Reinforcement learning increasingly shapes memory specificity from childhood to adulthood

Kate Nussenbaum[1,2], Catherine A. Hartley[1]

katenuss@princeton.edu; cate@nyu.edu

1. New York University
2. Princeton University

Abstract

In some contexts, abstract stimulus representations can effectively promote the pursuit of reward, 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 based on experienced reward statistics, as well as how the specificity of these learning representations influences subsequent memory. Across two experiments (total N = 224), we found that children, adolescents, and adults flexibly up- and down-weighted more detailed versus broader stimulus representations, depending on the reward structure of the environment. The representations used for learning shaped mnemonic specificity; placing greater weight on detailed representations during value-guided learning enhanced subsequent memory for stimulus details, while placing greater weight on broader, categorical representations enhanced memory only for categorical information. Moreover, the relation between learning and memory strengthened with age; relative to adults, children demonstrated reduced coupling between the specificity of the representations used for value-based choice and the specificity of their subsequent memories. Our work demonstrates that from early in life, reward shapes the granularity with which the world is partitioned, which in turn exerts an increasing influence on how experiences are remembered into adulthood.

Experiences can be represented at multiple, nested levels of abstraction. Last Friday, you may have eaten pasta at a restaurant and then gone to a movie — but you may have also eaten carbonara at an Italian bistro and seen *Barbie* in IMAX at the newly renovated cinema by your apartment. The specificity with which you represent your experiences has functional consequences for future behavior — representing your meal as 'pasta' may help you decide whether to eat an unfamiliar pasta dish, but may prove unhelpful in the future if you face a choice between carbonara and alfredo. More abstract representations may facilitate the acquisition of generalizable knowledge, whereas more specific representations can be leveraged to guide decisions that require finer-grained distinctions between similar entities (McClelland et al., 1995). While choices about eating pasta may be relatively unimportant, the specificity of our representations influences how we learn from the outcomes of our actions (Dunsmoor & Murphy, 2015; Love et al., 2004; Shepard, 1987; Tenenbaum & Griffiths, 2001), form lasting memories that underpin our mental models of the environment (Knowlton & Squire, 1993; Kumaran et al., 2016; McClelland et al., 1995; O'Reilly et al., 2014; Zaki & Nosofsky, 2001), and ultimately, harness our past experiences to guide our future behavior.

The generality or specificity with which experiences are represented may be particularly consequential early in life. Children, who are equipped with more capacity-limited learning and memory systems, must navigate a world of less familiar structure (Ghetti & Fandakova, 2020; Hartley et al., 2021; Li et al., 2004). Recent developmental studies of value-based learning and of episodic memory have suggested that there may be systematic increases in the specificity with which experiences are represented from childhood to adulthood. Younger children show broader generalization of threat responses to novel stimuli (Glenn et al., 2012; Mednick & Lehtinen, 1957; Schiele et al., 2016), provide reports of autobiographical memories that lack rich detail (Fivush et al., 1984; Nelson & Gruendel, 1981; Price & Goodman, 1990), and perform poorly on lab-based tasks of mnemonic discrimination (Keresztes et al., 2017; Lambert et al., 2015; Ngo et al., 2018; Rollins & Cloude, 2018). Theoretical proposals have suggested that representing information with less specificity early in life may be adaptive — a bias toward more general representations may promote the recognition of shared features across diverse experiences, which may be particularly useful for children as they build semantic knowledge of the world (Keresztes et al., 2018; Ramsaran et al., 2019; Reyna, 2012).

Several recent findings, however, suggest that developmental change in the specificity of learning and memory representations may not follow a simple, context-invariant trajectory. While some studies of value-based learning have indeed seen broader generalization in younger participants (Glenn et al., 2012; Schiele et al., 2016), others have found that generalization *increases* with age (Schulz et al., 2019). Studies of developmental changes in episodic memory have similarly revealed mixed findings, particularly in later childhood and adolescence. While some work has suggested that mnemonic specificity increases through late childhood (Keresztes et al., 2017), other research has not found evidence for significant age-related change in the granularity with which information is remembered (Callaghan et al., 2021). Even at younger ages, mnemonic specificity is not static; it can be enhanced if information is made more salient (Ngo et al., 2019). Moreover, the extent to which specificity and generality trade off may also change with age; detailed and more abstracted representations can compete for expression during learning (Richards et al., 2014), but detailed memories can also support generalization (Tompary et al., 2020), perhaps to a greater extent in adults than in children (Ngo et al., 2021).

