Report

Neuron

An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence

Highlights

- New evidence for a role for the hippocampus in reinforcement learning in adolescents
- Enhanced cooperation between multiple learning systems in the adolescent brain
- Associations between learning and memory in behavior and brain for adolescents

Authors

Juliet Y. Davidow, Karin Foerde, Adriana Galván, Daphna Shohamy

Correspondence

jdavidow@fas.harvard.edu (J.Y.D.), ds2619@columbia.edu (D.S.)

In Brief

Davidow et al. discover adaptive consequences of reward sensitivity in adolescence. Adolescents showed better reinforcement learning and enhanced memory for positive feedback events, both related to prediction error-related activation in the hippocampus and greater hippocampal-striatal functional connectivity.

Davidow et al., 2016, Neuron *92*, 93–99
October 5, 2016 © 2016 Elsevier Inc.
http://dx.doi.org/10.1016/j.neuron.2016.08.031

CellPress

Neuron
Report

An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence

Juliet Y. Davidow,[1,5,*] Karin Foerde,[2,3] Adriana Galván,[4] and Daphna Shohamy[5,6,7,*]
[1]Department of Psychology, Harvard University, Cambridge, MA 02138, USA
[2]Department of Psychiatry, Columbia University Medical Center, New York, NY 10032, USA
[3]New York State Psychiatric Institute, New York, NY 10032, USA
[4]Department of Psychology, University of California, Los Angeles, CA 90095, USA
[5]Department of Psychology, Columbia University, New York, NY 10027, USA
[6]Zuckerman Mind Brain Behavior Institute and Kavli Institute for Brain Science, Columbia University, New York, NY 10027, USA
[7]Lead Contact
*Correspondence: jdavidow@fas.harvard.edu (J.Y.D.), ds2619@columbia.edu (D.S.)
http://dx.doi.org/10.1016/j.neuron.2016.08.031

SUMMARY

Adolescents are notorious for engaging in reward-seeking behaviors, a tendency attributed to heightened activity in the brain's reward systems during adolescence. It has been suggested that reward sensitivity in adolescence might be adaptive, but evidence of an adaptive role has been scarce. Using a probabilistic reinforcement learning task combined with reinforcement learning models and fMRI, we found that adolescents showed better reinforcement learning and a stronger link between reinforcement learning and episodic memory for rewarding outcomes. This behavioral benefit was related to heightened prediction error-related BOLD activity in the hippocampus and to stronger functional connectivity between the hippocampus and the striatum at the time of reinforcement. These findings reveal an important role for the hippocampus in reinforcement learning in adolescence and suggest that reward sensitivity in adolescence is related to adaptive differences in how adolescents learn from experience.

INTRODUCTION

Adolescents are highly sensitive to reward (Andersen et al., 1997; Brenhouse et al., 2008; Galván et al., 2006; Somerville and Casey, 2010; van Duijvenvoorde et al., 2014), which has been linked to the emergence of maladaptive behaviors (Brenhouse and Andersen, 2011; Galván, 2013; Spear, 2000). It has been suggested that this reward sensitivity may also be adaptive by promoting learning and exploration, which are critical for transitioning to independence (Casey, 2015; Spear, 2000). However, evidence for enhanced learning in adolescence and associated neural mechanisms have remained elusive. We sought to test the hypothesis that adolescents would be better than adults at learning from reinforcement and that this benefit would be related to enhanced activity in brain regions that support learning and memory, particularly the striatum and the hippocampus.

Advances in understanding neural mechanisms of reinforcement learning in adults have leveraged computational reinforcement learning models to quantify trial-by-trial learning signals in the brain (Daw et al., 2005, 2011; O'Doherty et al., 2003). Such models highlight the important role of prediction errors (PEs), which reflect the extent to which reinforcement received on a given trial deviates from what is expected. By reflecting trial-by-trial deviations between predictions and outcomes, prediction errors provide a learning signal that updates subsequent behavior. fMRI studies in adults and adolescents have shown that prediction errors correlate with blood-oxygen-level-dependent (BOLD) activity in the striatum (e.g., Christakou et al., 2013; Cohen et al., 2010; Hare et al., 2008; O'Doherty et al., 2003; van den Bos et al., 2012). Despite some reports of enhanced striatal activity in adolescents, reports of developmental differences in prediction error-related striatal activity are mixed (Christakou et al., 2013; Cohen et al., 2010; van den Bos et al., 2012), and so far, none have shown a link between enhanced striatal BOLD activity in adolescents and enhanced learning. This suggests that, to the extent that adolescents' reward sensitivity could be related to benefits for learning, these may be accounted for by other brain systems.

A natural brain candidate region for supporting reinforcement learning in adolescence is the hippocampus, known for its role in long-term episodic memory (e.g., Davachi, 2006; Gabrieli, 1998; Squire et al., 2004). The hippocampus also contributes to reward-related behaviors, including reinforcement learning, reward-guided motivation, and value-based decision making. Studies in adults show that the hippocampus and the striatum interact cooperatively to support both episodic encoding and reinforcement learning (Adcock et al., 2006; Bunzeck et al., 2010; Wimmer and Shohamy, 2012). These findings suggest that reward sensitivity in adolescence could be

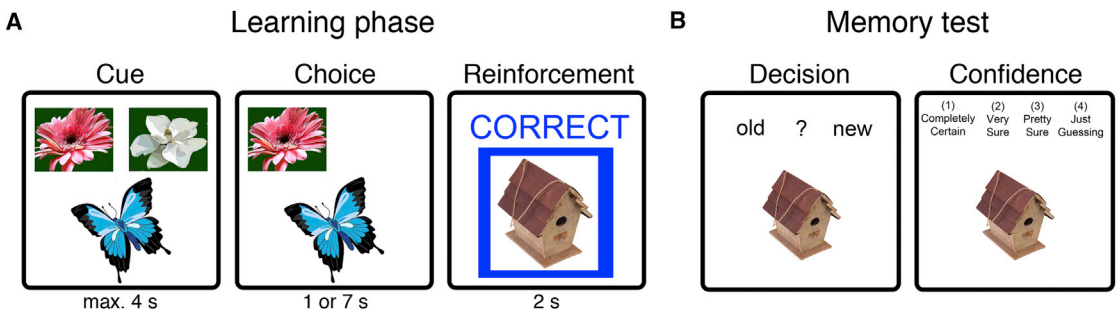

Figure 1. Behavioral Task to Assess Trial-by-Trial Incremental Learning and Episodic Memory

(A) Learning phase: on each trial, a centrally presented cue appeared below two targets. Participants pressed a button to predict which flower a butterfly would land on and received probabilistic reinforcement along with a trial-unique picture of a commonplace object.

(B) Memory test: participants saw a picture of an object, judged whether the picture was "old" or "new," and then rated their level of confidence in that choice.

related to enhanced hippocampal activity, to better reinforcement learning, and to better episodic memory for rewarding events. But, so far, the role of the hippocampus in reinforcement learning in adolescence has not been studied.

We used a learning task in combination with fMRI and reinforcement learning models to address this gap. We hypothesized that, compared to adults, (1) adolescents would be better at learning from reinforcing outcomes; (2) adolescents would show a greater relation between reinforcement learning and episodic memory for rewarding events during learning; and (3) these differences in learning would be related to enhanced activity in the hippocampus and stronger coupling between the hippocampus and the striatum.

Participants learned incrementally, based on trial-by-trial reinforcement, to associate cues with outcomes (Figure 1A). The association between cues and outcomes was probabilistic, requiring continual use of reinforcement to update choices. Reinforcement was simply the word "correct" or "incorrect" and was not motivated by monetary incentives to avoid confounds related to the motivational significance of monetary reward across age groups. To test episodic memory for reinforcement events, we included a unique picture of an object that was incidental to the reinforcement itself in each outcome (Figure 1B). This design allowed us to measure (1) incremental learning based on trial-by-trial reinforcement, (2) episodic memory for reinforcement events, which are positive versus negative, and (3) the role of the hippocampus and the striatum in both forms of learning.

RESULTS

Enhanced Reinforcement Learning in Adolescents

We tested whether adolescents (n = 41, 13–17 years old) differed from adults (n = 31, 20–30 years old) at learning from reinforcements, comparing (1) overall performance and (2) estimated learning rates from the reinforcement learning model. Learning performance was quantified as the percent of trials for which participants responded with the outcome most often associated with a given cue (e.g., Poldrack et al., 2001; Shohamy et al., 2004). A repeated measures (RM)-ANOVA (block × group) revealed that both age groups showed significant learning, but, consistent with our prediction, adolescents' learning exceeded that of adults (Figure 2A; main effect of block: $F_{3,210} = 20.2$, $p = 0.000$; block × group interaction: $F_{3,210} = 4.04$, $p = 0.008$). Similar results were found for optimal choice by trial (mixed-effect regression, main effect of trial: $z = 7.13$, $p = 0.000$; group × trial interaction: $z = -2.97$, $p = 0.003$), and we also found a better fit of the interaction model ($\chi^2 = 8.2$, $p = 0.004$) after penalizing for model complexity (Akaike, 1974).

To further characterize trial-by-trial responses, we applied a standard reinforcement learning model to each participant's choice data (Equation 1 in Supplemental Experimental Procedures). We chose to fit a canonical model, which represents a standard class of models used extensively in studies of brain correlates of reward prediction errors in adults (see Daw et al., 2011). We estimated a learning rate parameter for each participant (α), which reflects the extent to which feedback on each trial is used to update later choices. Here, a lower learning rate is better because the probabilistic associations between cues and targets are fixed; a lower learning rate suggests that learning is guided by accumulating evidence over a greater number of trials rather than shifting behavior based on the outcome of any single trial (e.g., Daw, 2011).

Importantly, the model provided a good fit to the observed behavior across both groups (one-way t test comparing a null model, $t_{71} = -39.70$, $p = 0.000$, Akaike's Information Criterion [AIC] used to penalize model complexity), and the model fits did not differ between them (independent samples t test, $t_{70} = 1.35$, $p = 0.2$). Consistent with their overall better learning, adolescents had a lower learning rate than adults ($t_{70} = -3.0$, $p = 0.004$; Figure 2B), indicating more incremental learning. Moreover, across groups, there was a significant negative correlation between learning rate and improved performance on the task ($r_{70} = -0.43$, $p = 0.000$; Figure S1A), indicating that lower learning rates were indeed related to better performance. Reaction times decreased over time for both groups, with no differences between them, suggesting that differences in learning are not due to general differences in responses to task demands (Figure S1B).

