# Reinforcement learning influences memory specificity across development

**Kate Nussenbaum (katenuss@nyu.edu) & Catherine A. Hartley (cate@nyu.edu)**
Department of Psychology, NYU
New York, NY, 10003, USA

## Abstract

**In some reward-learning contexts, abstract stimulus representations can effectively guide behavior, whereas in others, more detailed representations are needed to guide choice. Here, using a novel reinforcement learning task, we asked how children, adolescents, and adults flexibly adjust the specificity of the representations used for learning across contexts, as well as how the specificity of the representations used during learning influences subsequent memory. We found that across development, participants up-weighted more detailed information when doing so was beneficial. Further, participants who placed greater weight on detailed information during learning also demonstrated enhanced mnemonic specificity for the stimuli they encountered.**

**Keywords:** reinforcement learning; adaptive learning; memory; mnemonic specificity; development

## Introduction

Studies of value-based learning and episodic memory encoding suggest that relative to adults, children may represent information with less specificity (Michalska et al., 2016; Schiele et al., 2016; Ramsaran, Schlichting, & Frankland, 2019; Keresztes, Ngo, Lindenberger, Werkle-Bergner, & Newcombe, 2018). However, to effectively guide behavior across diverse contexts, learning and memory systems should flexibly adapt to the statistics of diverse contexts. Broad generalization gradients may effectively guide action in some contexts, but selecting adaptive actions in other environments requires more granular representations of experiences (Santoro, Frankland, & Richards, 2016).

Here, we developed and used a novel reinforcement learning task and subsequent test of recognition memory to address how the flexible adaptation of the specificity of value-learning changes across development and whether developmental change in mnemonic specificity emerges from corresponding changes in value-learning computations.

## Methods

### Participants and task

148 participants between the ages of 8 and 26 completed a two-part, online study. In the first session, participants completed a reinforcement-learning task that comprised six blocks. On every trial, they saw a stimulus and had to decide whether to approach it (and win or lose points) or avoid it (and see how many points they would have won or lost had they approached). Within each block, stimuli comprised five exemplars drawn from three categories (e.g., in the 'pets' block, stimuli comprised 5 unique dogs, 5 unique cats, and 5 unique rabbits). Critically, unbeknownst to participants, three blocks were category-predictive, such that the stimulus category (e.g., dogs) determined the mean of the normal distribution ($SD = 1.5$) from which rewards were randomly sampled on each trial. In each category-predictive block, one category was good (6 $\geq$ mean reward $\geq$ 3), one category was bad (-6 $\leq$ mean reward $\leq$ -3), and one was neutral (mean reward = 0, though 0 itself was never a possible outcome of 'approaching'). The other three blocks were exemplar-predictive, such that the individual exemplars were pseudo-randomly assigned deterministic point values between -9 and 9. Within each block, six stimuli repeated 6 times, three stimuli repeated 3 times, and six stimuli were only presented once.

One week after completing the learning task, participants completed a memory test. On each trial, participants saw an image, and had to determine if it was definitely new, maybe new, maybe old or definitely old. The test comprised all 90 old images that participants saw during learning, as well as 48 novel category foils (e.g., hamsters), and 54 novel exemplar foils (e.g., new dogs).

### Modeling reinforcement learning

We fit multiple variants of a temporal-difference learning model to our choice data. The models assumed that participants tracked stimulus values at both the category (c) and exemplar (e) levels, and integrated them to determine whether to approach or avoid each stimulus (s). At the group level, our best-fitting model included a single learning rate but separate inverse temperature parameters for scaling stimulus values at the category and exemplar level, such that:

$$p(approach|s) = \frac{e^{\beta_c * V(c) + \beta_e * V(e)}}{e^{\beta_c * V(c) + \beta_e * V(e)} + e^0}$$

The model allowed these inverse temperature parameters to vary across block conditions (category- vs. exemplar-predictive). On each trial (t), participants updated $V(c)$ and $V(e)$ based on the prediction errors they experienced (i.e., $(r - V(c)_t)$ and $(r - V(e)_t)$, scaled by a learning rate. Our best-fitting model also included a single free parameter for the initial category- and exemplar-level stimulus values.

