1
2
3
4
5
6
7      **Positive reward prediction errors strengthen incidental memory encoding**
8
9
10
11
12    Anthony I. Jang[1*], Matthew R. Nassar[1*], Daniel G. Dillon[2,3] and Michael J. Frank[1]
13
14
15
16
17
18    [1]Department of Cognitive, Linguistic, and Psychological Sciences; Brown Institute for
19    Brain Science, Brown University, Providence RI 02912-1821
20    [2] Center for Depression, Anxiety and Stress Research, McLean Hospital, Belmont MA
21    02478
22    [3] Harvard Medical School, Boston MA 02115
23    *Authors contributed equally and are listed alphabetically
24
25
26
27
28
29
30
31
32
33
34
35    Corresponding Author:
36    Matthew R. Nassar
37    Department of Cognitive, Linguistic and Psychological Sciences
38    Brown University
39    Providence, RI 02912-1821
40    Phone: 607-316-4932
41    E-mail: matthew_nassar@brown.edu
42
43
44
45

## Abstract

The dopamine system is thought to provide a reward prediction error signal that facilitates reinforcement learning and reward-based choice in corticostriatal circuits. While it is believed that similar prediction error signals are also provided to temporal lobe memory systems, the impact of such signals on episodic memory encoding has not been fully characterized. Here we develop an incidental memory paradigm that allows us to 1) estimate the influence of reward prediction errors on the formation of episodic memories, 2) dissociate this influence from other factors such as surprise and uncertainty, 3) test the degree to which this influence depends on temporal correspondence between prediction error and memoranda presentation, and 4) determine the extent to which this influence is consolidation-dependent. We find that when choosing to gamble for potential rewards during a primary decision making task, people encode incidental memoranda more strongly even though they are not aware that their memory will be subsequently probed. Moreover, this strengthened encoding scales with the reward prediction error, and not overall reward, experienced selectively at the time of memoranda presentation (and not before or after). Finally, this strengthened encoding is identifiable within a few minutes and is not substantially enhanced after twenty-four hours, indicating that it is not consolidation-dependent. These results suggest a computationally and temporally specific role for putative dopaminergic reward prediction error signaling in memory formation.

**Introduction**

Behaviors are often informed by multiple kinds of memories. For example, a decision about what to eat for lunch might rely on average preferences that have been slowly learned over time and that aggregate over many previous experiences, but it might also be informed by specific, temporally precise memories (e.g., ingredients seen in the fridge on the previous day). These different kinds of memories prioritize different aspects of experience. For example, reinforcement learning typically aggregates information across relevant experience to form general preferences that are used guide behavior [1], whereas episodic memories allow access to details about specific, previously experienced events or scenes with limited interference from other similar ones. Computational modeling suggests that these two kinds of memories have different representational requirements and are likely subserved by anatomically distinct brain systems [2-4]. In particular, a broad array of evidence suggests that reinforcement learning is implemented through cortico-striatal circuitry in the prefrontal cortex and basal ganglia [5-8] whereas episodic memory appears to be reliant on synaptic changes in medial temporal structures, especially the hippocampus [1,9-12].

While the distinct anatomy of these systems allows them to operate over different representations, they are not independent. Medial temporal regions provide direct inputs into striato-cortical regions [2,3,13-15] and both regions receive shared information through common intermediaries [5-8,16]. Furthermore, both systems receive neuromodulatory inputs that undergo context dependent fluctuations that can affect synaptic plasticity and alter information processing in both systems [17,18]. Recent work in computational neuroscience has highlighted potential roles for two neuromodulators, dopamine and norepinephrine, in implementing and optimizing reinforcement learning in a changing world. In particular, dopamine is thought to supply a reward prediction error (RPE) signal that gates Hebbian plasticity in the striatum, facilitating repetition of rewarding actions [5,6,19,20]. In untrained animals dopamine prediction error signals are observed in response to primary rewards, but with experience dopamine signals precede to the earliest cue predicting future reward [5]. Such cue-induced dopamine signals are thought to serve a motivational role [21], biasing behavior toward the effortful and risky actions that could allow for acquisition of the upcoming reward [22-27]. In contrast, norepinephrine has been proposed to provide a salience signal that amplifies overall learning from surprising events, irrespective of valence, or during periods of uncertainty [28-30]. In many experimental tasks, such a salience signal would look very similar to a reward prediction error signal, but careful experimental design can result in their dissociation [31].

While normative roles for dopamine and norepinephrine have frequently been discussed in terms of their effects on reinforcement learning and motivational systems, such signals likely also affect processing in medial temporal memory systems [32-35]. For example, dopamine can enhance LTP [36] and replay [37] in the hippocampus, which could potentially provide a mechanism to prioritize behaviorally relevant information for longer

1   term storage [32]. More recent work using optogenetics to perturb hippocampal dopamine
2   inputs revealed a biphasic relationship, whereby low levels of dopamine suppress
3   hippocampal information flow but higher levels of dopamine facilitate it [35]. Given that the
4   highest levels of dopamine are typically elicited by dopamine bursts [38], especially with
5   larger reward prediction errors [5], this result suggests that memory encoding in the
6   hippocampus might be enhanced for unexpectedly positive events.
7
8   However, despite relatively strong evidence that dopaminergic projections signal reward
9   prediction errors [5,39], and that dopamine release in the hippocampus can facilitate
10  memory encoding in non-human animals [40],  evidence for a positive effect of reward
11  prediction errors on memory formation in humans is scarce. Monetary incentives and
12  reward expectation can be manipulated to improve episodic encoding, even of incidental
13  memories, but it is not clear that such effects are driven by reward prediction errors
14  rather than motivational signals or reward value per se [18,33,41,42]. The few studies that
15  have closely examined the relationship between reward prediction error signaling and
16  episodic memory have not found evidence for a memory advantage after positive
17  prediction errors [43,44]. However, there are a number of technical factors that could mask
18  a relationship between reward prediction errors and incidental memory formation in
19  standard paradigms. For instance, such tasks typically have not controlled for salience
20  signals, such as surprise and uncertainty, that can be closely related to RPEs, and that
21  are thought to exert independent effects on episodic encoding through a separate
22  noradrenergic neuromodulatory system [28-30,45].
23
24  Here we combine a novel behavioral paradigm with computational modeling to clarify
25  the impact of prediction error and surprise, elicited during reinforcement learning, on
26  episodic encoding. Our paradigm required subjects to encode images in the context of a
27  learning and decision-making task, and then to complete a surprise recognition
28  memory test for the images. The decision-making task required subjects to decide
29  whether to accept or reject a risky gamble based on the value of potential payouts and
30  the reward probabilities associated with two image categories, which they learned
31  incrementally based on trial-by-trial feedback. Our design allowed us to measure and
32  manipulate reward prediction errors at multiple timepoints and to dissociate those RPEs
33  from other computational factors that are often correlated with them. In particular, our
34  paradigm and computational models allowed us to manipulate and measure surprise
35  and uncertainty, which have also been implicated in learning and episodic encoding,
36  and are often closely related to RPEs in many tasks. However, both surprise and
37  uncertainty are thought be conveyed through noradrenergic modulation (rather than
38  dopamine, which is thought to reflect RPE). We also assessed the degree to which
39  relationships between these computations and episodic encoding depend on
40  consolidation by testing recognition memories either immediately after the decision task
41  or after a 24 hour delay.
42
43  Our results reveal that subjects were more likely to remember images presented in trials
44  in which they accepted the risky gamble. Moreover, the degree of this memory benefit

1    increased with the RPE experienced at the time of image presentation, but not by the
2    RPE, surprise, or uncertainty associated with either the previous or subsequent trial
3    outcome. These results were replicated in an independent sample, which also
4    demonstrated sensitivity to counterfactual information about choices the participants did
5    not make. Collectively, these data demonstrate a key role for reward prediction errors in
6    episodic encoding, clarify the timescale and computational nature of interactions
7    between reinforcement learning and encoding, and make testable predictions about the
8    neuromodulatory mechanisms underlying both processes.
9
10
11   **Results**
12
13   The goal of this study is to elucidate how computational factors that govern trial-to-trial
14   learning and decision-making might impact episodic memory encoding and retrieval. To
15   do so, we designed a two-part task that included a learning and decision-making phase
16   (Fig 1a) followed by a recognition memory phase (Fig 1c) (see Methods for additional
17   details regarding the task, participants, and analysis). During the learning phase, on
18   each trial subjects decided whether to accept ("play") or reject ("pass") an opportunity to
19   gamble based on the potential reward payout. The magnitude of the potential reward
20   was revealed numerically at the start of each trial, but its probability had to be learned
21   via feedback. Subjects were presented with a visual image and were told that the
22   probabilities of reward would depend on whether the (trial-unique) image belonged to
23   one of two categories (animate or inanimate). The reward probability was yoked across
24   categories, such that p(animate) = 1 – p(animate). On each trial, the subject integrated
25   information about the trial payout (selected at random independently for each trial) and
26   learned probability in order to decide whether to "play" or "pass" (Fig 1a). Furthermore,
27   sequential presentation of value, probability and trial outcome information allowed us to
28   manipulate subject reward expectations dynamically within a trial to separately elicit
29   reward prediction errors before, during, and after image presentation (Fig 1b).
30
31   The learning phase thus presented subjects with a series of unique stimuli (images from
32   animate/inanimate categories) in a context that allowed for the statistical dissociation of
33   three computational factors thought to govern learning from feedback: prediction error,
34   surprise, and uncertainty. Dissociation of these three factors was achieved in part
35   through the independent manipulation of reward probability and value, and in part
36   through occasional change points in the assignments of the yoked reward probabilities
37   to each of the two image categories (Fig 1d; [46]). This design yielded estimates of
38   surprise from an ideal observer model that spiked at improbable outcomes, including—
39   but not limited to—those observed after changes in the reward probabilities (Fig 1e).
40   Estimates of uncertainty changed more gradually and tended to be highest during
41   periods following surprise (i.e., when the reward outcomes are volatile, one becomes
42   more uncertain about the learned probability; Fig 1f). Reward prediction errors at time of
43   feedback were highly variable across trials and more related to the probabilistic (and
44   bivalent) trial outcomes than to transitions in the reward structure (Fig 1g).

1        While repeated-choice bandit tasks typically involve a prediction error at time of
2 feedback, our task also provided some information about the probability of a rewarding
3 outcome during the decision phase, coincident with the image presentation. This
4 allowed us to examine the effects of a separate, image reward prediction error signal
5 (Fig 1b). On some trials, the presented image category was associated with a higher-
6 than-expected reward probability leading to a positive image reward prediction error,
7 whereas on other trials the presented image suggested a lower than expected reward
8 probability leading to a negative image prediction error (Fig 1h). The image prediction
9 errors could be thought of as the dopamine signal that would be expected to occur in
10 response to the probabilistic reward cue provided by the memoranda themselves.

11
12
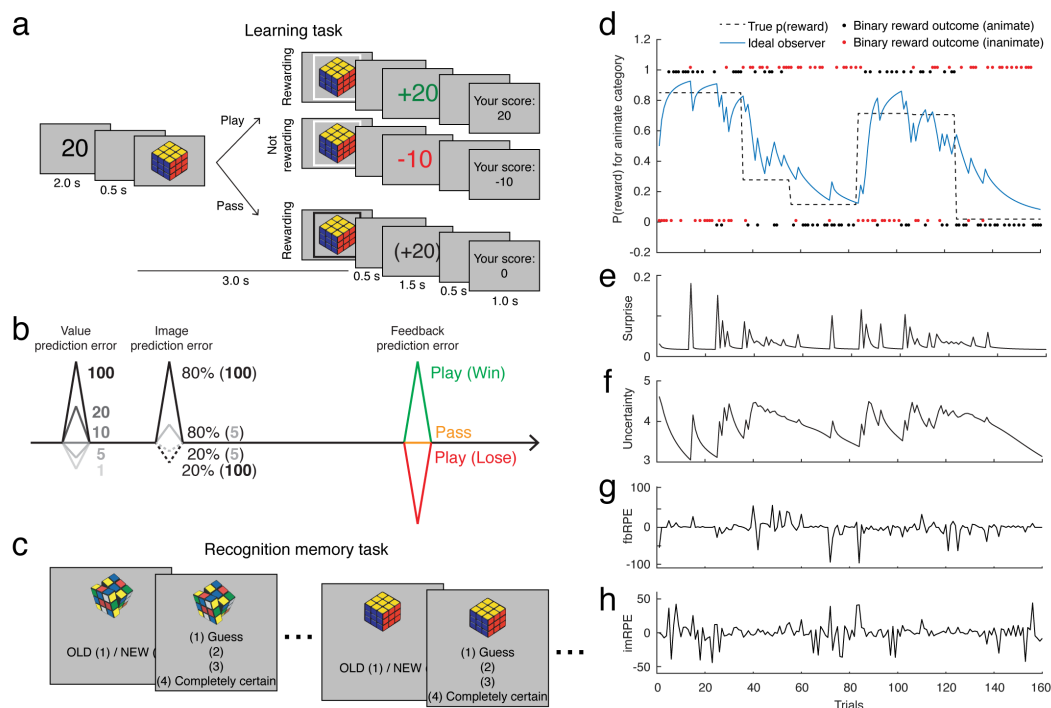


13
14
15 **Figure 1:** Dissociating effects of reward prediction error, surprise and uncertainty on incidental memory
16 encoding. **a,** In the learning task subjects were informed about the potential value of a successful gamble
17 for the current trial (two seconds, 20 for example trial) and then shown a unique image belonging to one
18 of two categories (animate/inanimate) for a total of three seconds, during which the subject would decide
19 whether to accept ("play") or reject ("pass") the gamble. After a brief delay, subjects were informed about
20 their earnings and, if they rejected the gamble, saw counterfactual information regarding the trial outcome.
21 At the end of each trial a cumulative total score was displayed. **b,** The learning task facilitates the
22 manipulation of reward prediction errors before, during, and after image presentation. Prediction errors
23 elicited before image presentation would reflect the value information presented at trial outset (leftmost
24 peak; colors indicate value), prediction errors at image presentation would reflect reward probability
25 information conveyed by the image and its interaction with trial value (middle peak; solid/dashed lines
26 reflect high/low reward probabilities), and prediction errors occurring after image presentation would
27 convey information about trial outcome (rightmost peak; colors reflect trial outcome). **c,** After the learning
28 task, subjects completed a surprise recognition memory test in which each image was either "old" (i.e.,
29 from the learning task) or "new" (i.e., a novel lure). Subjects were asked to provide a binary answer along