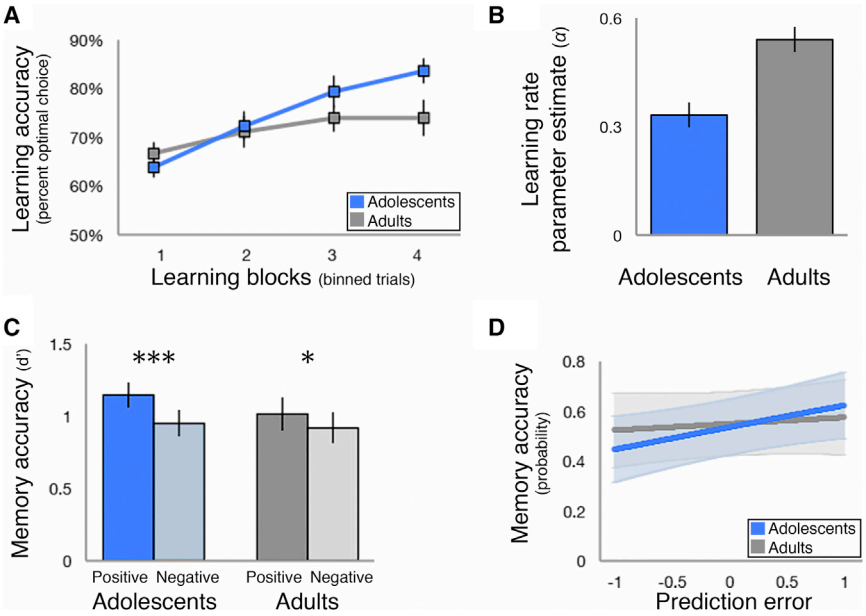

Figure 2. Behavioral Results: Adolescents Differ from Adults in Reinforcement Learning and in Association between Reinforcement Learning and Episodic Memory

(A) Learning accuracy. Both groups learned over time, but adolescents' learning exceeded adults'. Points reflect mean optimal choice for 24 or 30 (fMRI) trials; error bars show ±1 SEM.

(B) Learning rate parameter estimates from a reinforcement learning model. Adolescents had a lower learning rate than adults, reflecting more incremental updating of choice based on reinforcement. Error bars show ±1 SEM.

(C) Memory accuracy (d′) for trial-unique pictures that had been presented during reinforcement events in the learning task. Memory accuracy was computed separately by presented reinforcement to determine whether adolescents differed in their memory for positive and negative events. Adolescents and adults had better memory for images that accompanied positive, rather than negative, reinforcement. Error bars show ±1 SEM between participants. ***$p < 0.000$, *$p < 0.05$.

(D) The relationship between trial-by-trial reinforcement learning signals and later episodic memory for the reinforcement event. Only adolescents showed a reliable relationship between the magnitude of prediction error learning signals and likelihood of remembering episodic details of the reinforcement event. Lines show association between level of prediction error and the predicted probability from the fitted model for memory accuracy. Error bars around the fitted line show ±1 SEM.

Memory Positivity Bias in Adolescents and Adults

We first assessed episodic memory for the trial-unique objects that were presented during learning, separating trials by whether subjects had been shown positive ("correct") versus negative ("incorrect") outcomes. We found a significant effect of reinforcement (RM-ANOVA, $F_{1,70} = 24.6$, $p = 0.000$; no effect of group, $F_{1,70} = 1.6$, $p = 0.2$; no interaction, $F_{1,70} = 1.2$, $p = 0.3$; Figure 2C; Supplemental Information; Table S1), indicating that both groups showed a "positivity bias"—better memory for positive, rather than negative, reinforcement events.

Trial-by-Trial Prediction Errors Are Associated with Episodic Memory in Adolescents, but Not in Adults

We next tested whether reinforcement learning measures were related to episodic memory using model-derived estimates of trial-by-trial prediction errors (δ) (Equation 1 in Supplemental Experimental Procedures). Prediction errors provide an estimate of how surprising each trial's outcome was, which we used as a within-participant regressor for both behavioral and brain imaging analysis.

We found that prediction errors were related to memory accuracy and that this effect significantly interacted with group (mixed-effect regression interaction: PE × group, $z = 2.4$, $p = 0.02$; no main effect of PE, $p = 0.2$; or group, $p = 0.7$). This interaction reflected a significant relationship between prediction error and memory among the adolescents ($z = 5.2$, $p = 0.000$; Figure 2D), but not the adults ($z = 1.3$, $p = 0.2$). Thus, in adolescents, but not adults, episodic memory for outcomes was related to prediction errors. A similar effect was found for the relationship between reinforcement learning and the positivity bias in episodic memory across participants (Figures S1C and S1D).

Prediction Error Signals in the Hippocampus in Adolescents

A subset of 25 adolescents and 22 adults underwent fMRI while performing the learning task (behavioral effects in the fMRI group were the same as in the full behavioral sample; see Figures S1E–S1I). To interrogate the brain systems underlying differences in behavior between groups, we regressed prediction errors against BOLD activity within each participant and compared the groups in regions of a priori interest in the hippocampus and the striatum (for whole brain results, see Table S2; for analyses of value in the ventromedial prefrontal cortex [vmPFC], see Supplemental Information).

We found that prediction errors were correlated with BOLD activity in the striatum in both groups, with no significant differences between them (see Figure S2D; Table S2). In the hippocampus, by contrast, adolescents had significantly greater prediction error-related BOLD activity than adults (Figure 3C; Figure S2C).

Given the behavioral link between reinforcements and memory in the adolescents, we investigated whether episodic memory was related to functional connectivity between the hippocampus and the striatum. We used a psychophysiological interaction (PPI) analysis with the time series from a hippocampal seed as the physiological variable (Figure 4A) and reinforcement valence of the outcome event (correct > incorrect) as the psychological variable. We found significant connectivity between the hippocampus and the putamen in adolescents (but not

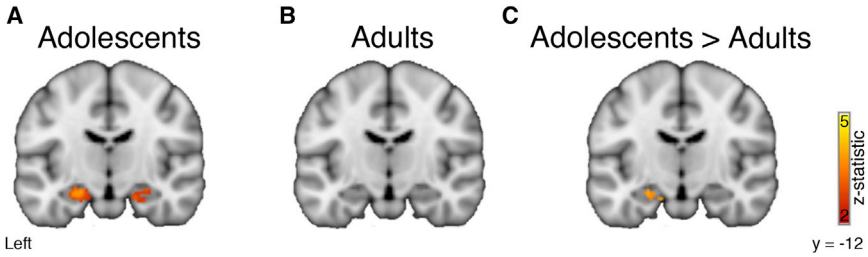

Figure 3. Greater Prediction Error-Related Activation in the Hippocampus in Adolescents

(A) Adolescents (n = 25) showed significant activation bilaterally in the hippocampus in two clusters (left: family-wise error (FWE)-p < 0.01, z = 4.15, peak [−16, −8, −20]; right: FWE-p < 0.03, z = 3.23, peak [24, −20, −12]).

(B) Adults (n = 22) did not show above threshold activation in the hippocampus.

(C) Direct comparisons between groups within the hippocampus showed significantly greater activation in the left hippocampus in the adolescents than in the adults (FWE-p < 0.03, z = 3.54, peak [−16, −8, −22]). See Figures S2A–S2C.

adults) that was greater for correct than incorrect outcomes (z = 2.68, family-wise error (FWE)-p < 0.01, 155 voxels, peak [−16, 10, −6]; Figure 4B). We then extracted the interaction value for each participant from the PPI and correlated this measure of learning-related connectivity with an independent within-participant behavioral measure of memory positivity bias (Figure 4C). We found a significant correlation between connectivity during learning and the extent to which memories for positive reinforcement events were enhanced for the adolescents (r = 0.62, p = 0.000), but not the adults (r = 0.05, p = 0.84), and a significant difference in the correlations between the groups (comparison of Fisher z transformed correlation coefficients z = 2.16, p = 0.03).

DISCUSSION

The negative implications of reward sensitivity in adolescents have been well documented, but much less is known about the possible adaptive side for learning. Our results show that adolescents were better at learning from outcomes, outperforming adults. We also found that in adolescents, but not adults, trial-by-trial reinforcement learning is related to episodic memory for reinforcement events, such that memory was better for surprisingly positive versus negative outcomes. These behavioral benefits were related to heightened prediction error-related BOLD activity in the hippocampus and to stronger functional connectivity between the hippocampus and the striatum at the time of reinforcement. Finally, only in the adolescents, functional connectivity between these learning systems was related to the extent of bias toward better memory for positive reinforcement events.

This is the first demonstration of a role for the hippocampus in reinforcement learning in adolescents. Our results imply that, as adolescents navigate through new life experiences, learning from reinforcement is linked to how episodic memories are shaped and to the extent to which they are biased toward encoding more of the good than the bad. This feature of learning is important to consider in relation to decision making because it speaks to the sorts of biases that adolescents may encounter when they draw on prior experience to inform current decisions.

It is important to note that the adolescents in our study were not better at all types of learning; rather, the benefits were selective to reinforcement-based updating and reward-related memory. Overall, episodic memory in the adolescents was not better than in adults, and there were no differences between the groups in memory for just positive or just negative learning events. Instead, the groups differed specifically in the strength of the interaction between these two forms of learning.

These findings suggest that, in adolescents, there is less differentiation between different forms of learning and their neural substrates when compared with adults. One possible interpretation of this finding is that it may be related, in part, to the known delay in development of prefrontal control mechanisms in adolescence (Somerville and Casey, 2010 for review). Although it is not known precisely how the arbitration between different learning and decision systems takes place in the adult brain, it has been suggested that the prefrontal cortex may play an important role (Daw et al., 2005; Poldrack and Packard, 2003). Indeed, an influential model of adolescent decision making posits a dynamic imbalance between appetitive motivational brain systems, including the striatum, and inhibitory control systems in the prefrontal cortex (Galván, 2013; Somerville and Casey, 2010). Our findings extend this framework and show that the striatum may be just one learning system, along with the hippocampus, that has relatively greater influence during adolescence. Specifically, our findings suggest that the functional development of midbrain dopaminergic reward systems and their connectivity with the striatum and the hippocampus in adolescence is positioned to affect both strengthening of reward-guided habits and actions, as well as episodic memory for motivational events. Future studies will need to assess the role of control and flexibility to identify whether prefrontal systems regulate the interactions between the striatum and the hippocampus.