Here, we suggest that these varied developmental trajectories of the specificity of value associations and episodic memory may reflect emerging adaptivity in the representations used for learning. The relative costs and benefits of representing experiences more abstractly versus more specifically do not just vary across the lifespan — they vary across the multiple, diverse learning environments that children, adolescents, and adults experience every day. In some contexts, more general representations can guide adaptive choice, and in others, more specific representations are needed. At the dog park, for example, walkers should represent the individuating features of each dog so they can learn to approach those that are friendly and avoid those that bite; in the woods, however, hikers can ignore the specific features of wolves and represent them more generally because they should avoid all of them — attempting to individuate each one may needlessly tax cognitive resources and prevent effective generalization. Adaptive value-guided learning thus requires the flexibility to adjust the specificity of value associations to the reward statistics of the environment (Lengyel & Dayan, 2007; Love, 2005; Love et al., 2004; Santoro et al., 2016). Some research suggests that the ability to dynamically tune value-learning computations to the optimal settings for particular environments improves from childhood to adulthood (Nussenbaum et al., 2022; Nussenbaum & Hartley, 2019). Other work, however, suggests that adults may approach new learning problems with stronger prior beliefs about the information most relevant for guiding behavior and show less flexibility in updating them in the face of new information (Decker et al., 2015; Liquin & Gopnik, 2022). Developmental changes in the specificity of value-learning computations may be driven by changes in the extent to which learning representations are dynamically shaped by the statistics of varied learning environments.

The specificity with which information is represented during learning may in turn influence the specificity with which information is encoded in memory, such that detailed information is preserved when it is useful for guiding behavior. A growing body of work has revealed a tight coupling between value learning and episodic encoding (Biderman et al., 2020; Gershman & Daw, 2017) — across development, the statistics of the environment (e.g., surprise, reward) govern both how value associations are learned as well as what information is attended and prioritized in memory (Calderon et al., 2021; Cohen et al., 2019; Davidow et al., 2016; Jang et al., 2019; Kalbe & Schwabe, 2020; Ngo et al., 2019; Rouhani & Niv, 2019, 2021; Starita et al., 2019; Wittmann et al., 2011). Further, individual and developmental differences in how people learn value associations relate to the information that they subsequently remember (Davidow et al., 2016; Rosenbaum et al., 2022). Despite research that indicates a strong influence of learning computations on *what* information is prioritized in memory, it is unclear how value-learning computations influence the adaptive *specificity* of memory representations across development.

Thus, our goals in this study were twofold. First, we sought to characterize how children, adolescents, and adults flexibly adapt the specificity with which they represent information during value-guided learning. We hypothesized that participants would increase their use of more specific representations when doing so was necessary for making good choices, to a greater extent with increasing age. Second, we asked how the specificity of the information used during value-based choice influences the specificity with which information is represented in memory. We hypothesized that across age, participants would demonstrate more specific memory for information encountered in the context in which detailed information was needed to guide choice. Further, we hypothesized that individual and developmental differences in the specificity of learning computations would be reflected in subsequent memory, such that people who placed more weight on detailed information during learning would show corresponding enhancements in memory specificity

We tested these questions across two reinforcement-learning experiments in which stimuli comprised unique exemplars drawn from broader categories. We manipulated the reward structure of the learning task across blocks, such that in some contexts, reward contingencies were determined by unique exemplars, whereas in others, they were governed by the broader categories. In both experiments, we found that participants across age flexibly adapted their use of exemplar-level and categorical information to make effective choices across contexts. In line with our hypothesis, individual differences in learning were reflected in subsequent memory, such that the specificity of memory was shaped by the specificity of value-guided learning. Further, we found that the influence of learning on memory strengthened across development, such that adults demonstrated a tighter coupling between the specificity of their learning computations and subsequent memory representations. Our findings reveal that the specificity of learning and memory does not follow a single developmental trajectory; instead, the structure of the environment shapes the specificity of the representations that children, adolescents, and adults use to guide choice, which are in turn, increasingly reflected in memory across development.