1    with a 1-4 confidence rating for each image. Importantly, the new images were semantically matched to
2    the old images such that accurate responding depended on the retrieval of detailed perceptual information
3    from encoding. **d,** Reward probabilities during the learning task were determined by image category, were
4    yoked across categories, and were reset occasionally to require learning (black dotted line). Binary
5    outcomes indicating whether the gamble would be rewarded (red/black dots) were governed by reward
6    probabilities and could be used by an ideal observer model to infer the underlying reward probabilities
7    (blue). **e-g,** The ideal observer learned in proportion to the surprise associated with a given trial outcome
8    (**e**) and the uncertainty about its estimate of the current reward probability (**f**), both of which fluctuated
9    dynamically throughout the task and were dissociable from reward prediction error signals at time of
10    feedback (**g**) and at time of image presentation (**h**).

11

12    Analyses of data from 199 subjects suggest that they (1) integrated reward probability
13    and value information and (2) utilized reward prediction errors, surprise, and uncertainty
14    to gamble effectively. Subjects increased the proportion of play (gamble) responses as
15    a function of both trial value and the ground truth category reward probability (Fig 2a).
16    To capture trial-to-trial dynamics of subjective category probability assessments, we fit
17    play/pass data from each subject with a set of reinforcement learning models. The
18    simplest such model fit betting behavior as a weighted function of reward magnitude
19    and probability, with probabilities updated on each trial with a fixed learning rate. More
20    complex models (see Methods for details) considered the possibility that this learning
21    rate might itself be adjusted according to other factors such as surprise, uncertainty, or
22    whether the subject had decided to gamble on the trial. Consistent with previous work
23    [30,31,47,48], the best fitting model adjusted learning according to normative measures of
24    both surprise and uncertainty (Fig 2c). Coefficients describing the effects of surprise and
25    uncertainty on learning tended to be positive across subjects, indicating that subjects
26    were more responsive to feedback if it was surprising or was provided during a period of
27    uncertainty (Fig 2d; surprise: $t = 2.69$, df = 198, $p = 0.0078$; uncertainty $t = 6.38$, df =
28    198, $p < 0.001$). Thus, subject behavior in the learning task is best described as using
29    surprise and uncertainty to scale the extent to which reward prediction errors are used
30    to adjust subsequent behavior.

31
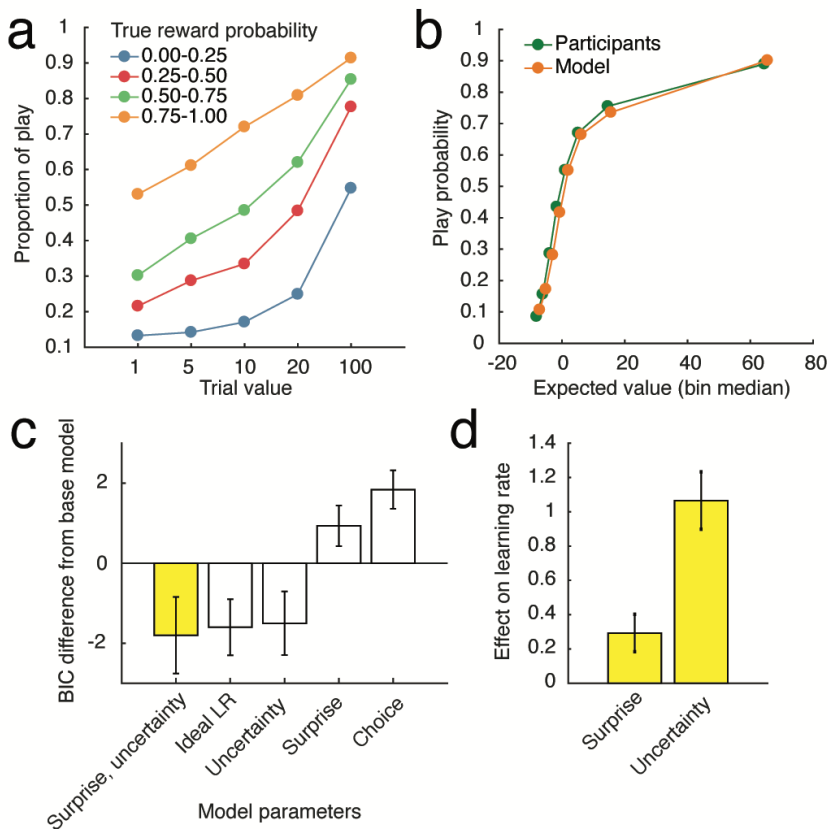
1
2 **Figure 2:** Subjects' gambling behavior integrated reward value and subjective reward probability
3 estimates, which were updated as a function of surprise and uncertainty. **a,** Proportion of trials in which
4 the subjects chose to play, broken down by reward value and true reward probability. **b,** Subject choice
5 behavior and model-predicted choice behavior. The model with the lowest BIC, which incorporated the
6 effects of surprise and uncertainty on learning rate, was used to generate model behavior. Expected
7 values for all trials were divided into 8 equally sized bins for both subject and model-predicted behavior. **c,**
8 Bayesian information criterion of five reinforcement learning models with different parameters that affect
9 learning rate. **d,** Mean maximum likelihood estimates of surprise and uncertainty parameters of the best
10 fitting model (first bar on **c**). Error bars indicate standard error across subjects.

11

12

13 Behavior in the recognition task confirmed that subjects formed memories about the
14 unique images displayed in the learning task. Subjects completed a surprise recognition
15 memory task either five minutes (no delay, $n = 109$) or twenty-four hours (24 hour delay,
16 $n = 90$) after completion of the learning task. During the recognition memory task,
17 subjects viewed all the "old" images from the learning task plus a semantically matched
18 set of "new" foil images that were not shown previously. For each image, subjects
19 provided a binary response indicating whether the image was old or new, along with a
20 1-4 confidence measure.

21

22 Subjects in both delay conditions reliably identified images that had been presented in
23 the learning task with accuracy above chance (Fig 3a; mean(sem) d' = 0.85(0.042) for
24 no delay and 0.50(0.030) for 24 hour delay condition). The reliability of memory reports

1 was greater when subjects expressed higher levels of confidence (mean(sem) d' =
2 1.12(0.062) for no delay and 0.65(0.039) for 24 hour delay condition) than when they
3 reported lower confidence (mean(sem) d' = 0.44(0.038) for no delay and 0.25(0.037) for
4 24 hour delay condition).
5
6 To aggregate information provided in the binary reports and confidence measure we
7 transformed these sources of data to create a single 1-8 memory score, such that 8
8 reflected a high confidence "old" response and 1 reflected a high confidence "new"
9 response. As expected, the true proportion of "old" images was higher when memory
10 scores were highest, and the relationship between memory scores and ground truth was
11 monotonic and roughly linear across both delay conditions (Fig 3b). Thus, subjects
12 incidentally formed lasting memories for the images displayed in the learning task, and
13 memory scores provided a measure of the subjective memory strength associated with
14 each image.
15
16 Recognition memory for an image depended on the context in which that image was
17 presented during the learning task. Memory scores for previously viewed images were
18 higher for images observed on trials in which subjects gambled (play) than for images
19 observed on trials in which they passed (Fig 3c; FigS1). The difference between
20 memory scores for old and new items was larger for play than pass trials (t=3.30,
21 dof=198, p =0.001) and this did not differ across delay conditions (t=0.38, dof=198, p =
22 0.70) (Fig 3d). Larger memory scores were produced, at least in part, by an increase in
23 the sensitivity of memory reports. Across all possible memory scores, hit rate was
24 higher for play trials than for pass trials and the area under subject-specific ROC curves
25 generated in this way was greater for play as compared to pass trials (Fig 3e; t=3.12,
26 dof=198, p = 0.002) but did not differ across delay conditions (t=0.34, dof=197, p=.73).
27 In principle, this enhanced memory encoding could be driven by positive reward
28 prediction errors, which would occur in response to high probability images (figure 1b)
29 and motivate play decisions (figure 2a). In subsequent analyses we test this idea
30 directly.
31
32
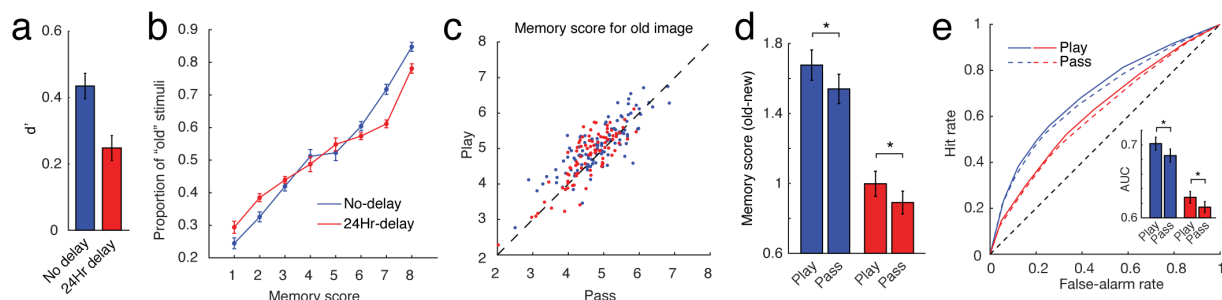


33
34 **Figure 3:** Recognition memory is stronger for stimuli presented on trials in which subjects accepted the
35 gamble. **a,** Average d' for the two delay conditions. **b,** Average proportion of image stimuli that were "old"
36 (presented during the learning task), separated by memory score. **c,** Mean memory score of "old" images
37 for play trials versus pass trials. Each point represents a unique subject. A majority of subjects lie above
38 the diagonal, indicating better memory performance for play trials. **d**, Mean pairwise difference in memory