The current study aimed to evaluate the link between reinforcement learning and episodic memory by concurrently presenting incidental trial-unique stimuli with reinforcement. An important direction for future research will be to determine whether these findings extend to goal-directed episodic encoding. In adults, striatal activity has been shown to relate to goal-directed modulation of episodic memory (Han et al., 2010). Prior work in adolescents has shown greater sensitivity to reward-predictive cues in the striatum (Galván et al., 2006). Together with the current findings, this suggests that goal-directed cue processing in adolescents may elicit greater cooperation between the hippocampus and the striatum and better goal-directed encoding. This possibility remains to be tested.

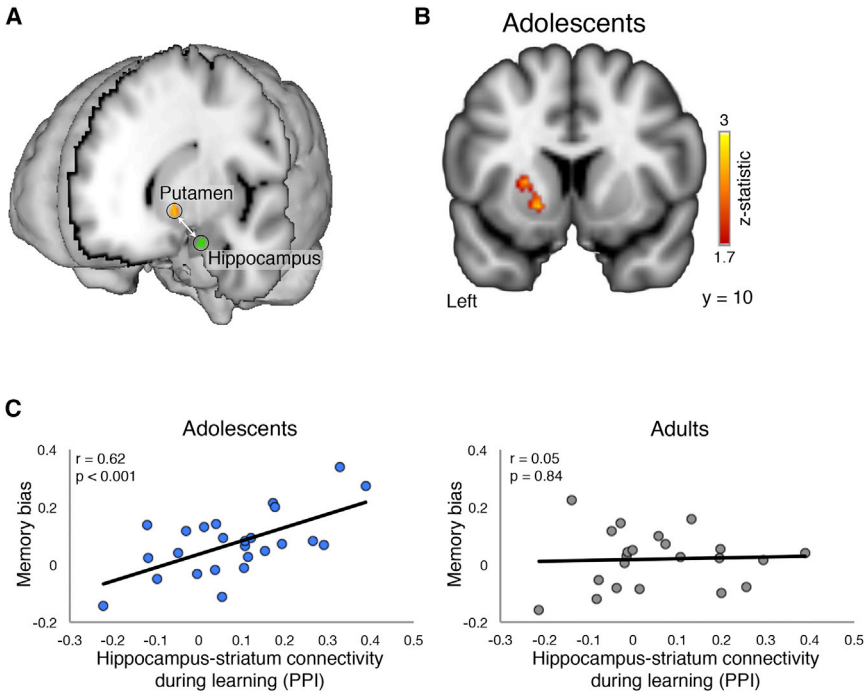

Figure 4. Functional Connectivity during Learning Relates to Memory Positivity Bias in Adolescents

(A) Time series within the hippocampus showed functional coupling with the putamen for the contrast of correct > incorrect presented reinforcement.

(B) Interaction between the physiological and psychological regressors in adolescents (limited to a hypothesis-driven search within the bilateral striatum; z = 2.68, FWE-p < 0.01, peak [−16, 10, −6]).

(C) Result of the interaction term from the PPI was extracted for each participant and correlated with behavioral memory bias. There was an association between learning-related connectivity and the enhancement of memory for positive reinforcement in the adolescents, but not in the adults.

Another important question is how subregions of the striatum contribute to learning and interact with the hippocampus. We found prediction error-related BOLD activity in the ventral striatum, as has been shown repeatedly (e.g., Bartra et al., 2013; Clithero and Rangel, 2014). This region is connected with the hippocampus (e.g., Haber and Knutson, 2010) and interacts with it functionally (e.g., Kahn and Shohamy, 2013). However, our functional connectivity analysis revealed activity in a separate region in the putamen that correlated with the hippocampus. Research in adults identified a similar region displaying connectivity with the hippocampus during cue-value learning (Wimmer et al., 2014). While functional connectivity in BOLD data does not necessarily reflect anatomical connectivity, these findings raise important questions for future work about the interacting circuits supporting reinforcement learning and episodic memory in adolescence.

Our findings are generally consistent with studies of episodic memory in development (Ghetti et al., 2010; Ofen et al., 2012). Previous research has shown that adolescents have adult-like recognition memory (Ghetti et al., 2010), whereas younger children have worse episodic encoding (Ghetti et al., 2010) and retrieval (DeMaster et al., 2014; Lloyd et al., 2009). Findings regarding developmental changes in the hippocampus have been mixed. Studies of item recognition report no differences in the hippocampus during development (Ofen et al., 2007). But other studies indicate that changes in the hippocampus do continue into adolescence (Daugherty et al., 2016; Lee et al., 2014) and are related to differences in associative memory performance in adolescents (Ghetti et al., 2010; DeMaster et al., 2014).

Many new experiences occur during adolescence. Some work suggests that, at least when looking back from adulthood, adolescence is a time in which particularly powerful and positive memories are formed (Haque and Hasking, 2010; Rubin and Berntsen, 2003; Thomsen et al., 2011). Of course, adolescence is also a time when psychopathology may begin to emerge (Casey et al., 2015; Ernst et al., 2009; Padmanabhan and Luna, 2014). Both perspectives emphasize the importance of learning from experiences during this time of development. The heightened sensitivity of striatal learning systems may put reward-seeking actions into overdrive but can also confer a benefit in learning from predictable, but variable, outcomes, as we show here. Our findings demonstrate that this reinforcement sensitivity has implications for what kinds of memories are formed in adolescence and how these memories drive behavior.

EXPERIMENTAL PROCEDURES

All recruitment, screening, consent and assent, and testing procedures were approved by the University of California, Los Angeles, Institutional Review Board (IRB) and Columbia University IRB. For descriptions of experimental materials and procedures, see Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.08.031.

AUTHOR CONTRIBUTIONS

J.Y.D., K.F., A.G., and D.S. designed the study; J.Y.D. and A.G. collected the data; J.Y.D. analyzed the data under the supervision of D.S., A.G., and K.F.; and J.Y.D., K.F., A.G., and D.S. wrote the manuscript.

ACKNOWLEDGMENTS

The authors thank Katherine Duncan, Brad Doll, and Raphael Gerraty for assistance with data analysis and Lucy Owen, Elizabeth Pierce, Kathy Do, Emily Barkley-Levenson, Diane Goldenberg, Kaitlyn Breiner, and Elica Rahdar for helping with data collection. This research was supported in part by The

National Science Foundation (DGE-11-44155 to J.Y.D., BCS 0963750 to A.G., and Career Award 0955494 to D.S.).

Received: December 23, 2015
Revised: July 10, 2016
Accepted: August 10, 2016
Published: October 5, 2016

REFERENCES

Adcock, R.A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., and Gabrieli, J.D.E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron *50*, 507–517.

Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. Automat. Contr. *19*, 716–723.

Andersen, S.L., Rutstein, M., Benzo, J.M., Hostetter, J.C., and Teicher, M.H. (1997). Sex differences in dopamine receptor overproduction and elimination. Neuroreport *8*, 1495–1498.

Bartra, O., McGuire, J.T., and Kable, J.W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage *76*, 412–427.

Brenhouse, H.C., and Andersen, S.L. (2011). Developmental trajectories during adolescence in males and females: a cross-species understanding of underlying brain changes. Neurosci. Biobehav. Rev. *35*, 1687–1703.

Brenhouse, H.C., Sonntag, K.C., and Andersen, S.L. (2008). Transient D1 dopamine receptor expression on prefrontal cortex projection neurons: relationship to enhanced motivational salience of drug cues in adolescence. J. Neurosci. *28*, 2375–2382.

Bunzeck, N., Dayan, P., Dolan, R.J., and Duzel, E. (2010). A common mechanism for adaptive scaling of reward and novelty. Hum. Brain Mapp. *31*, 1380–1394.

Casey, B.J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. Annu. Rev. Psychol. *66*, 295–319.

Casey, B.J., Glatt, C.E., and Lee, F.S. (2015). Treating the developing versus developed brain: translating preclinical mouse and human studies. Neuron *86*, 1358–1368.

Christakou, A., Gershman, S.J., Niv, Y., Simmons, A., Brammer, M., and Rubia, K. (2013). Neural and psychological maturation of decision-making in adolescence and young adulthood. J. Cogn. Neurosci. *25*, 1807–1823.

Clithero, J.A., and Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. Soc. Cogn. Affect. Neurosci. *9*, 1289–1302.

Cohen, J.R., Asarnow, R.F., Sabb, F.W., Bilder, R.M., Bookheimer, S.Y., Knowlton, B.J., and Poldrack, R.A. (2010). A unique adolescent response to reward prediction errors. Nat. Neurosci. *13*, 669–671.

Daugherty, A.M., Bender, A.R., Raz, N., and Ofen, N. (2016). Age differences in hippocampal subfield volumes from childhood to late adulthood. Hippocampus *26*, 220–228.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. Curr. Opin. Neurobiol. *16*, 693–700.

Daw, N.D. (2011). Trial-by-trial data analysis using computational models. In Decision Making, Affect, and Learning: Attention and Performance XXIII, M.R. Delgado, E.A. Phelps, and T.W. Robbins, eds. (Oxford University Press).

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron *69*, 1204–1215.

DeMaster, D., Pathman, T., Lee, J.K., and Ghetti, S. (2014). Structural development of the hippocampus and episodic memory: developmental differences along the anterior/posterior axis. Cereb. Cortex *24*, 3036–3045.

Ernst, M., Romeo, R.D., and Andersen, S.L. (2009). Neurobiology of the development of motivated behaviors in adolescence: a window into a neural systems model. Pharmacol. Biochem. Behav. *93*, 199–211.

Gabrieli, J.D.E. (1998). Cognitive neuroscience of human memory. Annu. Rev. Psychol. *49*, 87–115.

Galván, A. (2013). The teenage brain: sensitivity to rewards. Curr. Dir. Psychol. Sci. *22*, 88–93.

Galván, A., Hare, T.A., Parra, C.E., Penn, J., Voss, H., Glover, G., and Casey, B.J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. J. Neurosci. *26*, 6885–6892.