## Results

### Learning

Participants made increasingly optimal approach-avoid choices as a function of increasing age and within-block trial number ($ps < .001$). Participants also learned to make more optimal choices faster in the category-predictive block relative to the exemplar-predictive block ($ps < .001$). The

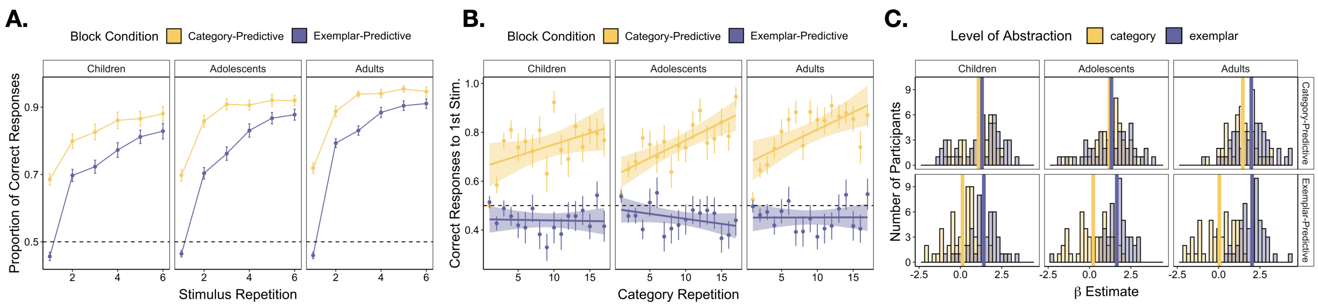


Figure 1: (A) Optimal approach-avoid choices across stimulus repetitions, age groups and block conditions. (B) Optimal approach-avoid choices for the first appearance of each stimulus across age groups and same-category trials. (C) Category-level and exemplar-level inverse temperature parameters across age groups and block conditions. Lines show age-group means.

effect of block condition varied by age ($p = .002$) such that age differences in optimal choices were more pronounced in the exemplar-predictive block (Fig. 1A).

Participants also used stimulus category to guide responses to novel, within-category exemplars. We found that optimal responses to new stimuli were more likely as a function of increasing category repetition, but only in the category-predictive condition ($p < .001$; Fig. 1B). In addition, this interaction effect grew stronger with increasing age ($p = .002$).

Did participants across age flexibly adjust the extent to which they used categorical versus exemplar-level value estimates to guide choice? To address this question, we examined inverse temperature parameter estimates from our computational model. In line with our hypothesis, we observed a block condition x level of abstraction interaction effect, $F(1,438) = 74.8, p < .001$, such that participants demonstrated higher values of $\beta_c$ in the category-predictive block (Fig. 1C). Contrary to our hypothesis, however, we did not observe a significant age x block condition x level of abstraction interaction effect, meaning we did not observe evidence for age differences in the extent to which participants flexibly adapted their weighting of categorical and exemplar-level information across blocks.

## Memory

Did the extent to which exemplar-level information could be used to guide learning influence subsequent recognition memory? To address this question, we used each participant's memory confidence responses to construct receiver operating characteristic curves for each level of memory specificity in each block condition. We then used the area under the curve (AUC) as a theory-neutral measure of memory performance (Brady, Robinson, Williams, & Wixted, 2022).

Participants demonstrated better category versus exemplar memory, as well as for stimuli encountered in the exemplar-predictive relative to the category-predictive blocks of the learning task ($ps < .001$) We did not observe a significant level of abstraction x block condition interaction effect, $F(1,442.2) = .12, p = .73$, nor did we observe any significant interactions with age ($ps > .10$).