1   score between the "old" images and their semantically-matched foil images. **e,** ROC curves for play vs.
2   pass trials. Area under the ROC curves (AUC) shown in the inset. AUC was greater for play versus pass
3   trials, indicating better detection of old versus new images for play trials compared to pass trials. Colors
4   indicate time between encoding and memory testing; blue = no delay, red = 24 hour delay.
5
6
7   The degree to which memory was enhanced on play trials depended on the magnitude
8   of the reward prediction error at time of image presentation (see Fig 1a,b). In particular,
9   memory scores for play trials increased as a function of the reward prediction error
10  computed at the time of image presentation (Fig 4a,b; t = 4.29, df = 198, p = 3 x 10$^{-5}$) in
11  a manner that did not depend on delay condition (t = 0.09, df = 198, p = 0.93).
12  Moreover, this effect was most prominent in the subjects that displayed gambling
13  behaviors that were the most sensitive to trial-to-trial fluctuations in probability and value
14  (spearman correlation of gambling GLM coefficients (probability & value) with memory
15  score image RPE coefficient: ρ = 0.14, p = 0.04).
16
17  To better understand the nature of this relationship, we show subsequent memory
18  scores related to the constituent components of the reward prediction error signal. A key
19  feature of an error in reward prediction at the time of image presentation is that it should
20  depend directly on reward probability (i.e., the probability associated with the image
21  category relative to the average reward probability across categories) as this is the
22  reward information cued by the image. In contrast, the trial value should not directly
23  influence RPE because this value was already indicated prior to image presentation,
24  and should thus only inform reward predictions themselves, and not their errors.
25  Consistent with a selective effect of RPE at time of image presentation, subsequent
26  memories were stronger for play trials in which the image category was associated with
27  a higher reward probability (Fig 4c,d; t = 3.31, df = 198, p = 0.001) but not for play trials
28  with higher potential outcome value, which if anything were associated with slightly
29  lower memory scores (Fig 4e,f; t = -1.97, df = 198, p = 0.051). Reward probability
30  effects were stronger in subjects that displayed more sensitivity to probability and value
31  in the gambling task (spearman correlation of gambling GLM coefficients (probability &
32  value) with memory score image reward probability coefficient: ρ = 0.16, p = 0.02).
33
34
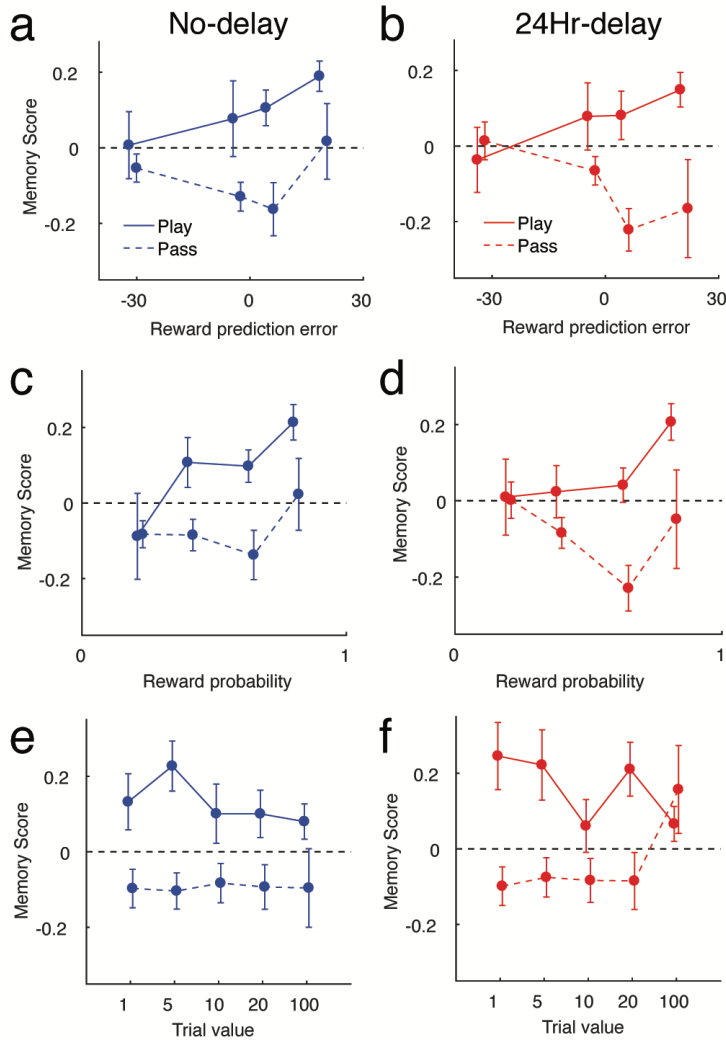
**Figure 4:** Subsequent memory strength depends on the reward prediction error at time of image presentation, but not trial value. **a,b,** Reward prediction error during image presentation shows a positive association with subsequent recognition memory for the image in both no delay (**a**; blue) and 24 hour delay (**b**; red) conditions. **c-f,** Reward probability estimates (**c,d**), but not reward value (**e,f**) retained the positive association with subsequent memory. This suggests that prediction error that occurs during image presentation, but not the overall value of the image, is driving the subsequent memory effect. Colors indicate time between encoding and memory testing; blue = no delay, red = 24 hour delay.

While subject gambling behavior depended critically on the reward prediction error, uncertainty, and surprise associated with trial feedback, to our surprise none of these factors influenced subsequent memory for the images. Specifically, memory scores were not systematically related to the reward prediction error experienced at time of feedback on the trial preceding image presentation (Fig 5a; t = 0.25, dof = 198, p = 0.80) or immediately after image presentation (Fig 5b; t = -1.19, dof = 198, p = 0.26). Similarly, the surprise associated with feedback preceding (Fig 5b; t = -1.71, dof = 198, p = 0.088) or following (Fig 5d; t = 1.24, dof = 198, p = 0.16) image presentation was not systematically related to subsequent memory scores. Furthermore, there were no

1  obvious relationships between subsequent memory score and uncertainty, again
2  despite the apparent impact of uncertainty on subject gambling behavior (fig 5e,
3  previous trial uncertainty, t=1.70, dof=198, p=0.091; fig 5f, current trial uncertainty, t = -
4  0.63, dof = 198, p = 0.53).

5

6
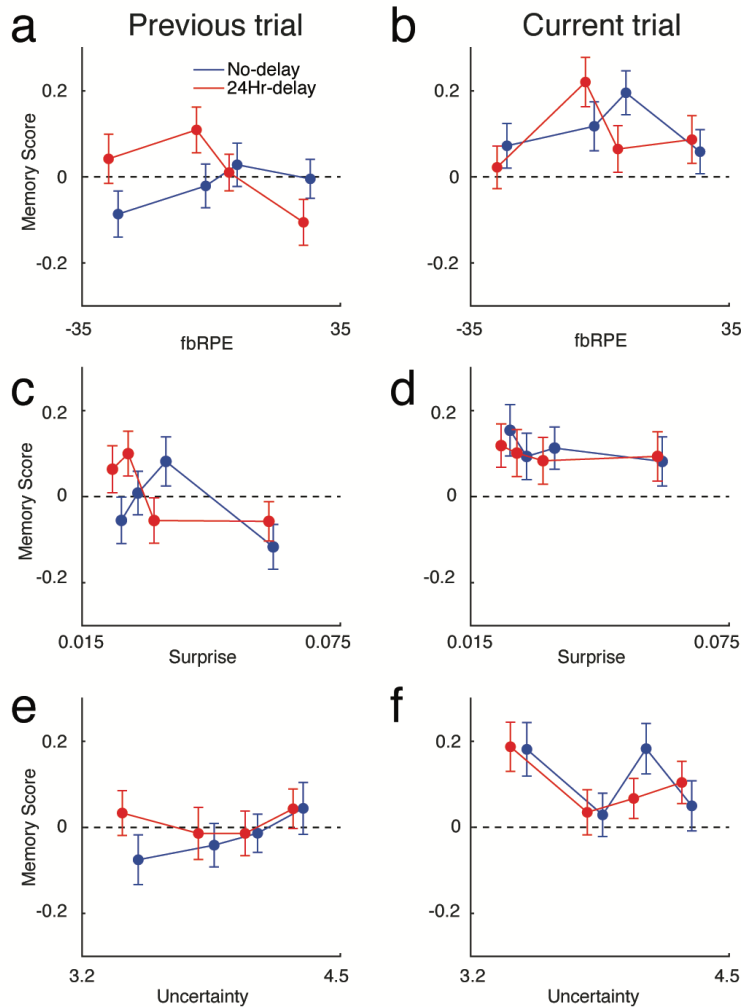


7
8  **Figure 5:** Subsequent memory is not affected by unexpected rewards, surprise, or outcome uncertainty
9  during the feedback preceding or following image presentation. **a,b,** Reward prediction error during the
10  feedback phase of the previous trial (**a**) or current trial (**b**) did not affect subsequent memory. **c,d,**
11  Surprise associated with the feedback phase of the previous (**c**) or current (**d**) trial did not affect
12  subsequent memory. **e,f,** Uncertainty during the previous (**e**) or current (**f**) trial did not affect subsequent
13  memory. Colors indicate time between encoding and memory testing; blue = no delay, red = 24 hour
14  delay.

15

16

17

18

19

20

1  To better estimate the contributions of learning-related computations to subsequent
2  memory strength, we constructed a hierarchical regression model capable of 1) pooling
3  information across subjects and delay conditions in an appropriate manner, 2)
4  estimating the independent contributions of each factor while simultaneously accounting
5  for all others, and 3) accounting for the differences in memory scores attributable to the
6  images themselves. The hierarchical regression model attempted to predict memory
7  scores by estimating coefficients at the level of items and subjects, as well as estimating
8  the mean parameter value over subjects and the effect of delay condition for each
9  parameter (Fig 6a).
10       Consistent with the results presented thus far, the hierarchical regression results
11  support the notion that encoding was strengthened by the decision to gamble (play vs.
12  pass) and reward prediction errors elicited at the time of image presentation, but not by
13  the computational factors that controlled learning rate (surprise and uncertainty). Play
14  trials were estimated to contribute positively to encoding, as indexed by uniformly
15  positive values for the posterior density on the play/pass parameter (Fig 6b top row of
16  column 2; table 1). The reward probability associated with the displayed category was
17  positively related to subsequent memory on play trials (Fig 6b column 3; table 1), as
18  was its interaction with value (Fig. 6b column 5; table 1)–although there was no reliable
19  effect of value itself (Fig 6b columns 6; table 1). The direction of the interaction effect
20  suggests that subjects were more sensitive to image probability on trials in which there
21  were more points available to be won, consistent with the more nuanced predictions of a
22  reward prediction error at time of image presentation (Fig 1b). All observed effects were
23  selective for the old items that subjects observed in the task, as the same model fit to
24  the new, foil images yielded coefficients near zero for each of these terms (Fig S2).
25  Consistent with our previous analysis, coefficients for the uncertainty and surprise terms
26  were estimated to be near zero (Fig 6b rightmost columns; table 1).
27       In addition to supporting our previous analyses, our model allowed us to examine
28  the extent to which any subsequent memory effects required a consolidation period. In
29  particular, any effects on subsequent memory that were stronger in the 24hr delay
30  condition vs. the immediate condition might reflect an effect of post-encoding processes.
31  Despite evidence from animal literature that dopamine can robustly affect memory
32  consolidation (e.g., Bethus et al., 2010), we did not find strong support for any of our
33  effects being consolidation dependent (note lack of positive coefficients in bottom row of
34  Fig 6b, which would indicate effects stronger in the 24 hour condition). As might be
35  expected, subjects in the no delay condition tended to have higher memory scores
36  overall (Fig 6b bottom of column 1; table 1) but their memory scores also tended to
37  change more as a function of reward probability (Fig 6b bottom of column 3; table 1)
38  than did their counterparts in the 24hr delay condition. These results reveal the
39  expected decay of memory over time, and suggest that the image prediction errors
40  induced by categories associated with higher reward probability are associated with an
41  immediate and decaying boost in memory accuracy.
42
43       In summary, behavioral data and computational modeling support a role for
44  computation of surprise, uncertainty and reward prediction error in the learning phase of