Ghetti, S., DeMaster, D.M., Yonelinas, A.P., and Bunge, S.A. (2010). Developmental differences in medial temporal lobe function during memory encoding. J. Neurosci. *30*, 9548–9556.

Haber, S.N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology *35*, 4–26.

Han, S., Huettel, S.A., Raposo, A., Adcock, R.A., and Dobbins, I.G. (2010). Functional significance of striatal responses during episodic decisions: recovery or goal attainment? J. Neurosci. *30*, 4767–4775.

Haque, S., and Hasking, P.A. (2010). Life scripts for emotionally charged autobiographical memories: a cultural explanation of the reminiscence bump. Memory *18*, 712–729.

Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., and Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J. Neurosci. *28*, 5623–5630.

Kahn, I., and Shohamy, D. (2013). Intrinsic connectivity between the hippocampus, nucleus accumbens, and ventral tegmental area in humans. Hippocampus *23*, 187–192.

Lee, J.K., Ekstrom, A.D., and Ghetti, S. (2014). Volume of hippocampal subfields and episodic memory in childhood and adolescence. Neuroimage *94*, 162–171.

Lloyd, M.E., Doydum, A.O., and Newcombe, N.S. (2009). Memory binding in early childhood: evidence for a retrieval deficit. Child Dev. *80*, 1321–1328.

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., and Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. Neuron *38*, 329–337.

Ofen, N., Kao, Y.-C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., and Gabrieli, J.D.E. (2007). Development of the declarative memory system in the human brain. Nat. Neurosci. *10*, 1198–1205.

Ofen, N., Chai, X.J., Schuil, K.D.I., Whitfield-Gabrieli, S., and Gabrieli, J.D.E. (2012). The development of brain systems associated with successful memory retrieval of scenes. J. Neurosci. *32*, 10012–10020.

Padmanabhan, A., and Luna, B. (2014). Developmental imaging genetics: linking dopamine function to adolescent behavior. Brain Cogn. *89*, 27–38.

Poldrack, R.A., and Packard, M.G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. Neuropsychologia *41*, 245–251.

Poldrack, R.A., Clark, J., Paré-Blagoev, E.J., Shohamy, D., Creso Moyano, J., Myers, C., and Gluck, M.A. (2001). Interactive memory systems in the human brain. Nature *414*, 546–550.

Rubin, D.C., and Berntsen, D. (2003). Life scripts help to maintain autobiographical memories of highly positive, but not highly negative, events. Mem. Cognit. *31*, 1–14.

Shohamy, D., Myers, C.E., Grossman, S., Sage, J., Gluck, M.A., and Poldrack, R.A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. Brain *127*, 851–859.

Somerville, L.H., and Casey, B.J. (2010). Developmental neurobiology of cognitive control and motivational systems. Curr. Opin. Neurobiol. *20*, 236–241.

Spear, L.P. (2000). The adolescent brain and age-related behavioral manifestations. Neurosci. Biobehav. Rev. *24*, 417–463.

Squire, L.R., Stark, C.E.L., and Clark, R.E. (2004). The medial temporal lobe. Annu. Rev. Neurosci. *27*, 279–306.

Thomsen, D.K., Pillemer, D.B., and Ivcevic, Z. (2011). Life story chapters, specific memories and the reminiscence bump. Memory *19*, 267–279.

van den Bos, W., Cohen, M.X., Kahnt, T., and Crone, E.A. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. Cereb. Cortex *22*, 1247–1255.

van Duijvenvoorde, A.C.K., Op de Macks, Z.A., Overgaauw, S., Gunther Moor, B., Dahl, R.E., and Crone, E.A. (2014). A cross-sectional and longitudinal analysis of reward-related brain activation: effects of age, pubertal stage, and reward sensitivity. Brain Cogn. *89*, 3–14.

Wimmer, G.E., and Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. Science *338*, 270–273.

Wimmer, G.E., Braun, E.K., Daw, N.D., and Shohamy, D. (2014). Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. J. Neurosci. *34*, 14901–14912.

Neuron, Volume *92*

Supplemental Information

An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence

Juliet Y. Davidow, Karin Foerde, Adriana Galván, and Daphna Shohamy

Supplemental results

Testing a standard reinforcement learning model without priors on estimated parameters

When allowing the learning rate and inverse temperature parameters to take on values between negative infinity and positive infinity, the model did not arrive at a solution for 15 adolescents (37% of the group) and 11 adults (35% of the group). For the remaining sample for whom the model could solve, 26 adolescents and 20 adults, the model provided a good fit to the observed behavior across both groups (one-way t-test comparing a null model, $t_{45} = -28.99$, $p < 0.000$) and the model fits did not differ between the groups (independent samples t-test, $t_{44} = 1.8$, $p = 0.08$). The learning rate parameter was lower in the adolescents ($t_{44} = -2.06$, $p = 0.046$; adolescent mean 0.38, standard error of the mean (SEM) 0.06, adult mean 0.55, SEM 0.06) and there was no difference between groups in the inverse temperature parameter ($t_{44} = -0.17$, $p = 0.86$; adolescent mean 5.6, SEM 1.1, adult mean 6.1, SEM 3.4).

Memory accuracy for incidental images

There was a significant overall effect of memory accuracy (d-prime) in the combined sample (one-way t-test against 0, $t_{71} = 15.5$, $p < 0.000$) and in each group separately (adolescents, $t_{40} = 12.2$, $p < 0.000$; adults, $t_{30} = 9.5$, $p < 0.000$) showing that all participants had good memory for the incidental pictures that were presented at the time of reinforcement. As reported in the main manuscript there was a main effect of reinforcement in a repeated measures analysis of variance (RM-ANOVA), such that there was better memory for images that were presented with positive than negative reinforcement, in adolescents ($t_{40} = 4.4$, $p < 0.000$) and adults ($t_{30} = 2.4$, $p = 0.02$).

Testing the association of memory accuracy with trial-by-trial expected value for cues estimated from a reinforcement learning model

In an analysis parallel to the association with trial-by-trial prediction error and later memory accuracy, we investigated whether there was also such an association between expected value as estimated by the reinforcement learning model for cues and later memory accuracy. We found no significant associations (main effect of Expected value, $Z = 0.8$, $p = 0.4$; main effect of Group, $Z = -0.3$, $p = 0.8$; interaction, $Z = -1.4$, $p = 0.2$). This supports the interpretation that differences in value representation, on their own, do not explain the differences we observe between groups in the relationship between reinforcement learning and episodic memory.

Targeted analyses of responses in the ventromedial prefrontal cortex (vmPFC) during cue and reinforcement

Given the known role of vmPFC in value and reinforcement learning (Bartra et al., 2013; Clithero and Rangel, 2014), we explored whether there were differences in value representation in adolescents and adults that might explain differences in learning. We used a general linear model to detect activation at the time of reinforcement for events that were positive greater than negative. We applied a correction threshold of $Z > 2.3$ ($p < 0.01$ one-tailed) to the whole brain in a group-level general linear model, to observe average activation maps in each age group separately, and to compare activation between groups. Within the vmPFC, we found significant activation in each group (adolescents, $Z = 4.37$, $p < 0.000$, peak [-12, 44, -12]; adults, $Z = 3.77$, $p = 0.008$, peak [6, 42, -12]), but no differences between them at the time of reinforcement. We performed a similar analysis at the time of cue onset, evaluating the correlation of estimated expected value for the cue from the reinforcement learning model. This analysis again revealed no significant differences in vmPFC activation between the groups.

