracy and memory confidence separately was a better model.

Computational models

To capture learning strategies participants used during the hypothesis testing task, we compared reinforcement learning (RL) models with different task representations participants may have used to solve the task (Fig. 2A).

Item updating RL with decay

Item updating is a learning strategy in which individuals learn values for each of the 27 keys in the hypothesis testing task. This model does not generalize between keys, suggesting a key treasure memorization strategy. No higher-order hypothesis about features is formed. After choosing a key, K, the value of this key, V_k, is updated as $V_k^{new} = V_k^{old} + \alpha(R - V_k^{old})$, where α is the learning rate, and R is the outcome (treasure or no treasure). Note that there were no monetary rewards in the task. Values of unchosen keys decay according to $V_k^{unchosen} = V_k^{unchosen} \times \gamma$, where γ is the decay parameter. At the beginning of each block, the V_k is initialized to 0.

Feature updating RL with decay

Feature updating RL implies a higher-order concept learning strategy in which participants generalize across keys to isolate the reward-predictive feature. Participants learn feature weights: $W_f^{new} = W_f^{old} + \alpha(R - W_f^{old})$. The value of each key, V(K), is the sum of values of its three features: $V(K) = \sum_{f \in K} W_f$. The feature weights can be thought of as a form of selective attention that enhances the more reward-predictive features of each key (Leong et al. 2017). As in the item model, the values of unchosen features are decayed, with the decay rate controlled by γ. Note that inclusion of the decay parameter improved the performance of both models, consistent with previous reports (Niv et al. 2015). Values of the nine different features are initialized at 0 at the beginning of each run.

Decision–making

For both models, we modeled choices using a softmax decision function over the values of keys, $p(K_i) = e^{mV(K_i)} / \sum_{j=1}^{3} e^{mV(K_j)}$, to calculate the chosen probability of each key for a given trial, i. m is the inverse temperature parameter that controls the level of noise in the decision process. We fit models by maximizing the log likelihood of observed choices for all other subjects using SciPy's minimize function with the BFGS method.

Model comparison

To compare different models, we used a "leave one subject out" cross-validation approach. For each model, the model was fit to the data for all other subjects. This approach avoids overfitting by pooling data from multiple subjects (Ballard et al. 2018). We computed the average likelihood per block for each subject. This metric varies from 0 to 1, where 1/3 is the chance level.

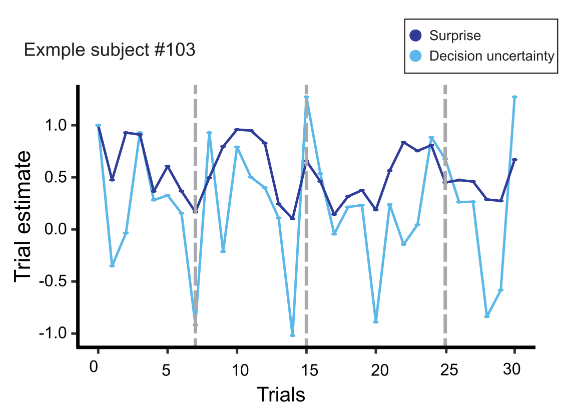

Figure 7. Correlation of decision uncertainty and surprise. An individual example of changes of decision uncertainty and surprise in one block. Gray dashed line indicates switching trials.

Estimating 24-h memory performance using decision uncertainty and surprise derived from reinforcement learning models

Individuals explore an uncertain environment with the goal of reducing uncertainty. Varied levels of decision uncertainty are important for individuals to update beliefs about the uncertain environment, which guides future decision-making. To understand the underlying mechanism of how exploration influences memory, we derived decision uncertainty from RL models (Fig. 7). We also derived surprise from RL models given previous research suggesting that surprise enhances memory (Rouhani et al. 2018) (see the Supplemental Material).

Decision uncertainty

Decision uncertainty estimates how uncertain participants were about the choice on each trial. Decision uncertainty is defined as $1-(P_k^{highest} - P_k^{second\ highest})$, where P is the probability of action for each key under the softmax function. Higher values indicate higher decision uncertainty. We used difference between the highest chosen probability and the second highest chosen probability for a given trial, instead of entropy, because during hypothesis testing, participants can rule out one possibility easily and ultimately need to decide between two options. Entropy depends on chosen probability of all three options. Therefore, we decided to use decision difficulty instead of entropy to capture decision uncertainty during hypothesis testing.

Surprise

Surprise estimates the prediction errors at the outcome (treasure revealed or withheld). surprise is defined as $R - V_k^{chosen}$.

Higher values indicate larger surprise. Surprise is positively correlated with decision uncertainty ($r = 0.83$, $P < 0.01$). In the feature updating model, because the expected value of a key (V_k) is the sum of weights on the individual features, the expected value of a key can be >1. A positive reward of 1 can lead to a slightly negative RPE. We used signed prediction errors as surprise.

Because previous research suggested that binned values are more effective to capture nonlinear relationships (Jang et al. 2019), we divided decision uncertainty and surprise into five quantiles, with roughly equal number of trials in each quantile. Binned decision uncertainty and surprise were then submitted into a mixed effect model with the fixed effects as subject and object images as random effect.