1    our paradigm. However, only decisions to gamble and the instantaneous reward
2    prediction error at the time of image presentation were related to subsequent memory.
3    To better understand the reward prediction error effect, and to ensure the reliability of
4    our findings, we conducted a second experiment.
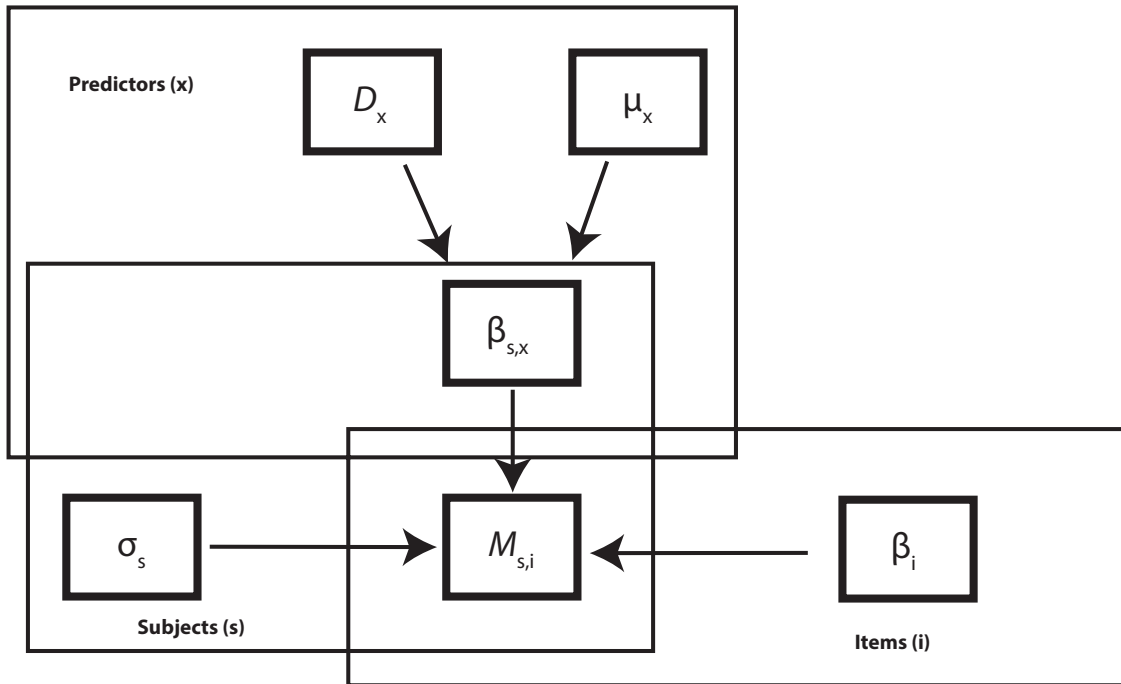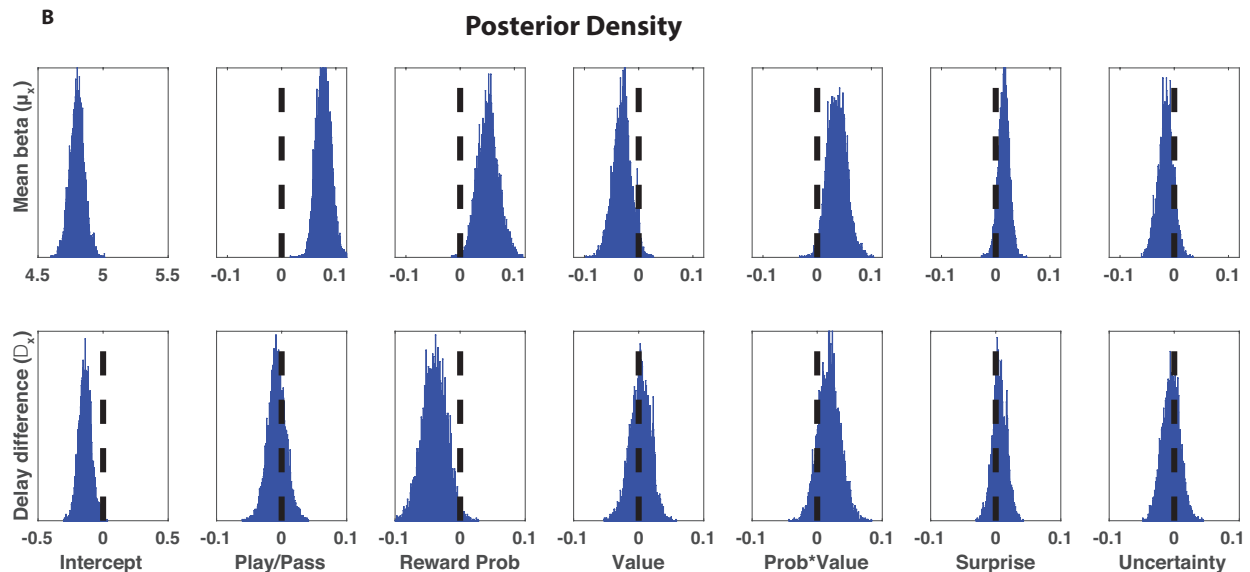


5
6    **Figure 6:** Hierarchical regression model reveals effects of choice and positive prediction errors on
7    recognition memory encoding. **a:** Graphical depiction of hierarchical regression model. Memory scores for
8    each subject and item ($M_{s,i}$) were modeled as normally distributed with subject specific variance ($\sigma_s$) and a
9    mean that depended on the sum of two factors: 1) subject level predictors related to the decision context

1. in which an image was encountered (i.e., whether the subject played or passed) linearly weighted
2. according to coefficients ($\beta_{s,x}$) and 2) item level predictors specifying which image was shown on each
3. trial and weighted according to their overall memorability across subjects ($\beta_i$). Coefficients for subject level
4. predictors were assumed to be drawn from a global mean value for each coefficient ($\mu_x$) plus an offset
5. related to the delay condition ($D_x$). Parameters were weakly constrained with priors that favored mean
6. coefficient values near zero and low variance across subject and item specific coefficients. **b:** Posterior
7. probability densities for mean predictor coefficients ($\mu_x$; top row) and delay condition parameter difference
8. ($D_x$; bottom row) estimated through MCMC sampling over the graphical model informed by the observable
9. data ($M_{s,i}$).

10.
11.

12. **Experiment 2**

13.

14. Our previous findings suggested that variability in the strength of memory encoding was
15. related to computationally-derived RPE signals and the gambling behavior that elicited
16. them. However, the yoked reward probabilities in Experiment 1 ensured that the reward
17. probability associated with the image category presented would be perfectly anti-
18. correlated with the reward probability associated with the other category that was not
19. presented. Thus, while a high reward category item would increase the expected reward
20. relative to before the trial and hence elicit an RPE, we were unable to disentangle
21. whether the observed effects were driven by the reward probability directly, the
22. counterfactual reward associated with the alternate category, or, as would be predicted
23. by a true reward prediction error, their difference. To address this issue we conducted a
24. second experiment in which expectations about reward probability were manipulated
25. independently of the actual reward probability on each trial allowing us to distinguish
26. between these alternative explanations.

27.

28. Specifically, we modified the design of our task such that the learning phase included
29. separate manipulations of reward probability for the two image categories. Thus, during
30. some trials both categories would be associated with a high reward probability and in
31. some sessions both would be associated with a low reward probability (Fig 7a-b). In this
32. design, RPEs are relatively small when the reward probabilities are high for both
33. categories and deviate much more substantially when the reward probability differs
34. across image category (Fig 7c). Thus, if the factor boosting subsequent memory scores
35. is truly a reward prediction error, it should depend positively on the reward probability
36. associated with the image category, but negatively with the reward probability
37. associated with the other category.

38.

39. Subjects in both delay conditions reliably identified images presented in the learning
40. task with accuracy above chance (Fig 7d; mean(sem) d' = 0.91(0.058) for no delay and
41. 0.54(0.034) for 24 hour delay condition). We also observed a robust replication of the
42. difference in memory scores for trials in which subjects either gambled or passed on the
43. offer to gamble (Fig 7e; t=3.89, dof=172, p<0.001; difference across delay condition, t=-
44. 0.60, dof=171, p=0.55).

45.

1    In the new experimental design, we could analyze variability in memory scores for each
2    old image as a function of its associated category reward probability ("image category")
3    and the reward probability associated with the other category ("other category"). Based
4    on the RPE account, we hypothesized that we would see a positive effect of the
5    experienced reward (image category) and a negative effect of the reward that would
6    have otherwise been available (other category) on subsequent memory. Consistent with
7    this prediction, for both delay conditions, there was a cross-over effect whereby memory
8    scores scaled positively with the probability associated with the image category (Fig 7f,g
9    t=2.37, dof=172, p=0.019), and negatively with the probability associated with the other
10   category (t=-2,27, dof=172, p=0.024). These effects did not differ across delay
11   conditions (current category, t=0.55, dof=171, p=0.58; other category, t=0.36, dof=171,
12   p=0.72).
13
14
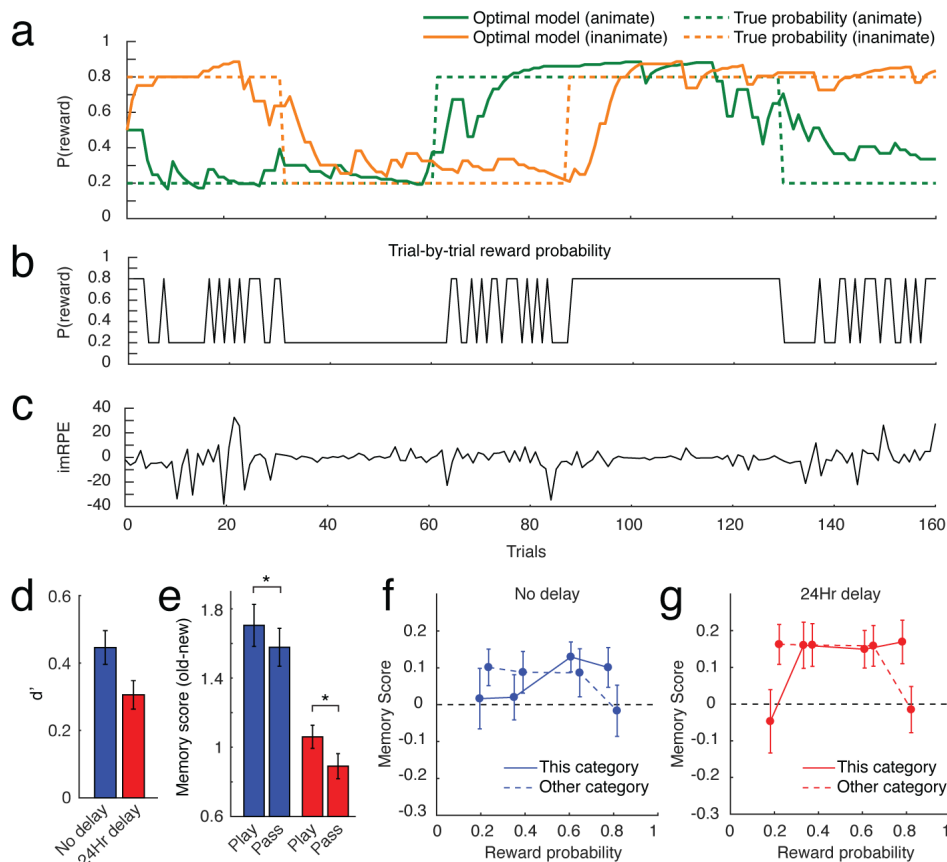


15
16   **Figure 7:** Experiment 2 allowed us to separately estimate the effects of both category probabilities on
17   subsequent memory recall. **a:** In the new learning task, the true reward probabilities of the two categories
18   were independent, and were restricted to either 0.2 or 0.8. The change-points occurred randomly,
19   ensuring that each subject completed a unique learning task. Restrictions were applied so that each task
20   contained at least one block (constituting at least 20 trials) of the four possible reward probability
21   combinations (0.2/0.2, 0.2/0.8, 0.8/0.2, 0.8/0.8). **b:** Trial-by-trial reward probability shows stretches of
22   stable reward probability (0.2/0.2, 0.8/0.8), or varying reward probability (0.2/0.8, 0.8/0.2). **c;** The
23   variability of imRPE is influenced by the reward probability conditions. **d**: Average d' for both delay
24   conditions. **e**: Mean pairwise difference in memory score between the "old" images and their semantically-

1  matched foil images. **f,g**: There is a positive association between reward probability of the currently
2  shown stimulus category and subsequent recognition memory, and a negative association between the
3  reward probability of the other stimulus category and memory. In panels d-f, colors indicate time between
4  encoding and memory testing; blue = no-delay, red = 24 delay.
5
6  To better estimate the effects of image category, other category, and play/pass behavior
7  on subsequent memory we fit memory score data with a modified version of the
8  hierarchical regression model that included separate reward probability terms for the
9  "image" and "other" categories. Posterior density estimates for the play/pass coefficient
10  were greater than zero (Fig 8a; table 1), replicating our previous finding. The posterior
11  density for the "image category" and "other category" probabilities was concentrated in
12  the region over which image category was greater than other category (mean [95% CI]
13  image category coefficient – other category coefficient: 0.052 [0.015,0.94]) and
14  supported independent and opposite contributions of both category probabilities (Fig 8b;
15  table 1). These results, in particular the negative effect of "other category" probability on
16  the subsequent memory scores, are more consistent with an effect of prediction error
17  than a direct effect of reward prediction itself. More generally, these results support the
18  hypothesis that reward prediction errors elicited at the time of image presentation
19  enhance the degree to which such images are encoded in episodic memory systems.
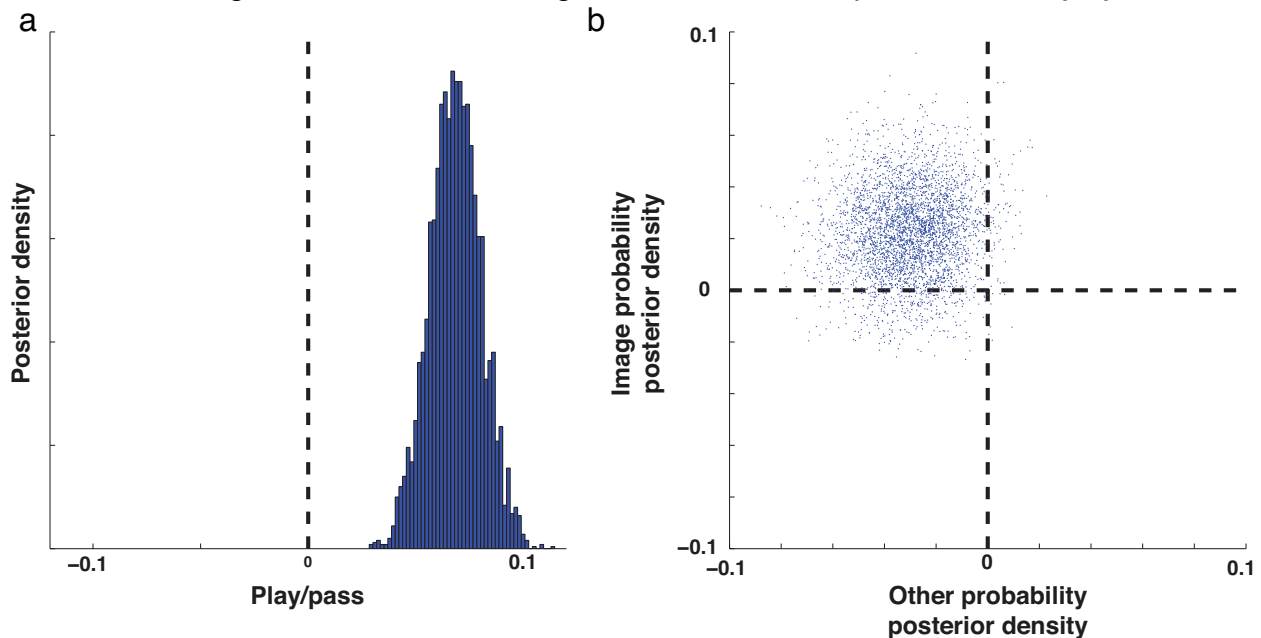