Supplemental Figures and Tables

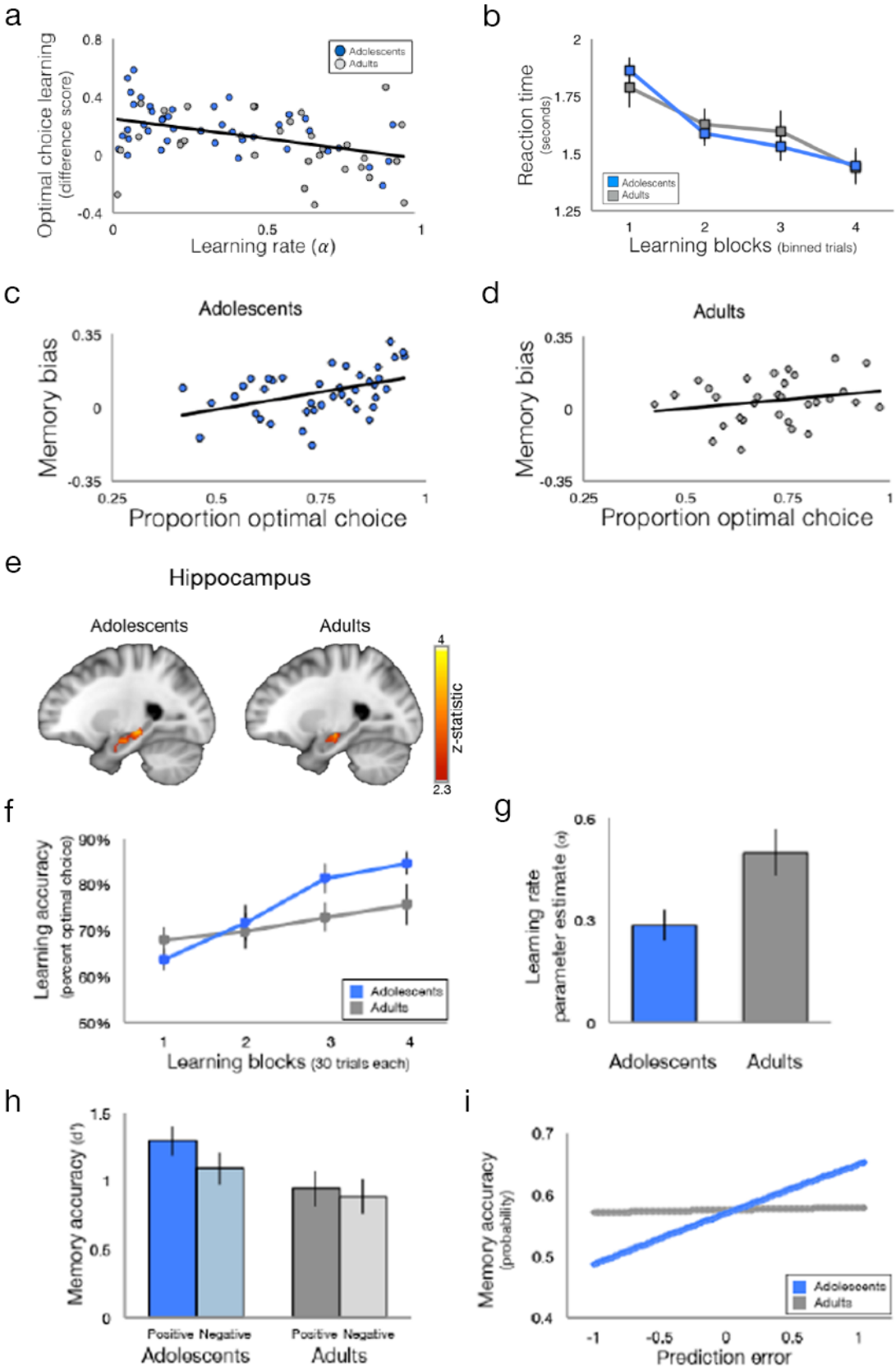

Figure S1. Related to Figure 2. (a) A negative correlation between learning rate and improvement in task performance. To investigate factors that could contribute to the differences in learning, we compared the estimated learning rate parameter (from the reinforcement learning model) and improvement in choices during the task (computed as the difference between mean optimal choice proportion in block 4 from block 1; $y = block_4 - block_1$) which reflects the increase in proportion of choosing the optimal target from the start to the end of the learning phase. Consistent with the idea that a lower learning rate is better in this sort of task, we found a significant correlation for the whole sample, $r = -0.43$, $p = 0.0002$; groups are plotted in blue and grey respectively for visualization. **(b) No difference in reaction time between the groups.** Latency to respond from onset of trial, until target selection. Both groups show a speeding of responses over the course of learning, (RM-ANOVA, Block: $F_{3,210} = 26.5$, $p = 0.000$), with no differences between them (Group: $F_{1,70} = 0.001$, $p = 0.9$; Block X Group: $F_{3,210} = 1.0$, $p = 0.4$). Points reflect mean reaction time per block; bars show ± one SEM. **(c-d) Positivity bias in memory is related to learning performance in adolescents, but not adults.** To relate behavioral measures of learning and memory across participants, we generated a single numeric index of memory sensitivity to reinforcement valence reflecting each participant's "positivity bias" in memory. This was calculated by computing a difference score for each participant between the proportion of remembered images when the outcome was positive vs. negative. We correlated this index of positivity bias with learning as the averaged proportion of optimal choice over the entire learning task. Among the adolescents, the positivity bias in memory was correlated with learning, such that adolescents who learned better also had a higher positivity bias ($n = 41$, $r = 0.43$, $p = 0.006$). Adults did not show this relationship ($n = 31$, $r = 0.22$, $p = 0.2$). A comparison of Fisher-Z transformed correlation coefficients was not significant ($Z = 0.93$, $p = 0.4$). **(e) Subsequent memory analysis of fMRI results in the hippocampus.** We investigated memory related activation in the hippocampus by comparing remembered and forgotten images from the behavioral memory test completed after the scan to brain activity at the time the image was initially presented during the learning phase. Memory regressors for fMRI analysis for each participant were derived from remembered and forgotten trials. To categorize a trial as remembered, the participant must have correctly identified the image as "old" and must have a d-prime above zero for the confidence-rating that the participant provided for that image. Similarly, forgotten images were defined as items mis-identified as "new" with a positive d-prime score for the associated confidence-rating. Any trial where a participant reported that they were "Just Guessing" was excluded. We examined BOLD correlates for remembered greater than forgotten pictures in each group separately and then directly compared between the adolescent and adult groups. These analyses were restricted to the hippocampus, our *a priori* subsequent memory ROI (see **Supplemental Table S1**). Within the adolescents this analysis revealed one significant cluster in the left hippocampus (Family-Wise Error (FWE)-p cluster < 0.001, $Z = 3.71$, 261 voxels, peak [-16, -30, -10]). For the adults, there was activation in bilateral hippocampus (Right: FWE-p cluster < 0.02, $Z = 3.63$, 96 voxels, peak [18, -14, -18]. Left: FWE-p cluster < 0.03, $Z = 3.38$, 93 voxels, peak [-22, -18, -12]), however in a direct contrast, there was no above threshold activation for either group over the other. Whole brain analysis results are in **Supplemental Table S1**. **(f-i) Behavioral results in fMRI subsample.** Behavioral results from the subset of participants who completed the probabilistic learning task while undergoing fMRI (n=25 adolescents; n=22 adults). Key behavioral findings within this sample replicate the results in the full behavioral sample reported in the main manuscript. Namely, the adolescents show (1) significantly better reinforcement learning than adults, (2) significantly lower learning rates than adults, (3) a significantly stronger association between prediction error learning signals and episodic memory for outcome events (within participants) and (4) a significant correlation between reinforcement learning and a "positivity bias" in episodic memory (within participants). **(f) Overall learning.** A RM-ANOVA (Block X Group) on the percent of trials for which participants responded with the outcome most often associated with that cue revealed that both age groups showed significant learning over time but, consistent with our prediction, adolescents' learning exceeded that of adults (main effect of Block: $F_{3,138} = 20.95$, $p < 0.0001$; Block X Group interaction $F_{3,138} = 4.49$, $p = 0.005$). Similar results were found for optimal choice by trial, showing learning for the full sample (mixed-effect regression main effect: $Z = 2.49$, $p = 0.013$), but better learning among the adolescent group (mixed-effect regression interaction: $Z = 2.65$, $p = 0.008$) with a significantly better fit of the group interaction model ($\chi^2=6.42$, $p = 0.01$) after penalizing model complexity using Akaike's Information Criterion (AIC). **(g) Learning rate parameter estimate from a reinforcement learning model.** Consistent with their overall better learning performance, adolescents had a lower learning rate parameter than adults ($t_{46} = 2.68$, $p = 0.01$), indicating more incremental learning. Importantly, the model provided a good fit to the observed behavior across both groups (one-way t-test comparing a null model, $t_{47} = -35.4$, $p < 0.0001$) and the model fits did not differ between them (independent samples t-test, $t_{46} = 0.77$, $p = 0.45$). **(h) Memory accuracy by reinforcement.** We compared memory for the trial-unique objects that were presented during learning, separating trials into those with positive ("correct") versus negative ("incorrect") outcomes. Adolescents had better

memory for positive than negative reinforcement events (two- tailed paired-samples test $t_{25} = 3.56$, $p = 0.002$). The adults showed significant memory overall (one-way t-test, $n = 22$, $t_{21} = 8.1$, $p = 0.0000001$; 1.1 ± 0.14), but no difference between positive and negative reinforcement ($t_{21} = 1.04$, $p = 0.3$). **(i) Prediction error and predicted probability of memory accuracy.** The level of prediction error evoked by the reinforcement on a particular trial was related to the probability of later memory accuracy, with an interaction effect for group (mixed-effect regression interaction: $Z = 2.78$, $p = 0.006$). This relationship between prediction error and memory accuracy was driven by the adolescents ($Z = 4.2$,
$p = 0.00003$), and no such relationship was found for the adults ($Z = 0.2$, $p = 0.8$).