Acknowledgments

We thank Amanda Nguyen for assistance with data collection. This work was supported, in part, by grants from the National Institutes of Health (R03-DA046733 to D.V.S., and R21DA043568 and K01MH111991 to V.P.M.) and the Brain and Behavioral Research Foundation (NARSAD Young Investigator Award to V.P.M.).

Author contributions: X.S., D.V.S., and V.P.M. designed research. X.S. performed the research. X.S. and I.C.B. analyzed the data. X.S, I.C.B, D.V.S., and V.P.M. wrote the paper.

References

Ballard I, Miller EM, Piantadosi ST, Goodman ND, McClure SM. 2018. Beyond reward prediction errors: human striatum updates rule values during learning. *Cereb Cortex* **28:** 3965–3975. doi:10.1093/cercor/bhx259

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *J Stat Softw* **67:** 1–48. doi:10.18637/jss.v067.i01

Choung O, Lee SW, Jeong Y. 2017. Exploring feature dimensions to learn a new policy in an uninformed reinforcement learning task. *Sci Rep* **7:** 17676. doi:10.1038/s41598-017-17687-2

Friston K, Kiebel S. 2009. Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B Biol Sci* **364:** 1211–1221. doi:10.1098/rstb.2008.0300

Gottlieb J, Lopes M, Oudeyer P-Y. 2016. Motivated cognition: neural and computational mechanisms of curiosity, attention, and intrinsic motivation. In *Advances in motivation and achievement* (ed. Kim S et al.), vol. 19, pp. 149–172. Emerald Group Publishing Limited, Bradford, West Yorkshire, England.

Gruber M, Ranganath C. 2019. How curiosity enhances hippocampus-dependent memory: the prediction-appraisal-curiosity-exploration (PACE) framework. *Trends Cogn Sci* **23:** 1014–1025. doi:10.1016/j.tics.2019.10.003

Gruber MJ, Gelman BD, Ranganath C. 2014. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* **84:** 486–496. doi:10.1016/j.neuron.2014.08.060

Jang AI, Nassar MR, Dillon DG, Frank MJ. 2019. Positive reward prediction errors during decision-making strengthen memory encoding. *Nat Hum Behav* **3:** 719–732. doi:10.1038/s41562-019-0597-3

Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y. 2017. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* **93:** 451–463. doi:10.1016/j.neuron.2016.12.040

Markant DB, Ruggeri A, Gureckis TM, Xu F. 2016. Enhanced memory as a common effect of active learning: enhanced memory through active learning. *Mind Brain Educ* **10:** 142–152. doi:10.1111/mbe.12117

Metcalfe J, Kornell N. 2003. The dynamics of learning and allocation of study time to a region of proximal learning. *J Exp Psychol Gen* **132:** 530–542. doi:10.1037/0096-3445.132.4.530

Monosov IE. 2020. How outcome uncertainty mediates attention, learning, and decision-making. *Trends Neurosci* **43:** 795–809. doi:10.1016/j.tins.2020.06.009

Murphy C, Dehmelt V, Yonelinas AP, Ranganath C, Gruber MJ. 2021. Temporal proximity to the elicitation of curiosity is key for enhancing memory for incidental information. *Learn Mem* **28:** 38–39.doi:10.1101/lm.052241.120

Murty VP, Adcock RA. 2014. Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cereb Cortex* **24:** 2160–2168. doi:10.1093/cercor/bht063

Murty VP, DuBrow S, Davachi L. 2015. The simple act of choosing influences declarative memory. *J Neurosci* **35:** 6255–6264. doi:10.1523/JNEUROSCI.4181-14.2015

Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC. 2015. Reinforcement learning in multidimensional environments relies on attention mechanisms. *J Neurosci* **35:** 8145–8157. doi:10.1523/JNEUROSCI.2978-14.2015

Rouhani N, Niv Y. 2021. Signed and unsigned reward prediction errors dynamically enhance learning and memory. *Elife* **10:** e61077. doi:10.7554/eLife.61077

Rouhani N, Norman KA, Niv Y. 2018. Dissociable effects of surprising rewards on learning and memory. *J Exp Psychol Learn Mem Cogn* **44:** 1430–1443. doi:10.1037/xlm0000518

Ruggeri A, Markant DB, Gureckis TM, Bretzke M, Xu F. 2019. Memory enhancements from active control of learning emerge across development. *Cognition* **186:** 82–94. doi:10.1016/j.cognition.2019.01.010

Snodgrass JG, Vanderwart M. 1980. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J Exp Psychol Hum Learn* **6:** 174–215. doi:10.1037//0278-7393.6.2.174

Voss JL, Gonsalves BD, Federmeier KD, Tranel D, Cohen NJ. 2011. Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. *Nat Neurosci* **14:** 115–120. doi:10.1038/nn.2693

Wilson RC, Niv Y. 2012. Inferring relevance in a changing world. *Front Hum Neurosci* **5:** 189. doi:10.3389/fnhum.2011.00189

Received June 22, 2021; accepted in revised form February 25, 2022.

Decision uncertainty during hypothesis testing enhances memory accuracy for incidental information

Xinxu Shen, Ian C. Ballard, David V. Smith, et al.

Learn. Mem. 2022, **29:**
Access the most recent version at doi:10.1101/lm.053458.121