20
21  **Figure 8**: Memory scores in experiment two depend on subject gambling behavior and the probabilities
22  associated with both image categories. Memory score data from experiment two was fit with a version of
23  the hierarchical regression model described in figure 6A to replicate previous findings and determine
24  whether reward probability effects were attributable to both observed and unobserved category
25  probabilities. **a:** Posterior probability estimates of the mean play/pass coefficient were greater than zero
26  and consistent with those measured in the first experiment. **b:** Image category probability (observed)
27  coefficients are plotted against other category probability (unobserved) coefficients revealing that subjects
28  tended to have higher memory scores for images that were associated with high reward probabilities
29  (upward shift of density relative to zero) and when the unobserved image category was associated with a
30  low reward probability (leftward shift of density relative to zero).
31

1    Despite the general agreement between the two experiments, there was one noteworthy
2    discrepancy. While hierarchical models fit to both datasets indicated higher probability of
3    positive coefficients for the interaction between value and probability (e.g. positive
4    effects of probability on subsequent memory are greater for high value trials), the 95%
5    credible intervals for these estimates in experiment two included zero as a possible
6    coefficient value (table 1) indicating that the initial finding was not replicated in the
7    strictest sense.
8
9    To better understand this discrepancy, and make the best use of the data from both
10    experiments we extended the hierarchical regression approach to include additional
11    coefficients capable explaining differences across the two experiments and fit this
12    extended model to the combined data. As expected, this model provided evidence for a
13    memory advantage on play trials, and an amplification of this advantage for trials with a
14    high reward probability based on image category (Fig S4; table 1). Across the combined
15    dataset there was also positive effect of the interaction between value and probability
16    (Fig S4; table 1), supporting our initial observation in experiment one. Furthermore, we
17    observed that the reward probability effect was greater in the no delay condition (Fig S4;
18    table 1), with no evidence for any memory effects being stronger in the 24 hour delay
19    condition (table 1; all other delay difference ps >.19).
20
21
22    **Discussion**
23
24    An extensive prior literature has linked dopamine to reward prediction errors elicited
25    during reinforcement learning [5,6,19,39,49-51], and a much smaller literature has suggested
26    that dopamine can also influence the encoding and consolidation of episodic memories
27    by modulating activity in the medial temporal lobes [18,40,52]. To date, however, there has
28    been mixed evidence regarding the relationship between prediction error signaling and
29    memory encoding. Here we used a novel two-stage learning and memory paradigm
30    along with computational modeling to better characterize how prediction error signals
31    affect the strength of incidental memory formation.
32
33    We found that memory encoding was stronger for trials in which the subjects observed
34    an image that was associated with high reward probability (figure 5c,d). This effect was
35    only evident for trials in which subjects accepted the risky offer (consistent with the fact
36    that trial outcome was always zero on pass trials), was evident even after controlling for
37    other potential confounds (figure 6b column 3), and was amplified for trials in which
38    more points were on the line (figure 6b column 5). These results are all consistent with a
39    direct effect of reward prediction error at time of image presentation on memory
40    encoding (figure 5a,b). This interpretation is bolstered by evidence that individuals that
41    were more sensitive to value and probability in the decision making task showed reward
42    prediction error memory benefits to a greater degree. Experiment two further supported
43    the reward prediction error interpretation by demonstrating that memory benefits were
44    composed of equal and opposite contributions of the reward probability associated with

1    the observed image category and that of the unobserved, counterfactual, one (Figure

2    7f,g, Figure 8b). Together, these results provide evidence for the hypothesis that reward

3    prediction errors enhance the encoding of simultaneously presented incidental visual

4    information for subsequent memory.

5

6    We also found that subjects encoded memoranda to a greater degree on trials in which

7    they selected a risky bet (figure 3). This finding is consistent with a positive relationship

8    between prediction error signaling and memory strength, in that subject behavior

9    provides a proxy for the subjective reward probability estimates (figure 2A). However,

10   this behavioral effect was prominent in both experiments, even after controlling for

11   model-based estimates of reward prediction error (figures 6b & 8a). Therefore, while we

12   suspect that this result may at least partially reflect the direct impact of reward prediction

13   error, it may also reflect other factors associated with risky decisions. On play trials,

14   subjects view items while anticipating the uncertain gain or loss of points during the

15   upcoming feedback presentation whereas on pass trials, subjects are assured to

16   maintain their current score. The possibility that this difference in risk might contribute to

17   the subsequent memory effects observed for choice behavior would be consistent with

18   recent work showing that memoranda presented immediately prior to feedback are

19   better remembered if they preceded more uncertain feedback [53]. One potential

20   confound for these choice effects is the heightened state of attention that might occur

21   before receiving a more informative task outcome. While we are unable to rule this

22   possibility out completely, our study minimizes this possibility by presenting

23   counterfactual information on pass trials that is nearly identical to the experienced

24   outcome information. We find that subjects are slightly more influenced by outcome

25   information provided on play trials, suggesting that anticipatory attention might differ

26   somewhat between the two conditions (Fig S3a); however the degree of this difference

27   was small enough that model selection favored a model that did not distinguish between

28   play and pass trials for learning (Fig 2c). Furthermore, there was no relationship

29   between the degree to which subjects modulated learning from feedback on play versus

30   pass trials and the degree to which they showed subsequent memory improvements on

31   play trials (Fig S3b). Given that risky decisions tend to be preceded by higher levels of

32   dopamine, the possibility of a direct effect of risk taking on subsequent memory would

33   still be consistent with a dopaminergic mechanism [22-24,27].

34

35   This relationship between reward prediction errors and memory is consistent with a

36   broad literature highlighting the effects of dopaminergic signaling on hippocampal

37   plasticity [32,35,36] and memory formation [40] as well as an equally broad literature

38   suggesting that dopamine provides a reward prediction error signal [5,19] through

39   projections that extend both to the striatum and the hippocampus [17]. Our results support

40   the behavioral consequences that might be predicted to result from such mechanisms,

41   however they also refine them substantially. In particular, we show that the timing of

42   reward prediction error signaling relative to the memorandum is key; we saw no effect of

43   the reward prediction error elicited by prior or subsequent feedback on memory strength

44   (Figure 4a&b), despite strong evidence that this feedback was used to guide

1 reinforcement learning and decision making (Figure 2). Furthermore, we showed that
2 prediction error effects on memory emerged immediately after task performance and
3 were not enhanced after 24 hours (Figure 5a&b; Figure 6b column 3). These results are
4 somewhat at odds with previous literature suggesting that dopamine dependent memory
5 enhancement emerges only after an extended consolidation period [40]. It is unclear to
6 what extent we should expect generalization of these results to our study, given the
7 differences in experimental paradigm, timescale, memory demands, and species in the
8 two paradigms. However, our results open the door for future research to (1) directly
9 test whether prediction error driven memory enhancements are mediated by changes in
10 dopamine, and (2) characterize the conditions under which dopamine mediated
11 changes to memory encoding do and do not require a consolidation period.
12
13 More generally, our results provide insight into the apparent inconsistency in previous
14 behavioral studies that have attempted to link reward prediction error signals to memory
15 encoding. Consistent with previous work (e.g. [54]), our results emphasize the importance
16 of choice in the degree to which image valence contributed to memorability. Indeed, for
17 trials in which the subjects passively observed outcomes, we saw no relationship
18 between model derived reward prediction error estimates and subsequent memory
19 strength (Figure 5a&b dotted lines). This might help to explain the lack of a signed
20 relationship between reward prediction errors and subsequent memory strength in a
21 recent study by Rouhani and colleagues that leveraged a Pavlovian design that did not
22 require explicit choices to be made [44]. In contrast to our results, Rouhani and colleagues
23 observed a positive effect of absolute prediction error, similar to our model-based
24 surprise estimates, on subsequent memory. While we saw no effect of surprise on
25 subsequent memory, other work has highlighted a role for such signals as enhancing
26 hippocampal activation and memory encoding [55,56]. One potential explanation for this
27 discrepancy is in the timing of image presentation. Our study presented images only
28 briefly during the choice phase of the decision task. By contrast, Rouhani and
29 colleagues presented the memoranda for an extended period that also encompassed
30 the epoch containing trial feedback, potentially explaining why they observed effects
31 related to outcome surprise. More generally, the temporally selective effects of reward
32 prediction error observed here suggest that the reward prediction error effects may differ
33 considerably from other, longer timescale manipulations thought to enhance memory
34 consolidation through dopaminergic mechanisms [18,33,41,42].
35
36 Our results appear somewhat incompatible on first glance with those of Wimmer and
37 colleagues [43], who show that stronger prediction error encoding in the ventral striatum is
38 associated with weaker encoding of incidental information. We suspect that the
39 discrepancy between these results is driven by differences in the degree to which
40 memoranda are task relevant in the two paradigms. In our task, subjects were required
41 to encode the memoranda sufficiently to categorize them in order to perform the primary
42 decision-making task, whereas in the Wimmer study, memoranda were unrelated to the
43 decision task and thus might not be well-attended on all trials. Taken together, these
44 results suggest that reward prediction errors are most likely to enhance memory when

they are elicited by the memoranda themselves, with the potential influence of secondary tasks eliminated or at least tightly controlled.

In summary, our results demonstrate a role for reward prediction errors in prioritizing information for memory storage. We show that this role is temporally and computationally precise, independent of consolidation duration (at least in the current paradigm), and contingent on decision-making behavior. These data should help clarify inconsistencies in the literature regarding the relationship between reward learning and memory, and they make detailed predictions for future studies exploring the relationship between dopamine signaling and memory formation.

**Methods**

Experiment 1

**Experimental procedure**

The task consisted of two parts: the learning task and memory task. The learning task was a reinforcement learning task with random change-points in reward contingencies of the targets. The memory task was a surprise recognition memory task using image stimuli that were presented during the learning task and foils.

Subjects completed either the no delay or 24-hour delay versions of the task. In the no delay condition, the memory task followed the learning task only after a short break, during which a demographic survey was given. Therefore, the entire task was performed in one sitting. In the 24-hour delay condition, subjects returned 20-30 hours after completing the learning task to do the memory task.

**Subjects**

A total of 287 subjects (142, no delay condition; 145, 24hr-delay condition) completed the task via Amazon Mechanical Turk. From this, 88 subjects (33, no delay; 55, 24hr-delay) were excluded from analysis because they previously completed a prior version of the task or didn't meet our criteria of above-chance performance in the learning task. To determine whether a subject's performance was above-chance, we simulated random choices using the same task structure, then computed the total score achieved by the random performance. We then repeated such simulations 5000 times, and assessed whether the subject's score was greater than 5% of the score distribution from the simulations. The final sample had a total of 199 subjects (109, no delay, 90, 24hr-delay; 101 males, 98 females) with the age of 32.2 ± 8.5 (mean ± SD). Informed consent was obtained in a manner approved by the Brown University Institutional Review Board.

**Learning task**

The learning task consisted of 160 trials, where each trial consisted of three phases – value, image, and feedback (Figure 1a). During the value phase, the amount of reward associated with the current trial was presented in the middle of the screen for 1.5 s. This value was equally sampled from [1, 5, 10, 20, 100]. After an interstimulus interval (ISI) of 0.5 s, an image appeared in the middle of the screen for 3 s (image phase). During the image phase, the subject made one of two possible responses using the keyboard: PLAY (press 1) or PASS (press 0). When a response is made, a colored box indicating the subject's choice (e.g. black = play, white = pass) appeared around the image. The pairing of box color and subject choice was pseudorandomized across subjects. After this image phase, an ISI of 0.5 s followed, after which the trial's feedback was shown (feedback phase).