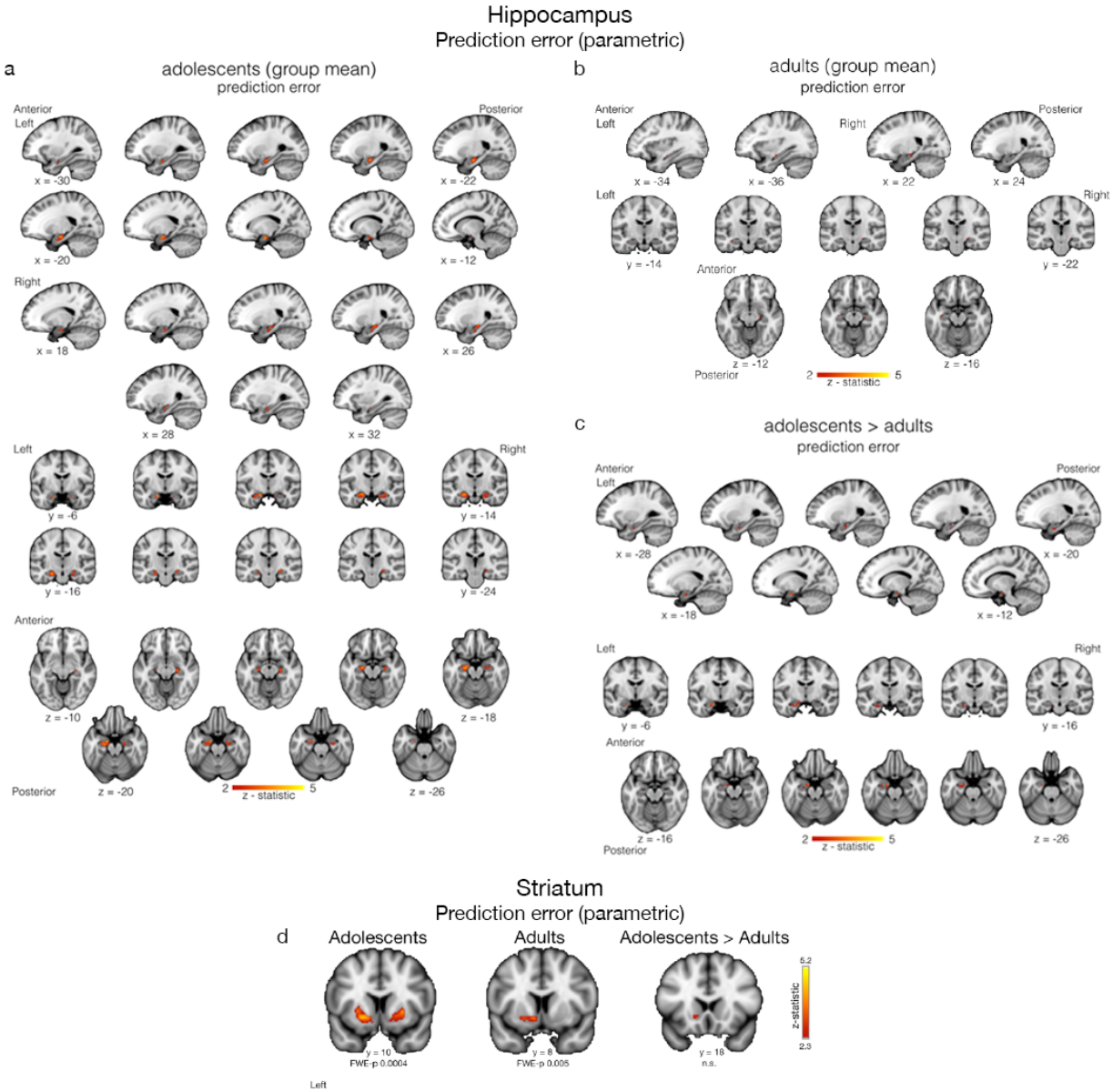

Figure S2. Related to Figure 3. (a-c) Full visualization of the extent of BOLD activity in the hippocampus correlating with prediction error. In the main manuscript (Figure 3) only one coronal plane is shown, and at the same location for 3 different maps that does not demonstrate the extent or peak of activity, so for purposes of visualization additional slices in all planes are presented. **(a) Adolescent group mean.** The adolescent group showed significant bilateral activation in the hippocampus correlated with prediction error. **(b) Adult group mean.** The adults showed bilateral activation in the hippocampus, but it did not surpass threshold for correction, (largest cluster $p > 0.1$, 13 voxels, peak [22, -20, -12], Z at peak = 2.5). **(c) Direct comparison of adolescent group greater**

than adult group. The adolescents had significantly greater activation in the left hippocampus than adults, and there were no significant findings for adults greater than adolescents. **(d) No significant differences between the groups in prediction error related activation in the striatum.** Prediction error was regressed as a parametric modulator with BOLD activity and examined within an *a priori* region of interest in the striatum (combined bilateral caudate, putamen, and accumbens, anatomically defined by the Harvard-Oxford probability atlas at a 75% probability threshold). Maps depict peak of presented contrast; see **Supplemental Table S2** for reports of local maxima. Contrasts within-group were thresholded at $Z>2.3$ one-tailed in FEAT. For between-group contrast correction was carried out with FSLs Cluster tool thresholded at $Z > 2.3$. Cluster extent correction thresholds were estimated by AFNIs 3dClustSim by providing the mask and smoothness within the mask, as calculated from the residuals using AFNIs 3dFWHMx (http://afni.nimh.nih.gov/). For adolescents, there was significant activation in the bilateral ventral striatum. For adults there was significant activation in the left ventral striatum. A comparison between groups did not meet correction threshold in either direction (depicted: adolescents > adults; largest cluster $p > 0.1$, 50 voxels, peak voxel [-14, 18, -4], Z at peak = 2.73). Whole brain results are in **Supplemental Table S2**. Contrasts within- and between-groups for whole brain analysis were thresholded at $Z > 2.3$ or $FDRp < 0.01$ one-tailed in FEAT. Whole brain contrasts between-groups yielded no significant differences at this threshold.

Table S1. Related to Figure 2c. Memory fMRI results (remembered > forgotten), peak and top 4 local maxima

Whole brain							
Anatomical Label (H-O atlas)	**Voxels**	**P**	**Z-score**	**MNI Coordinate**			
				X	**Y**	**Z**	
adults, *p-voxel < 0.01*							
L. temporal occipital fusiform cortex	2167	7.16e-11	4.04	-34	-58	-8	*cluster peak*
L. middle temporal gyrus			3.94	-66	-56	6	*local max*
			3.92	-54	-50	-8	*local max*
L. lateral occipital cortex			3.92	-36	-86	24	*local max*
L. inferior temporal gyrus			3.87	-40	-58	-6	*local max*
R. lateral occipital cortex	1044	3.76e-6	4.55	46	-62	-8	*cluster peak*
			3.53	42	-72	-8	*local max*
R. temporal occipital fusiform cortex			4.45	42	-48	-14	*local max*
			3.98	38	-52	-10	*local max*
R. inferior temporal gyrus			3.72	46	-58	-12	*local max*
L. frontal pole	683	0.000268	3.95	-8	66	28	*cluster peak*
			3.58	-8	62	2	*local max*
			3.56	-2	64	16	*local max*
R. superior frontal gyrus			3.55	8	56	18	*local max*
			3.54	-8	56	22	*local max*
L. frontal orbital cortex	437	0.00769	4.45	-36	32	-14	*cluster peak*
			3.52	-44	34	-6	*local max*
			3.39	-50	28	-10	*local max*
			3.02	-28	28	-20	*local max*
L. frontal pole			2.82	-52	40	-10	*local max*
L. amygdala	421	0.00974	4.23	-20	-6	-18	*cluster peak*
L. hippocampus			3.38	-22	-18	-12	*local max*
L. frontal orbital cortex			3.35	-28	8	-16	*local max*
L. parahippocampal gyrus			3.08	-16	-22	-24	*local max*
L. insular cortex			2.54	-34	8	-8	*local max*
R. postcentral gyrus	395	0.0144	3.87	36	-24	46	*cluster peak*
			3.26	36	-26	52	*local max*
R. precentral gyrus			3.65	40	-20	52	*local max*
			3.38	38	-20	66	*local max*
			3.25	38	-6	60	*local max*
R. amygdala	325	0.0429	3.67	12	-2	-20	*cluster peak*
			3.28	20	-6	-14	*local max*
			3.18	8	0	-12	*local max*
R. hippocampus			3.63	18	-14	-18	*local max*
			3.58	22	-14	-16	*local max*
adolescents, *p-voxel < 0.01*							
L. occipital pole	8867	5.99E-30	6.96	-32	-92	4	*cluster peak*
L. lateral occipital cortex			6.65	-42	-72	-12	*local max*
			6.16	-42	-76	-4	*local max*
L. occipital fusiform gyrus			6.3	-36	-78	-12	*local max*
L. inferior temporal gyrus			6.26	-44	-62	-8	*local max*
R. lateral occipital cortex	8229	1.99E-28	6.72	36	-88	-6	*cluster peak*
			6.47	36	-88	14	*local max*

			6.39	46	-64	-8	*local max*
			6.08	40	-64	-4	*local max*
R. occipital pole			5.87	16	-94	-2	*local max*
L. frontal operculum cortex	722	0.000164	4.15	-46	26	-2	*cluster peak*
L. frontal orbital cortex			4.06	-38	32	-10	*local max*
			3.56	-24	28	-12	*local max*
L. inferior frontal gyrus			3.81	-42	34	4	*local max*
			3.54	-46	28	14	*local max*
adolescents* < *adults, *p-voxel < 0.01*							
R. precuneous cortex	694	0.000233	4.37	4	-46	40	*cluster peak*
			3.46	0	-62	32	*local max*
			3.43	2	-58	32	*local max*
L. cingulate gyrus			3.55	-4	-40	44	*local max*
L. precuneous cortex			3.45	-4	-60	38	*local max*
adults* < *adolescents, *p-voxel < 0.01*							
R. lateral occipital cortex	3000	7.77E-14	4.96	36	-90	-6	*cluster peak*
			4.43	26	-90	-6	*local max*
			4.34	28	-82	18	*local max*
			4.28	34	-88	-12	*local max*
R. occipital fusiform gyrus			4.29	26	-84	-12	*local max*
L. occipital fusiform gyrus	2128	1.01E-10	4.44	-28	-84	-14	*cluster peak*
			4.27	-38	-76	-14	*local max*
			4.03	-40	-72	-12	*local max*
L. occipital pole			4.02	-26	-92	12	*local max*
L. temporal occipital fusiform cortex			3.65	-44	-52	-24	*local max*

ROI (bilateral hippocampus)

Anatomical Label (H-O atlas)	**Voxels**	**P**	**Z-score**	**MNI Coordinate**			
				X	**Y**	**Z**	
adults, *p-voxel < 0.01*							
R. hippocampus	96	0.0239	3.63	18	-14	-18	*cluster peak*
			3.58	22	-14	-16	*local max*
			2.59	16	-8	-20	*local max*
L. hippocampus	93	0.0257	3.38	-22	-18	-12	*cluster peak*
			3.32	-18	-12	-16	*local max*
adolescents, *p-voxel < 0.01*							
L. hippocampus	261	0.000852	3.71	-26	-30	-10	*cluster peak*
			3.69	-34	-30	-14	*local max*
			3.65	-22	-26	-12	*local max*
			3.17	-22	-10	-18	*local max*
			3.1	-24	-16	-18	*local max*
			2.95	-16	-8	-20	*local max*
adolescents* > *adults, *and* ***adults* > *adolescents***, *p-voxel < 0.01*							
no activation observed above threshold							

Table S2. Related to Figure 3. Prediction error fMRI results, peak and top 4 local maxima

Whole brain							
Anatomical Label (H-O atlas)	**Voxels**	**p-cluster**	**Z-score**	**MNI Coordinate**			
				X	**Y**	**Z**	
adults, *p-voxel < 0.01*							
R. occipital pole	5439	1.43E-19	4.83	14	-94	22	*cluster peak*
L. occipital pole			4.56	-30	-94	-10	*local max*
			4.46	-34	-92	-10	*local max*
			4.35	-12	-94	-14	*local max*
			4.34	-4	-92	-8	*local max*
R. cerebellum	372	0.035	4.26	38	-70	-38	*cluster peak*
			3.93	40	-66	-38	*local max*
			3.87	36	-60	-42	*local max*
			2.88	28	-46	-42	*local max*
			2.5	38	-78	-44	*local max*
adolescents, *p-voxel < 0.01*							
R. lateral occipital cortex	3377	6.95E-14	5.1	36	-84	4	*cluster peak*
			4.62	32	-90	12	*local max*
			4.6	34	-86	16	*local max*
			4.3	26	-84	12	*local max*
			4.05	52	-68	-8	*local max*
L. inferior temporal gyrus	3226	1.99E-13	4.83	-54	-62	-12	*cluster peak*
			4.62	-52	-52	-12	*local max*
L. lateral occipital cortex			4.54	-34	-90	4	*local max*
			4.48	-30	-88	-2	*local max*
L. middle temporal gyrus			4.49	-62	-48	-8	*local max*
L. amygdala	988	2.22E-05	4.65	-18	-8	-14	*cluster peak*
L. putamen			4.65	-18	10	-6	*local max*
			3.96	-26	0	-8	*local max*
L. accumbens (ventral striatum)			4.55	-12	6	-8	*local max*
L. hippocampus			3.64	-24	-14	-18	*local max*
L. frontal pole	785	0.000199	4.13	-48	44	6	*cluster peak*
			4.08	-48	40	12	*local max*
			3.94	-52	42	6	*local max*
			3.81	-42	36	14	*local max*
L. inferior frontal gyrus			3.62	-56	26	4	*local max*
R. precentral gyrus	578	0.00228	4.03	24	-26	72	*cluster peak*
			3.82	2	-28	64	*local max*
			3.73	2	-26	58	*local max*
			3.56	8	-26	74	*local max*
R. postcentral gyrus			3.43	26	-34	64	*local max*
adolescents > adults, *and* ***adults > adolescents***, *p-voxel < 0.01*							
no activation observed above threshold							