Finally, we asked whether individual differences in the extent to which participants weighted exemplar-level information during learning influenced subsequent exemplar-level memory. We examined how $\beta_e$, age, block condition, and their interactions influenced memory for individual exemplars. In line with our hypothesis, we found that $\beta_e$ significantly predicted memory, $F(1,287.6) = 26.2, p < .001$ (Fig. 2), most strongly in the exemplar-predictive condition (β x block condition effect: $F(1,173.2) = 10.4, p = .002.$). We further observed a β x block condition x age interaction effect, such that when exemplar-level information was useful for guiding choice, older participants who weighted it most strongly showed the best exemplar-level memory.

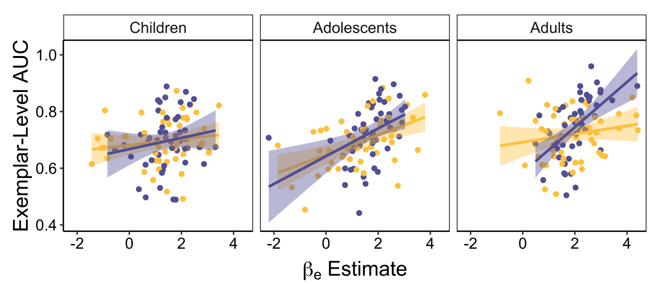


Figure 2: (A) Relation between $\beta_e$ estimates and exemplar-level memory (AUC) across age groups and block conditions.

## Conclusions

Taken together, our results show that participants across age could flexibly adapt reinforcement learning computations to the reward structure of the environment, using more specific representations to guide choice when doing so was necessary for gaining reward. Across development, memory specificity was shaped by reward-learning – the extent to which participants weighted exemplar-level information during learning influenced the extent to which they could differentiate old from new exemplars one week later. This relation grew stronger with age, suggesting that the prioritization in memory of information that is useful for decision-making may increase across development.

## Acknowledgments

We thank Juan Velez and Naiti Bhatt for help with stimulus design and data collection, as well as the NIMH (F31 MH129105 to K.N. and R01 MH126183 to C.A.H.) and the APA (Dissertation Research Award to K.N.) for funding.

## References

Brady, T. F., Robinson, M. M., Williams, J. R., & Wixted, J. T. (2022). Measuring memory is harder than you think: How to avoid problematic measurement practices in memory research. *Psychonomic Bulletin & Review*, 1–29. doi: 10.3758/s13423-022-02179-w

Keresztes, A., Ngo, C. T., Lindenberger, U., Werkle-Bergner, M., & Newcombe, N. S. (2018). Hippocampal Maturation Drives Memory from Generalization to Specificity. *Trends in Cognitive Sciences*, *22*(8), 676–686. doi: 10.1016/j.tics.2018.05.004

Michalska, K. J., Shechner, T., Hong, M., Britton, J. C., Leibenluft, E., Pine, D. S., & Fox, N. A. (2016). A developmental analysis of threat/safety learning and extinction recall during middle childhood. *Journal of Experimental Child Psychology*, *146*, 95–105. doi: 10.1016/j.jecp.2016.01.008

Ramsaran, A. I., Schlichting, M. L., & Frankland, P. W. (2019). The ontogeny of memory persistence and specificity. *Developmental Cognitive Neuroscience*, *36*, 100591. doi: 10.1016/j.dcn.2018.09.002

Santoro, A., Frankland, P. W., & Richards, B. A. (2016, 11). Memory Transformation Enhances Reinforcement Learning in Dynamic Environments. *Journal of Neuroscience*, *36*(48), 12228–12242. doi: 10.1523/jneurosci.0763-16.2016

Schiele, M. A., Reinhard, J., Reif, A., Domschke, K., Romanos, M., Deckert, J., & Pauli, P. (2016). Developmental aspects of fear: Comparing the acquisition and generalization of conditioned fear in children and adults. *Developmental Psychobiology*, *58*(4), 471–481. doi: 10.1002/dev.21393