Each trial had an assigned reward probability, such that if the subject chose PLAY, they would be rewarded according to that probability. If the subject chose PLAY and the trial was rewarding, they were rewarded by the amount shown during the value phase (Figure 1a). If the subject chose PLAY but the trial was not rewarding, they lost 10 points regardless of the value of the trial. If the choice was PASS, the subject neither earned nor lost points (+0), and was shown the hypothetical result of choosing PLAY (Figure 1a). During the feedback phase, the reward feedback (+value, -10, or hypothetical result) was shown for 1.5s, followed by an ISI (0.5 s), and a 1 s presentation of the subject's total accumulated score.

All image stimuli belonged to one of two categories: animate (e.g. whale, camel) and inanimate (e.g. desk, shoe). Each image belonged to a unique exemplar, such that there were no two images of the same animal or object. Images of the two categories had reward probabilities that were oppositely yoked. For example, if the living category has a reward probability of 90%, the non-living category had a reward probability of 10%. Therefore, the subjects only had to learn the probability for one category, and simply assume the opposite probability for the other category.

The reward probability for a given image category remained stable until a change-point occurred, after which it changed to a random value between 0 and 1 (Figure 1d). Change-points occurred with a probability 0.16 on each trial. To facilitate learning, change-points did not occur in the first 20 trials of the task and the first 15 trials following a change-point. Each subject completed a unique task with pseudorandomized order of images that followed these constraints.

The objective was to maximize the total number of points earned. Subjects were advised to pay close attention to the value, probability, and category of each trial in order to decide whether it is better to PLAY or PASS. Subjects were thoroughly informed about the possibility of change-points, and that the two categories were oppositely yoked. They underwent a practice learning task in which the reward probabilities for the two categories were 1 and 0 to clearly demonstrate these features of the task. Subjects were awarded a bonus compensation proportional to the total points earned during the learning and memory tasks.

**Memory task**

1  During the memory task, subjects viewed 160 "old" images from the learning task
2  intermixed with 160 "new" images (Figure 1c). Images were selected such that there
3  was a new image for each unique exemplar from the learning task. This was to ensure
4  that subjects had to make judgments about the actual image itself, rather than the fact
5  that they saw a particular exemplar (e.g. "I remember seeing THIS desk" vs. "I
6  remember seeing A desk").
7      The order of old and new images was pseudorandomized. On each trial, a single
8  image was presented, and the subject selected between OLD and NEW by pressing 1
9  or 0 on the keyboard, respectively (Figure 1c). Afterwards, they were asked to rate their
10 confidence in the choice from 1 (Guess) to 4 (Completely certain). Subjects were not
11 provided with correct/incorrect feedback on their choices.
12
13

14 **Bayesian optimal learning model**
15 The optimal learning model computed inferences over the probability of a binary
16 outcome that evolves according to a change-point process. The model was given
17 information about the true probability of a change-point occurring on each trial (H;
18 hazard rate) by dividing the number of change-points by the total number of trials for
19 each subject. For each trial, a change-point was sampled according to a Bernoulli
20 distribution using the true hazard rate ($CP \sim B(H)$). If a change-point did not occur (CP =
21 0), the predicted reward rate ($\mu_t$) was updated from the previous trial ($\mu_{t-1}$). When a
22 change-point did occur (CP = 1), $\mu_t$ was sampled from a uniform distribution between 0
23 and 1. The posterior probability of each trial's reward rate given the previous outcomes
24 can be formulated as follows:
25

$$p(\mu_t|X_{1:t}) \propto p(X_t|\mu_t) \sum_{CP_t} \sum_{\mu_{t-1}} p(\mu_t|CP_t, \mu_{t-1})p(CP_t)p(\mu_{t-1}|X_{1:t})p(X_{1:t-1}) \tag{1}$$

26 Where $p(X_t|\mu_t)$ is the likelihood of the outcomes given the predicted reward rate,
27 $p(\mu_t|CP_t, \mu_{t-1})$ represents the process of accounting for a possible change-point (when
28 CP = 1, $\mu_t \sim U(0,1)$), $p(CP_t)$ is the hazard rate, and $p(\mu_{t-1}|X_{1:t})$ is the prior belief of the
29 reward rate.
30     Using the model-derived reward rate, we quantified the extent to which each new
31 outcome influenced the subsequent prediction as the learning rate in a delta-rule:
32

$$B_{t+1} = B_t + \alpha_t \delta_t$$
$$\delta_t = X_t - B_t \tag{2}$$

33
34 where B is the belief about the current reward rate, α is the learning rate, and δ is the
35 prediction error, defined as the difference between the observed (X) and predicted (B)
36 outcome. Rearranging, we were able to compute the trial by trial learning rate:
37

$$\alpha = \frac{B_{t+1} - B_t}{X_t - B_t} \tag{3}$$

1    Trial by trial modulation of change-point probability (i.e. surprise) was calculated by
2    marginalizing over $\mu_t$:
3

$$p(CP_t|X_{1:t}) \propto p(X_t|\mu_t) \sum_{\mu_t} \sum_{\mu_{t-1}} p(\mu_t|CP_t, \mu_{t-1})p(CP_t)p(\mu_{t-1}|X_{1:t})p(X_{1:t-1}) \tag{4}$$

4
5    Uncertainty was determined by calculating the entropy of a discrete random variable X
6    (i.e. reward rate) with possible values $\{x_1, x_2, \ldots x_i\}$ for a finite sample (Shannon, 2001):
7

$$H(X) = - \sum_i P(x_i)\ln(p(x_i)) \tag{5}$$

8
9
10
11
12
13    **Descriptive analysis**
14
15    Memory scores for each image were computed by transforming the recognition and
16    confidence reports provided by the subject. On each trial of the recognition memory
17    task, subjects first chose between "old" and "new", then reported their confidence in that
18    choice on a scale of 1-4. We converted these responses so that choosing "old" with the
19    highest confidence (4) was a score of 8, while choosing "new" with the highest
20    confidence was a score of 1. Similarly, choosing "old" with the lowest confidence (1)
21    was a score of 5, while choosing "new" with the lowest confidence was a score of 4. As
22    such, memory scores reflected a confidence-weighted measure of memory strength
23    ranging from 1 to 8. These memory scores were used for all analyses involving
24    recognition memory.
25
26    Statistical analyses were performed in a between-subject manner. For each subject, we
27    computed the mean memory score of each trial type in question, then subtracted the
28    overall average memory score of the subject. Therefore, the memory scores used in our
29    analyses reflect the degree to which a certain trial condition led to better or worse
30    subsequent memory compared to average performance within each subject.
31
32    Relationships between computational factors and memory scores were assessed by
33    estimating the slope of the relationship between each computational factor and the
34    subsequent memory score separately for each computational variable and subject.
35    Statistical testing was performed using one sample t-tests on the regression coefficients

1 across subjects (for overall effects) and two sample t-tests for differences between
2 delay conditions (for delay effects). Regression coefficients (slopes) for individual
3 subjects were related to individual differences in decision making task performance
4 (coefficients from a GLM describing subject choices (play/pass) in terms of reward
5 probability and trial value) using spearman rank order correlation.
6
7 To generate the descriptive figures, we performed a binning procedure for each subject
8 to ensure that each point on the x axis contained an equal number of elements. For
9 each subject, we divided the y variable in question into quartiles and used the mean y
10 value of each quartile as the binned value. To plot data from all subjects on the same x
11 axis, we first determined the median x value for each bin per subject, then took the
12 average of the four bin median values across subjects. For figures containing more than
13 one plot, we shifted the x values of each plot slightly off-center to avoid overlap of
14 points.
15
16 A set of reinforcement learning models were fit directly to the subject behavior using a
17 constrained search algorithm (fmincon in Matlab) that maximized the total log posterior
18 probability of betting behavior given the optimal reward probability estimates, trial
19 values, and prediction errors (Fig c&d). All models contained four fixed parameters that
20 affected choice behavior: 1) a temperature parameter of the softmax function used to
21 convert trial expected values into action probabilities, 2) a value exponent term that
22 scales the relative importance of the trial value in making choices, 3) a play bias term to
23 indicate a tendency to attribute higher value to gambling behavior, and 4) an intercept
24 term for the effect of learning rate on choice behavior. The value estimated from
25 gambling on a given trial was given by:
26

$$V(t) = playBias + (P_{rew} * trialValue^{valExp} + (1 - P_{rew}) * (-10)^{valExp})$$

27
28 Where playBias is the play bias term, valExp is the value exponent, and $P_{rew}$ is the
29 reward probability inferred from the optimal model. The model fit with the above fixed
30 parameters (the base model) was then compared to models that contained additional
31 parameters that may affect trial-to-trial modulation of learning rate, including surprise,
32 uncertainty, learning rate computed from the optimal Bayesian model, and subject
33 choice behavior (play versus pass). In particular, learning rates were controlled through
34 a logistic function of a weighted predictor matrix that included an intercept term as well
35 as additional terms to capture the degree to which learning changed according to other
36 factors. Maximum likelihood weights for each predictor in the matrix (as listed above)
37 were estimated using gradient decent (fmincon in matlab) simultaneous with estimating
38 the decision related parameters described above. The best fitting model was
39 determined by computing the Bayesian information criterion (BIC) for each model, then
40 comparing these values to that of the base model [57]. Weak priors favoring normative
41 learning parameters were used to regularize parameter estimates for parameter
42 estimation but not model selection.
43

1   To compare subject behavior to model-predicted behavior, we simulated choice
2   behavior using the model with the lowest BIC, which incorporated surprise and
3   uncertainty variables in determining learning rate (Fig 2b). On each trial, we used the
4   expected trial value (V(t)) computed above, and the parameter estimates of the
5   temperature variable as inputs to a softmax function to generate choices.
6
7
8
9   **Hierarchical regression model**
10
11  Subject memory scores were modeled using a hierarchical mixture model that assumed
12  that the memory score reported for each item and subject would reflect a linear
13  combination of subject level predictors and item level memorability (Figure 6A). The
14  hierarchical model was specified in STAN (http://mc-stan.org) using the matlabSTAN
15  interface (http://mc-stan.org) [58]. In short, memory scores on each trial were assumed to
16  be normally distributed with a variance that was fixed across all trials for a given subject.
17  The mean of the memory score distribution on a given trial depended on 1) a trial-to-trial
18  task predictors that were weighted according to coefficients estimated at the subject
19  level and 2) item-to-item predictors that were weighted by coefficients estimated across
20  all subjects. Subject coefficients for each trial-to-trial task predictor were assumed to be
21  drawn from a group distribution with a mean and variance offset by a delay variable,
22  which allowed the model to capture differences in coefficient values for the two different
23  delay conditions. All model coefficients were assumed to be drawn from prior
24  distributions and for all coefficients other than the intercept (which captured overall
25  memory scores) prior distributions were centered on zero. The code used to specify the
26  hierarchical model is included as supplementary code.
27
28
29
30  Experiment 2
31
32  **Experimental procedure**
33
34  In experiment 2, the learning task was modified to dissociate reward rate from reward
35  prediction error. The reward probability of the two image categories (living vs. nonliving)
36  were independent and set to either 0.8 or 0.2, allowing for a 2x2 design (0.8/0.8, 0.8/0.2,
37  0.2/0.8, 0.2/0.2; Figure 7a). Change-points occurred with a probability 0.05 on every trial
38  for the two categories independently. Change points did not occur for the first 20 trials of
39  the task and the first 20 trials following a change point. Tasks were generated to contain
40  at least one block of each trial type in the 2x2 design. Each subject completed a unique
41  task with pseudorandomized order of images that followed these constraints. The task
42  instructions explicitly stated that the two image categories had independent reward
43  probabilities that need to be tracked separately. The rest of the task, including the
44  recognition memory portion, was identical to that of Experiment 1.

1
## Subjects
3  A total of 279 subjects (157, no delay condition; 122, 24hr-delay condition) completed
4  the task via Amazon Mechanical Turk. 105 subjects (64, no delay; 41, 24hr-delay) were
5  excluded from analysis because they previously completed a prior version of the task or
6  didn't meet our criteria of above-chance performance in the learning task. Therefore, the
7  final sample had a total of 174 subjects (93, no delay, 81, 24hr-delay; 101 males, 71
8  females, 2 no response) with the age of 34.0 ± 9.1 (mean ± SD). Informed consent was
9  obtained in a manner approved by the Brown University Institutional Review Board.

## Data availability
24 The data from both experiments and the scripts used to analyze and model the data are
25 available from the authors upon request.

## Author Contributions
28 All authors designed the experiment and wrote the manuscript. AIJ collected the data
29 and MRN developed the computational mdoels. MRN and AIJ designed and performed
30 behavioral analysis.