ROI (Striatum)							
Anatomical Label (H-O atlas)	**Voxels**	**p-cluster**	**Z-score**	**MNI Coordinate**			
				X	**Y**	**Z**	
adults, *p-voxel < 0.01*							
L. accumbens (ventral striatum)	222	0.00469	3.59	-8	8	-8	*cluster peak*
L. putamen			3.21	-28	0	6	*local max*
			3.15	-22	6	-8	*local max*
			2.75	-28	-12	6	*local max*
			2.74	-26	-6	-6	*local max*
			2.42	-26	-2	-8	*local max*
adolescents, *p-voxel < 0.01*							
L. putamen	388	0.000359	4.65	-18	10	-6	*cluster peak*
L. accumbens (ventral striatum)			4.55	-12	6	-8	*local max*
L. putamen			3.96	-26	0	-8	*local max*
			3.82	-24	8	0	*local max*
			3.35	-28	-14	6	*local max*
			3.27	-28	-8	-4	*local max*
R. putamen	209	0.00587	4.03	26	10	0	*cluster peak*
			3.63	16	6	-10	*local max*
adolescents > adults, *and* ***adults > adolescents***, *p-voxel < 0.01*							
no activation observed above threshold							

Supplemental Experimental Procedures

Forty-three (43) adolescents and 34 adults participated in the study. Among them, 28 adolescents and 25 adults were scanned with fMRI. Three adolescents were excluded from various analyses: two were excluded from all analyses, due to technical issues; one was excluded from brain imaging analyses due to excessive motion (absolute motion > 2.4 mm). Three adults were excluded from the final sample, two for technical issues, and one for an incidental neurological finding. This resulted in a total final sample of 41 adolescents (ages 13-17, mean age 15.9, SD = 1.4, 14 females) and 31 adults (ages 20-30, mean age 26.6, SD = 3.0, 18 females) for behavioral analyses and a subset of 25 adolescents (ages 13-17, mean age 15.9 years, SD = 1.4, 10 females) and 22 adults (ages 24-30, mean age 27.7 years, SD = 2.0, 13 females) for fMRI analyses.

All adult participants provided informed consent to participate in the study. All adolescent participants provided informed assent and had a parent provide informed consent and permission to participate. All participants were initially screened over the phone (for adolescent participants screening was completed with the adolescent and a parent). All participants were healthy, native-English speakers, with normal color vision, reported no history of psychiatric or neurological disorders, were not taking medication, and fMRI participants were all right-handed and had no contraindications for scanning. At the completion of study, participants were paid for their time; no additional monetary incentives were used as motivation in the task.

To estimate IQ, we administered the Wechsler Abbreviated Scale of Intelligence 2-part sub test, with adolescents scoring an average of 107 (SD = 10.4, range 86 to 126, n = 27) and adults scoring an average of 115 (SD = 9.5, range 99 to 138, n = 21). The subsample that completed the fMRI was similar, with adolescents scoring an average of 104 (SD = 10.3, range 90 to 126, n = 18) and adults scoring an average of 115 (SD = 10.6, range 99 to 138, n = 14). Adolescents reported estimated pubertal development using a self-report multiple choice questionnaire (Petersen et al., 1988), with a mode of stage 3, ranging from stages 2 – 5, reflecting that pubertal development in the sample was well underway but not complete. Providing racial and ethnic demographic information was optional; the fMRI sample was comprised of 24% Hispanic, 9% Asian, 19% African-American, 63% Caucasian, and 9% mixed-ethnicity participants recruited from the community.

Behavioral Task Procedures

Pre-task practice. Prior to performing the probabilistic learning task, adolescent and adult participants read instructions on a laptop and completed a practice round of the task (8 trials, different stimuli than used in the task) to become familiar with the general appearance of the task and to be certain that the instructions were correctly understood. The practice task was presented in MatLab (http://www.mathworks.com/) using the Psychophysics Toolbox (Brainard, 1997).

Learning task. Participants performed the probabilistic learning task (Foerde and Shohamy, 2011) on a computer or while undergoing fMRI. The task was presented in MatLab (http://www.mathworks.com/) using the Psychophysics Toolbox (Brainard, 1997) and presentation timing for fMRI of events and jitter durations were optimized for rapid event-related fMRI using OptSeq (Dale, 1999).

On each trial in the learning phase, participants saw one of four cues (butterfly; blue, purple, red, yellow) and had to predict which of two targets (flowers; pink, white) the butterfly was more likely to feed from. Participants had up to four seconds to make a response and were encouraged to respond as quickly as possible. Each cue was associated with one target on 80% of trials and with the other target on 20% of trials (**Figure 1A**). Thus most of the time choosing the optimal target for a cue results in "correct" reinforcement, but the other 20% of the time results in "incorrect" reinforcement, allowing the observation of learning rates over the course of the task as well as to estimate trial-by-trial expectations and prediction errors.

Participants pressed a button to choose either the left- or the right-sided target and then received reinforcement on their choice. Reinforcement was visually presented on the screen for two seconds, followed by a fixation-cross for a jittered inter-trial interval. Within participant, each target always appeared in the same location (i.e. fixed left and right) and was counterbalanced across participants. Within participant, the cue and target association was fixed over the entire task; the combinations were fully counterbalanced across participants. Learning behavior was analyzed using summary measures as well as trial-by-trial with reinforcement learning models and linear regression. For trial-by-trial regression, we fit a mixed effects linear model to each participant's optimal choice behavior using the *lmer*

package in R (Bates et al., 2015) associating accuracy with Trial X Group modeled as a fixed effect and Participants modeled as a random effect.

If a participant did not respond in time, the words "too late" were presented for the duration of the trial to preserve timing. No reinforcement was presented for such trials and thus there was no presentation of an “episodic” picture. These learning trials were discarded from behavioral analysis and modeled as a regressor of non-interest to be kept out of baseline for fMRI analysis. Such, invalid trials in the learning phase resulted in a different number of test items for the surprise memory test. There were very few of these invalid trials across all participants. For adults, the average response rate was 98% ± 0.03 SEM, with a minimum of 86%. For adolescents, the average response rate was 98% ± 0.03 SEM, with a minimum of 88%.

Post-learning test. Immediately following the learning phase was a test phase (1 fMRI run, 32 trials; behavior only sample - 1 run, 24 trials), where participants continued to make choices for the same cue-target associations, but no longer received reinforcement. Participants were instructed to continue choosing based on the associations that they had learned over the previous trials. This provided a measure for how well the associations were learned for each of the cues, in the absence of continued reinforcement.

Subsequent memory test. As described in the main text, after the learning task, participants were given a surprise test of recognition memory for the outcome images from the learning task. The pictures were orthogonal to the learning task in that they provided no information for learning the cue-target associations. Memory testing was completed on a computer, 30 minutes after the completion of the learning phase. For participants who did not undergo fMRI, the memory test was completed 10 minutes after the completion of the learning phase. Participants saw each picture that had appeared in the learning phase ("old" pictures) intermixed with an equal number of novel pictures ("new" pictures) that had not been seen in the learning phase. The order of the presentation of old and new pictures was random and unique to each participant. Confidence ratings provided an index of how sure participants felt about their memory judgments. Confidence ratings allowed the exclusion of “guess” trials from all behavioral analyses, and were used in generating behavioral memory regressors for fMRI analysis (see **Supplemental Figure S1e**).

We calculated an average memory accuracy score for each participant, excluding all memory test trials where participants indicated they were guessing. Memory accuracy was quantified by d-prime, both overall and separately for events where reinforcement was positive vs. negative. D-prime is a signal detection index and a more sensitive measure of accuracy because it takes into account the bias of any individual towards identifying items as old. We computed d-prime using normalized rates of correctly remembered trials (hits) and trials where participants mistakenly identified novel pictures as old (false alarms). For computation of d-prime for positive and negative reinforcement, normalized rate of false alarms was subtracted from normalized rates of hits when presented reinforcement was "correct" and "incorrect" respectively.

For correlating memory with functional connectivity during learning, we computed an index of valence bias in memory, by taking the difference score for each participant between the proportion of remembered images when the presented reinforcement outcome was positive vs. negative. Individuals who remembered many more images when the reinforcement was "correct" would have a positive difference score reflecting a "positivity bias" in memory, whereas those who remembered a similar proportion of images for both positive and negative reinforcements would have a value close to zero, reflecting little bias in memory.

Reinforcement Learning Model Analysis

To explore how reinforcement drives trial-by-trial updating of choices during learning, we fit a standard reinforcement learning model with two free parameters to each participant's observed choices in the learning phase (Daw, 2011; Sutton and Barto, 1998).

Limiting the prior probability distribution is appropriate for constraining possible solutions for estimated parameters, particularly when comparisons between groups is of interest (Daw, 2011). Additionally, a method for reducing noise in fitting parameters from individual participants comes from repeated estimations of LLE fits over many different initialization points for the minimization optimizer, and then checking for the smallest LLE and keeping the estimated parameters associated with that fit (Daw, 2011). This prevents solutions that fall into local minima. To address these pitfalls in reinforcement learning methods, we fit a model with a prior probability distribution limiting estimated parameters, α and β. Estimates for each free parameter were constrained with a probability density

function [β: *Gamma* (1.2, 5), α: *Beta* (1.1, 1.1)] (Daw et al., 2011). Priors were used to penalize the estimates output from the minimization optimizer. The minimization optimizer function was initialized at 20 random start points and then run 2,000 times, for 40,000 iterations per participant. The best solution, and its associated parameter estimates, among the iterations was selected. We also evaluated a model without the constraint of a prior on the free parameters (see **Supplemental Results**).