## Competing Interests
33 None of the authors have any competing interests to report, financial or otherwise.

## Acknowledgements

1
2
3
4
5
6
7
8
9

**Reference:**

1.  Sutton, R. & Barto, A. *Reinforcement learning: An introduction.* (MIT Press, 1998).
2.  Schacter, D. L. & Tulving, E. *Memory Systems 1994.* (MIT Press, 1994).
3.  McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* **102,** 419–457 (1995).
4.  O'Reilly, R. C. & Rudy, J. W. Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychological Review* **108,** 311–345 (2001).
5.  Schultz, W. A Neural Substrate of Prediction and Reward. *Science* **275,** 1593–1599 (1997).
6.  Frank, M. J. By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. *Science* **306,** 1940–1943 (2004).
7.  O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304,** 452–454 (2004).
8.  Sohal, V. S., Zhang, F., Yizhar, O. & Deisseroth, K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459,** 698–702 (2009).
9.  Squire, L. R. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review* **99,** 195–231 (1992).
10. Nadel, L. & Moscovitch, M. Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology* **7,** 217–227 (1997).
11. Wan, H., Aggleton, J. P. & Brown, M. W. Different contributions of the hippocampus and perirhinal cortex to recognition memory. *J. Neurosci.* **19,** 1142–1148 (1999).
12. Eichenbaum, H., Yonelinas, A. P. & Ranganath, C. The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* **30,** 123–152 (2007).
13. Heilbronner, S. R., Rodriguez-Romaguera, J., Quirk, G. J., Groenewegen, H. J. & Haber, S. N. Circuit-Based Corticostriatal Homologies Between Rat and Primate. *Biol. Psychiatry* **80,** 509–521 (2016).
14. Thierry, A. M., Gioanni, Y., Dégénétais, E. & Glowinski, J. Hippocampo-prefrontal cortex pathway: anatomical and electrophysiological characteristics.

*Hippocampus* **10,** 411–419 (2000).

15. Atallah, H. E., Frank, M. J. & O'Reilly, R. C. Hippocampus, cortex, and basal ganglia: insights from computational models of complementary learning systems. *Neurobiol Learn Mem* **82,** 253–267 (2004).

16. Eichenbaum, H. Prefrontal–hippocampal interactionsin episodic memory. *Nature Reviews Neuroscience* **18,** 547–558 (2017).

17. Floresco, S. B. Dopaminergic regulation of limbic-striatal interplay. *J Psychiatry Neurosci* **32,** 400–411 (2007).

18. Shohamy, D. & Adcock, R. A. Dopamine and adaptive memory. *Trends in Cognitive Sciences* **14,** 464–472 (2010).

19. Montague, P. R., Dayan, P. & Sejnowski, T. J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16,** 1936–1947 (1996).

20. Steinberg, E. E. *et al.* A causal link between prediction errors, dopamine neurons and learning. *Nature Publishing Group* **16,** 966–973 (2013).

21. Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews* **28,** 309–369 (1998).

22. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* (2003).

23. Dagher, A. The role of dopamine in risk taking: a specific look at Parkinson's disease and gambling. 1–12 (2014). doi:10.3389/fnbeh.2014.00196/abstract

24. Stopper, C. M., Tse, M. T. L., Montes, D. R., Wiedman, C. R. & Floresco, S. B. Overriding Phasic Dopamine Signals Redirects Action Selectionduring Risk/Reward Decision Making. *Neuron* 1–13 (2014). doi:10.1016/j.neuron.2014.08.033

25. Collins, A. G. E. & Frank, M. J. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review* **121,** 337–366 (2014).

26. Zalocusky, K. A. *et al.* Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature* **531,** 642–646 (2016).

27. Rutledge, R. B., Skandali, N., Dayan, P. & Dolan, R. J. Dopaminergic Modulation of Decision Making and Subjective Well-Being. *Journal of Neuroscience* **35,** 9811–9822 (2015).

28. Bouret, S. & Sara, S. J. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences* **28,** 574–582 (2005).

29. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46,** 681–692 (2005).

30. Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience* **15,** 1040–1046 (2012).

31. McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* **84,** 870–881 (2014).

32. Lisman, J. E. & Grace, A. A. The Hippocampal-VTA Loop: Controlling the Entry of

Information into Long-Term Memory. *Neuron* **46,** 703–713 (2005).

33. Wittmann, B. C. *et al.* Reward-Related fMRI Activation of Dopaminergic Midbrain Is Associated with Enhanced Hippocampus- Dependent Long-Term Memory Formation. *Neuron* **45,** 459–467 (2005).

34. Tully, K. & Bolshakov, V. Y. Emotional enhancement of memory: how norepinephrine enables synaptic plasticity. *Mol Brain* **3,** 15 (2010).

35. Rosen, Z. B., Cheung, S. & Siegelbaum, S. A. Midbrain dopamine neurons bidirectionally regulate CA3-CA1 synaptic drive. *Nature Publishing Group* **18,** 1763–1771 (2015).

36. Lemon, N. & Manahan-Vaughan, D. Dopamine D1/D5 Receptors Gate the Acquisition of Novel Information through Hippocampal Long-Term Potentiation and Long-Term Depression. *Journal of Neuroscience* **26,** 7723–7729 (2006).

37. McNamara, C. G., Tejero-Cantero, Á., Trouche, S., Campo-Urriza, N. & Dupret, D. Dopaminergic neurons promotehippocampal reactivation andspatial memory persistence. *Nature Publishing Group* **17,** 1658–1660 (2014).

38. Gan, J. O., Walton, M. E. & Phillips, P. E. M. Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature Neuroscience* **13,** 25–27 (2009).

39. Bayer, H. M. & Glimcher, P. W. Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron* **47,** 129–141 (2005).

40. Bethus, I., Tse, D. & Morris, R. G. M. Dopamine and Memory: Modulation of the Persistence of Memory for Novel Hippocampal NMDA Receptor-Dependent Paired Associates. *Journal of Neuroscience* **30,** 1610–1618 (2010).

41. Murty, V. P. & Adcock, R. A. Enriched Encoding: Reward Motivation Organizes Cortical Networks for Hippocampal Detection of Unexpected Events. *Cerebral Cortex* **24,** 2160–2168 (2013).

42. Murty, V. P., Tompary, A., Adcock, R. A. & Davachi, L. Selectivity in Postencoding Connectivity with High-Level Visual Cortex Is Associated with Reward-Motivated Memory. *Journal of Neuroscience* **37,** 537–545 (2017).

43. Wimmer, G. E., Braun, E. K., Daw, N. D. & Shohamy, D. Episodic Memory Encoding Interferes with Reward Learning and Decreases Striatal Prediction Errors. *Journal of Neuroscience* **34,** 14901–14912 (2014).

44. Rouhani, N., Norman, K. A. & Niv, Y. Dissociable effects of surprising rewards on learning and memory. *J Exp Psychol Learn Mem Cogn* (2018). doi:10.1037/xlm0000518

45. Preuschoff, K. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. 1–12 (2011). doi:10.3389/fnins.2011.00115/abstract

46. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nature Neuroscience* **10,** 1214–1221 (2007).

47. Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience* **30,** 12366–12378 (2010).

48. Nassar, M. R. *et al.* Age differences in learning emerge from an insufficient

representation of uncertainty in older adults. *Nature Communications* **7,** 11609 (2016).

49. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482,** 85–88 (2012).

50. Rutledge, R. B., Dean, M., Caplin, A. & Glimcher, P. W. Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *Journal of Neuroscience* **30,** 13525–13536 (2010).

51. Sadacca, B. F., Jones, J. L. & Schoenbaum, G. Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *Elife* **5,** (2016).

52. Frey, U. & Morris, R. G. Synaptic tagging: implications for late maintenance of hippocampal long-term potentiation. *Trends in Neurosciences* **21,** 181–188 (1998).

53. Stanek, J. K., Dickerson, K. C., Chiew, K. S., Clement, N. J. & Adcock, R. A. Expected reward value and reward uncertainty have temporally dissociable effects on memory formation. doi:10.1101/280164

54. Koster, R., Guitart-Masip, M., Dolan, R. J. & Düzel, E. Basal Ganglia Activity Mirrors a Benefit of Action and Reward on Long-Lasting Event Memory. *Cerebral Cortex* **25,** 4908–4917 (2015).

55. Long, N. M., Lee, H. & Kuhl, B. A. Hippocampal Mismatch Signals Are Modulated by the Strength of Neural Predictions and Their Similarity to Outcomes. *Journal of Neuroscience* **36,** 12677–12687 (2016).

56. Greve, A., Cooper, E., Kaula, A., Anderson, M. C. & Henson, R. Journal of Memory and Language. *Journal of Memory and Language* **94,** 149–165 (2017).

57. Schwartz, G. Estimating the Dimension of a Model. *The Annals of Statistics* **6,** 461–464 (1978).

58. Carpenter, B. *et al.* Stan: A Probabilistic Programming Language. *J. Stat. Soft.* **76,** (2017).

1
2
3
4
5
6
7
8
9
10   Supplementary figures:
11
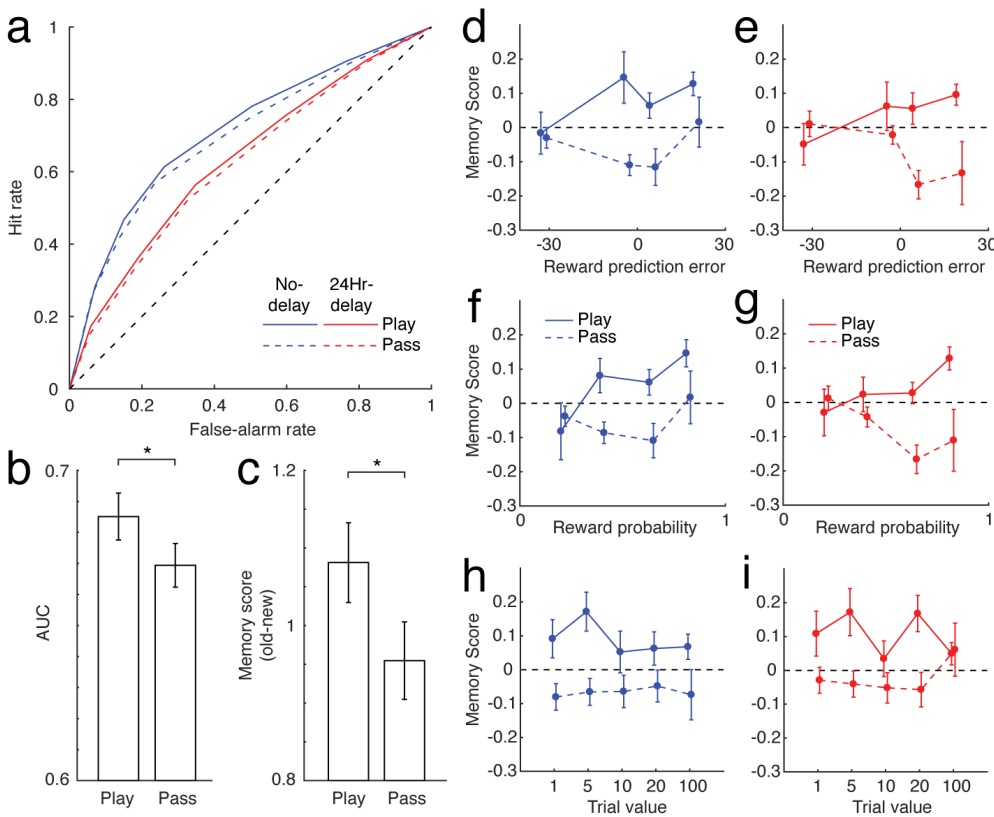


12
**Supplementary figure 1** Replication of main effects after omitting all trials in which subsequent memory
had low confidence ("guess"). To rule out the possibility that low confidence trials are driving our results or
adding unnecessary noise, we repeated the main findings of the study after omitting all trials in which the
confidence score was 1 (both target and foil). (a) The ROC curves for play/pass trials. (b) Area under the
ROC curve was greater for play versus pass (t(198) = 2.78, p = 0.0060; group difference, t(197) = -0.083,
p = 0.93). (c) Mean pairwise difference in memory score between the "old" images and their semantically-
matched foil images was greater for play versus pass (t(197) = 3.49, p < 0.001; group difference, t(196) =
0.89, p = 0.37). (d,e) Positive relationship between imRPE and subsequent memory (t(198)=2.48,
p=0.014; group difference, t(197)=-0.33, p=0.74). (f,g) Positive relationship between reward probability
and subsequent memory (t(198)=2.99, p=0.0031, group difference, t(197)=0.078, p=0.94). (h,i) No
relation between reward value and subsequent memory (t(198)=-1.74, p=0.084, group difference,
t(197)=0.82, p=0.41).

25
26
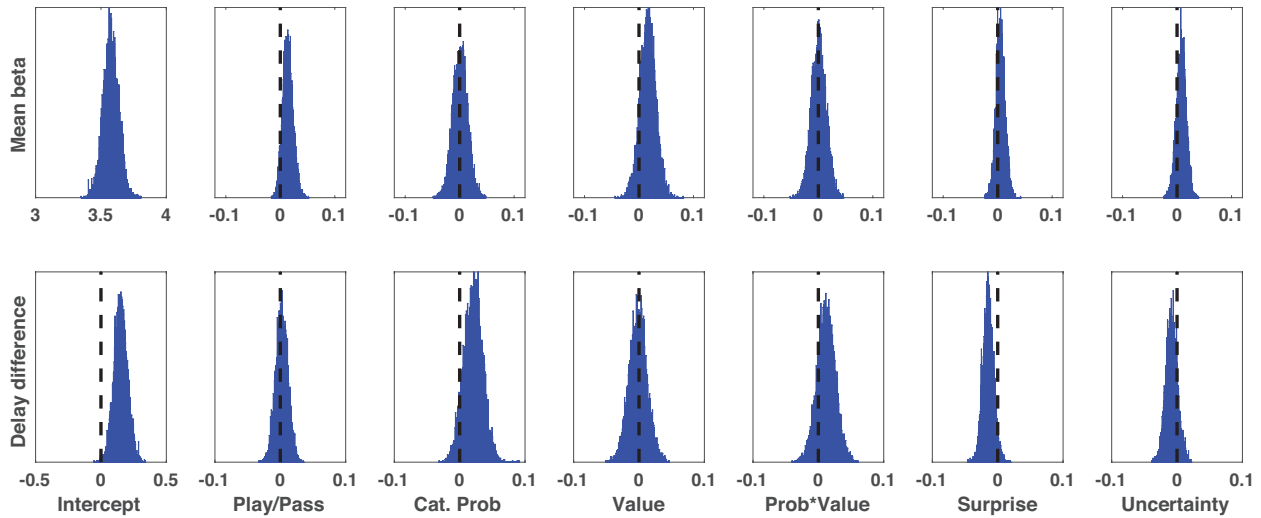
**Supplementary figure 2**: Hierarchical regression estimating memory scores for foil items. Posterior probability densities for mean predictor coefficients ($\mu_x$; top row) and delay condition parameter difference ($D_x$; bottom row) estimated through MCMC sampling over the graphical model described in figure 6A informed by data for foil items semantically matched to the images presented in the decision making task. Unlike fits to the target items (Figure 6B), coefficients for task predictors related to reward probability and play/pass did not deviate appreciably from zero. However, delay condition difference coefficients were positive for the intercept term, indicating that subjects in the 24 hour delay condition tended to report higher memory scores for foils, consistent with the poorer discriminability in the 24 hour delay condition (see figure 3C).
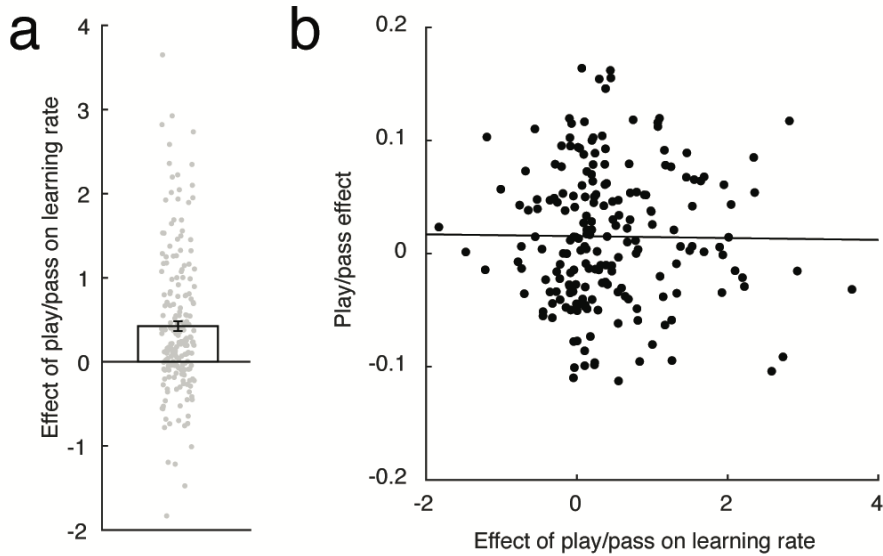
**Supplementary figure 3** Subject learning rates differed across choice conditions but this difference did not predict the strength of choice effects on subsequent memory. **a:** Parameter estimates quantifying the degree to which learning rate is modulated by subject play/pass decisions in the choice model (positive values indicate more learning on after play decisions) for all subjects in experiment 1 (gray points) and the mean/SEM values across subjects (bar/error line; mean parameter estimate = 0.42, t = 7.18, p < 0.001). **b:** The degree to which subjects (black points) adjusted learning rate according to play/pass behavior (abscissa) showed no association with the degree to which they enhanced memory of items presented on play trials (ordinate; spearman rho = 0.045, p = 0.53).
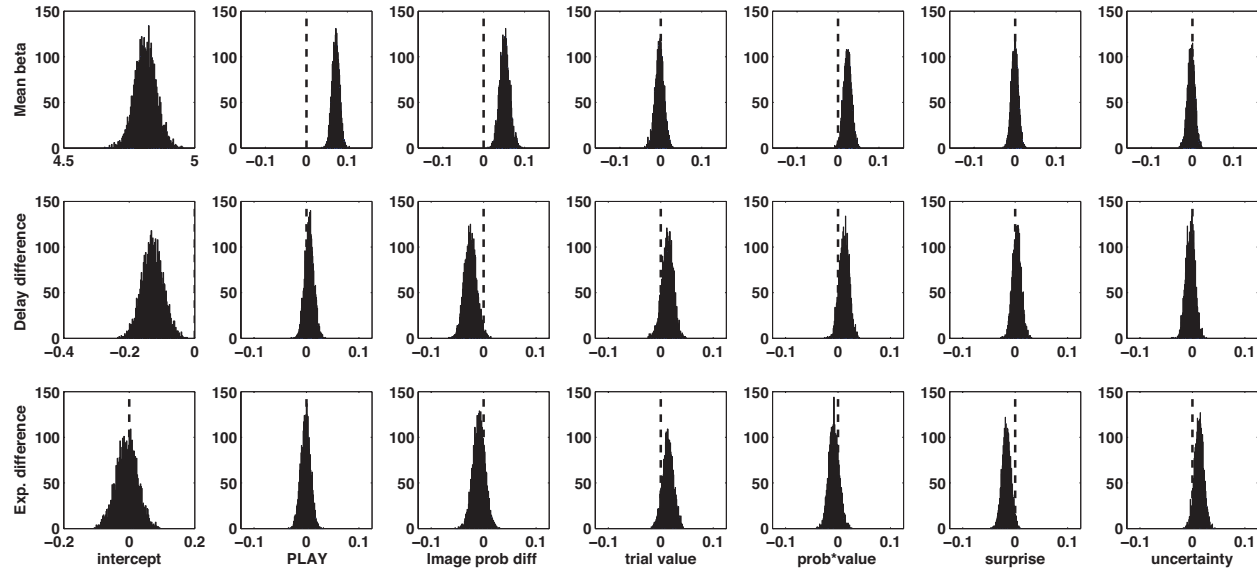
1
2 **Supplementary figure 4** Coefficient estimates for hierarchical model fit to all subjects across both
3 experiments. **Top:** Posterior probability density over mean coefficient estimates at the population level for
4 each parameter in the hierarchical model fit to subject memory scores. Leftmost column reflects the
5 intercept indicating overall memory scores for old items, whereas all other columns reflect the degree to
6 which learning task-related factors affected subsequent memory. The model included two separate terms
7 to model the probability associated with the shown image category ("cat prob(play)"; third column) and the
8 probability associated with the non-displayed image category ("non-cat prob(play"; fourth column) despite
9 the fact that these two terms were perfectly anti-correlated for all participants who completed experiment
10 1. **Middle:** Posterior probability density over delay difference estimates for each parameter in the
11 hierarchical model. **Bottom:** Posterior probability density over experiment difference estimates quantifying
12 the difference in coefficient values across the two experiments for each parameter in the hierarchical
13 model.
14
15
16
17
18
19

```
%% Specify hierarchical model:
mem_code = {
    'data {'
    '    int<lower=0> S; '     % Number of subjects
    '    int<lower=0> TPS; '   % trials per subject
    '    int<lower=0> N; '     % Number of trials
    '    int<lower=0> Ps; '    % number of predictors
    '    int<lower=0> Is; '    % number of items/foils
    '    vector[TPS] memScore[S]; // memory score 1-8 ' % single array of↵
vectors
    '    matrix[N,Ps] predMat; // predictor matrix '    % predictor matrix
    '    matrix[N,Is] itemPredMat; // item predictor matrix ' % predictor↵
matrix
    '    int delay[S];'
    '}'
    'transformed data {' % ok, wouldn't have thought this necessary, send↵
data in two dimensions and convert to 3...
    'matrix[TPS,Ps] predMatSub[S];'
    'matrix[TPS,Is] itemPredMatSub[S];'
    '    for (t in 1:TPS) {'
    '        for (s in 1:S) {'
    '            for (p in 1:Ps) {'
    '                predMatSub[s, t, p]<-predMat[(s-1)*TPS+t, p];'
    '            }'
    '            for (i in 1:Is) {'
    '                itemPredMatSub[s, t,i]<-itemPredMat[(s-1)*TPS+t, i];'
    '            }'
    '        }'
    '    }'
    '}'
    'parameters {'    % OK, we need mean and variance of coefficients
    '    vector[Ps] group_beta_mu; '
    '    vector[Ps] delayBetaOffset; '  % how much does group beta get↵
pushed around by delay condition?
    '    vector <lower=0> [Ps]  group_beta_sigma; '
    '    vector[Ps] sub_betas[S]; '
    '    real <lower=0> item_sigma;'
    '    real <lower=0> group_gammaScale;'
    '    real <lower=0> group_gammaShape;'
    '    real <lower=0> sub_sigmas[S]; '
    '    vector[Is] itemBs; '
    '}'
    'transformed parameters {'
    '    real <lower=0> sub_sigmaSquared[S]; '
    '    'for (s in 1:S) {'
    '        'sub_sigmaSquared[s]<-pow(sub_sigmas[s],2);'
    '    }'
    '}'
    'model {'
    '    group_beta_mu[1] ~ normal(4.5, 2);'       % Group mean for intercept↵
term -- somewhere around avg memory score
    '    group_beta_sigma[1] ~ gamma(1, 1);'       % Group sigma for↵
intercept term
    '    delayBetaOffset[1] ~ cauchy(0, 1);'   % relatively stiff prior↵
over the conditional difference.
    '    group_gammaScale ~ cauchy(0,3);'          % gamma distribution over↵
subject variance gamma distribution (Scale parameter)
    '    group_gammaShape ~ cauchy(0,3);'          % gamma distribution over↵
subject variance gamma distribution (Shape parameter)
    '    item_sigma ~ cauchy(0,1);'                % prior over variance
```

1
2

```
'    for (p in 2:Ps) {'
'        group_beta_mu[p] ~ normal(0, 2);'     % prior on group mean beta⬎
for each predictor -- set to zero with some variance
'        group_beta_sigma[p] ~ gamma(1, 1);'   % prior on group sigma for⬎
each predictor -- set wide gamma dist
'        delayBetaOffset[p] ~ cauchy(0, .3);'  % Set stiff prior on the⬎
delay condition effects for each parameter.
'    }'
'    for (i in 1:Is) {'
'        itemBs[i] ~ normal(0, item_sigma);'   % mean item effect is⬎
zero, variance half cauchy distributed.
'    }    %
'    for (s in 1:S) {'
'        for (p in 1:Ps) {'
'            sub_betas[s,p] ~ normal(group_beta_mu[p]+ delay[s] *⬎
delayBetaOffset[p] , group_beta_sigma[p]);'
'        }'
'        sub_sigmaSquared[s] ~ gamma(group_gammaShape,⬎
group_gammaScale);'
'        memScore[s] ~ normal(predMatSub[s] * sub_betas[s] +⬎
itemPredMatSub[s] * itemBs  , sub_sigmas[s]);'
'    }'
'}'
};

% Need to add sigmas for each subject...

fit = stan('model_code',mem_code,'data', mem_data, 'verbose', true, 'init',⬎
0, 'warmup', 1000, 'chains', 4, 'iter', 2000, 'algorithm', 'NUTS');

print(fit);
```

1
2    **Supplementary code**: Hierarchical regression model for experiment 1 spefified in matlabStan. All code
3    and data will be available by the authors upon request.
4
5
6
7
8
9
10
11
12
13
14
15
16