In **Equations 1** and **2**, Q is the estimated value, such that Q_t is the expected value on trial t and Q_{t+1} is the updated expected value on the trial following trial t. The difference between the expected outcome Q and the reinforcement received r on trial t, is the measure of prediction error (δ).

$$Q_{t+1} = Q_t + \alpha * \delta \qquad \textit{Equation 1}$$
$$\text{where } \delta = (r_t - Q_t)$$

One estimated parameter is a learning rate (α), which is a measure of the extent to which reinforcement on each trial is used to update choices. Values for α can range from 0 to 1. High learning rates closer to 1 indicate heavily weighing recent reinforcement for more rapid updating based on fewer trials; in a context where outcomes are probabilistic but probabilities are consistent, this can result in overweighting recent but rare reinforcement, and frequently shifting choice behavior (Sutton and Barto, 1998). A learning rate closer to 0 indicates more incremental learning in which changes in value accrue over a greater number of trials; this strategy allows for learning over several trials, such that surprising but rare reinforcement outcomes will not dramatically shift choice behavior (Sutton and Barto, 1998). In a probabilistic task like the one used here, a lower learning rate parameter is generally better, as it suggests that learning is guided by accumulating evidence over a greater number of trials rather than shifting behavior based on the outcome of any single trial. Another estimated parameter is the inverse temperature parameter (β) and follows a softmax distribution. This parameter can be thought of as a link that relates the value of an option and the actual choosing of that option (Daw, 2011). It is an index of noise in choice behavior (the extent to which a participant exploits the learned value vs. explores other options).

For each participant we also calculated the inverse log likelihood estimate (LLE; **Equation 2**) to describe how well the model fits each participant's observed choices C, using a minimization optimizer (fmincon in MatLab,

http://www.mathworks.com/) to find the global minimum for this test of model fit. The value term here is computed trial-by-trial by taking the difference of the expected value on trial t, Q_t, and the mean expected value for that stimulus from all trials Q_m. We used Akaike's Information Criterion (AIC; Akaike, 1974) to penalize the complexity of the model where k is the number of estimated parameters, and computed the chance level pseudo-R for two choices over the total number of trials T (Daw, 2011) and subtracted this from the AIC term for each participant.

$$LLE = LLE + \beta * Q_{t-m}(C_t) - \log \sum \exp(\beta * Q_{t-m}) \quad \textit{Equation 2}$$
where $Q_{t-m} = Q_t - Q_m$

$$AIC = (-2*(LLE))+(2*k) \quad \textit{Equation 3}$$

$$R = -T*LN(0.5) \quad \textit{Equation 4}$$

Parameters α and β were estimated for each participant as described above, and then respective group parameter averages were used for estimating prediction errors for individuals within the groups for behavioral and fMRI analyses (Daw, 2011).

Prediction errors were calculated by subtracting the estimated subjective value for a cue on each trial from the reinforcement received on that trial (i.e. 1 for "correct" reinforcement and 0 for "incorrect" reinforcement). Thus, if actual received outcome is larger than the expected outcome this results in a positive prediction error, and conversely received outcomes smaller than the expected outcome result in a negative prediction error. The more unexpected the reinforcement, the larger the prediction error estimate in its respective direction. Following previous work with a variant of the task (Foerde and Shohamy, 2011) we included both positive and negative prediction errors in the same parametric regressor for fMRI analyses. For trial-by-trial regression associating prediction error magnitude with later memory accuracy, we fit a mixed effects linear model to each participant's memory accuracy with Prediction error X Group as a fixed effect and Participants modeled as a random effect.

FMRI Data Acquisition and Preprocessing

Scanning was performed on a 3-Tesla Siemens Trio MRI scanner with a 12-channel head coil at the Ahmanson-Lovelace Brain Mapping Center at the University of California – Los Angeles. For each functional run during the learning phase of the behavioral task, after discarding 4 initial TRs (8 seconds) as the scanner stabilized, 200

volumes of T2*-weighted interleaved echo-planar (EPI) images were acquired in the transverse plane (slice thickness = 4 mm, 34 slices, voxel size = 3x3x4 mm^3, TR = 2000 ms, TE = 30 ms, FA = 90º, matrix = 64 x 64, FOV = 192 mm) for a total of 800 volumes. For the test phase all parameters were the same, except 175 volumes were collected. A high-resolution, T1-weighted magnetization-prepared rapid-acquisition gradient echo (MPRAGE) anatomical scan was acquired in the sagittal plane for each participant at the end of functional scanning (slice thickness = 1mm, voxel size = 1x1x1 mm^3, 160 slices, TR = 2170 ms, TE = 4.33 ms, FA = 7º, matrix = 256 x 256, FOV = 256 mm). Functional images were motion corrected, slice-time corrected using Fourier-space time-series phase-shifting, skull stripped using the FMRIB Software Library's (FSL) Brain Extraction Tool (Smith, 2002), spatially smoothed with a 5mm full width at half maximum (FWHM) Gaussian kernel, grand-mean intensity normalized, highpass filtered at 0.02 Hz, and registered to standard Montreal Neurological Institute (MNI) template. High-resolution MPRAGE images were skull-stripped with BET, segmented into probability maps for gray matter, white matter and CSF, and linearly registered to standard Montreal Neurological Institute (MNI) template with 12 degrees of freedom. All preprocessed images were visually inspected.

Motion

Functional images were motion corrected using MCFLIRT (Jenkinson et al., 2002) deriving 24 nuisance regressors related to motion (6-rigid body transformations – 3-translational, 3-rotational – their derivatives, their squares, and the square of the derivatives). One adolescent participant was excluded from all brain analyses for having motion greater than 2.5 standard deviations away from the whole sample (adolescents and adults) average motion. After removing this participant, the relative motion (t_{45} = -0.11, p = 0.9) and absolute motion (t_{45} = 0.29, p = 0.8) did not differ between adolescents (n = 25, mean absolute motion 0.25 ± 0.027 SEM, mean relative motion 0.07 ± 0.006 SEM) and adults (n = 22, mean absolute motion 0.24 ± 0.019 SEM, mean relative motion 0.07 ± 0.004 SEM).

FMRI Statistical Analysis

Imaging data were processed and analyzed using the FMRI Expert Analysis Tool (FEAT) from FSL v6.00 toolbox (www.fmrib.ox.ac.uk/fsl). Subject-level and group-level general linear models were analyzed in FEAT, using a general linear model (GLM) in FLAME1. For detecting activations correlated with prediction errors, the model included the onset of each reinforcement event weighted by its calculated parametric prediction error value,

orthogonalized to the reinforcement event. Onset of the choice for each trial and missed responses were also included in the GLM as regressors of non-interest. We used a within participant fixed-effects analysis to average regressed activation over the 4 functional runs (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT). Outputs from participant-level analyses were then analyzed in a group-level general linear model, to observe average activation maps in each age group separately, and to compare activation between groups.

Small Volume Correction

Given our hypotheses, centered on our *a priori* regions of interest in the striatum and hippocampus, group level whole brain results were masked for small volume correction. Correction was carried out in FEAT as part of post-stats, so that the mask was applied at a threshold of $Z > 2.3$ ($p < 0.01$ one-tailed) before cluster extraction for group means. The striatum and hippocampus ROIs were anatomically defined using the Harvard-Oxford probability atlas, thresholded at 75% probability. The striatum ROI was comprised of bilateral caudate, putamen, and nucleus accumbens (Desikan et al., 2006). The hippocampus ROI included the entire head, body, and tail, and visual inspection of the hippocampal ROI confirmed that the head was posterior to the amygdala (see **Supplemental Figure S2a-c**). Small volume correction for between-group contrasts was carried out by identifying clusters within masked regions using FSLs Cluster tool thresholded at $Z > 1.65$ ($p < 0.05$ one-tailed), and cluster extent correction thresholds were estimated by AFNIs 3dCulstSim by providing the mask and smoothness as calculated from the masked residuals using AFNIs 3dFWHMx (http://afni.nimh.nih.gov/).

Functional Connectivity Analysis

To examine functional connectivity between the striatum and the hippocampus, the two *a priori* regions of interests where we found prediction error related activation (see **Figure 3** and **Supplemental Figure S2**), we conducted a psycho-physiological interaction (PPI) analysis (O'Reilly et al., 2012). For the physiological regressor, we extracted timeseries in a region of the anterior hippocampus that consisted of a 3mm sphere drawn around the peak voxel of the adolescent group's mean response to prediction error. To be certain that this seed drawn around the peak voxel did not contain activation from regions outside of the hippocampus (e.g. the amygdala), we masked the sphere with the same anatomical ROI of hippocampus, described above. For the psychological regressor we used the contrast of correct > incorrect presented reinforcement. We included all other trial events in the GLM as regressors of non-

interest, as well as the physiological and psychological regressors, in addition to the interaction term of interest. The resulting map of activation of the interaction between the physiological and psychological regressors was masked with the anatomical ROI of the bilateral striatum, described above.

Supplemental References

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

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436.

Dale, A.M., 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp. 8, 109–114. doi:10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980. doi:10.1016/j.neuroimage.2006.01.021

Foerde, K., Shohamy, D., 2011. Feedback timing modulates brain systems for learning in humans. J. Neurosci. 31, 13157–13167. doi:10.1523/JNEUROSCI.2701-11.2011

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. NeuroImage 17, 825–841.

O'Reilly, J.X., Woolrich, M.W., Behrens, T.E.J., Smith, S.M., Johansen-Berg, H., 2012. Tools of the trade: psychophysiological interactions and functional connectivity. Soc. Cogn. Affect. Neurosci. 7, 604–609. doi:10.1093/scan/nss055

Petersen, A.C., Crockett, L., Richards, M., Boxer, A., 1988. A self-report measure of pubertal status: Reliability, validity, and initial norms. J. Youth Adolesc. 17, 117–133. doi:10.1007/BF01537962

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.

Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning : An Introduction, Adaptive Computation and Machine Learning. MIT Press, Cambridge, Mass.