Results

Experiment 1

In our first experiment, 151 participants between the ages of 8 and 25 years completed a six-block ‘approach/avoid’ reinforcement-learning task across which the specificity of the representations that could best guide choice varied (see Methods). Within each block of the learning task, participants completed 51 trials in which they had to decide whether to ‘approach’ or ‘avoid’ one of 15 unique stimuli, drawn from three broader categories, to earn the most points (Figure 1). The order of stimulus presentation was randomized, and within each broader category, two images repeated five times, one image repeated three times, and two images were only shown once during learning, which meant that novel images were introduced throughout each learning block. Critically, in half of the task blocks (*category-predictive* blocks), the three broader stimulus categories determined the average gains and losses associated with approaching each stimulus. In category-predictive blocks, stimulus values were sampled anew from Gaussian distributions on every trial, where the mean of the distribution was determined by stimulus category. One category was randomly determined to be 'good' such that the mean of its reward distribution was between 3 and 6; one category was 'neutral' such that the mean of its reward distribution was zero (though zero was never actually presented as an outcome); and one category was 'bad' such that the mean of its reward distribution was between -6 and -3. In the other half of the task blocks (*exemplar-predictive* blocks), each unique exemplar was assigned a deterministic positive or negative point value between -9 and 9, distributed such that the broader stimulus categories could not be used to guide effective approach/avoid decision making (Figure 1B). The order of the blocks was randomized for each participant, with the constraint that the first two blocks were always of different conditions.

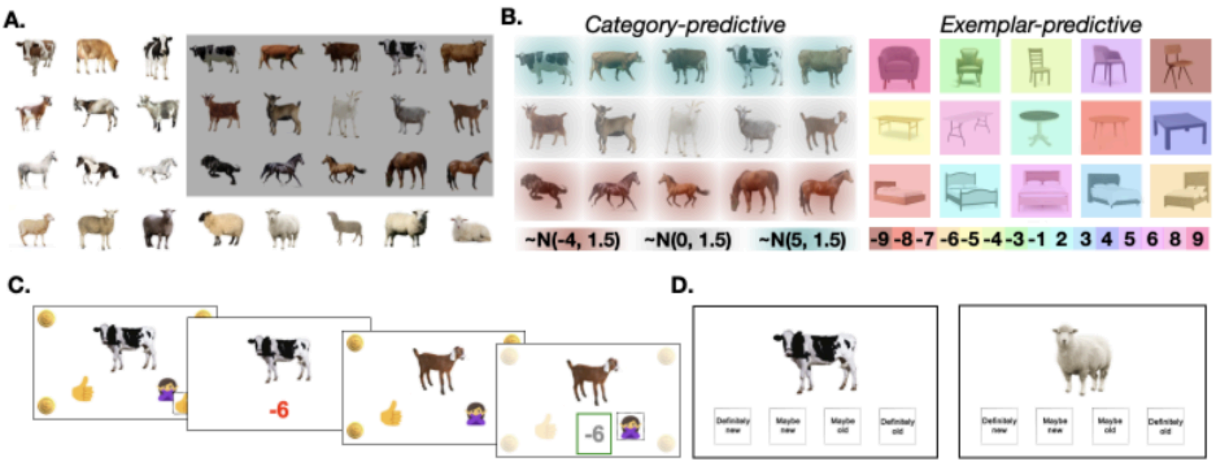


Figure 1. Experiment 1 task design. A) Each block of the reinforcement-learning task included 15 unique stimuli, which comprised five exemplars each drawn from three broader categories (shown in the gray box). For each stimulus set, three additional novel exemplars per sampled category and an additional category with eight novel stimuli were used in a test of subsequent memory (see panel D). B) In the category-predictive condition, rewards on every trial were sampled from normal distributions centered on means determined by the stimulus categories. In the exemplar-predictive condition, rewards on every trial were determined by the individual exemplars. C) On every trial of the reinforcement-learning task, participants chose whether to 'approach' or 'avoid' a stimulus. Participants won or lost points if they chose to approach the stimulus. While they did not win or lose any points if they chose 'avoid,' they saw counterfactual feedback showing how many points they *would have* won or lost had they approached. D) Approximately one week after completing the reinforcement-learning task, participants completed a test of recognition memory in which they had to decide whether stimuli were old or new on a four-point confidence scale.

Learning to 'approach' and 'avoid'

We first analyzed whether participants across age learned to approach stimuli with positive values and avoid those with negative values, via a logistic mixed-effects model with continuous age, within-block trial, block condition, within-condition block number, and their interactions as predictors. Participants increasingly made correct responses across trials within each block, odds ratio (OR) (and standard error) = 1.7 (.03), $\chi^2(1) = 284.9$, $p < .001$ (Figure 2A). Older participants made more correct responses than younger participants, OR = 1.3 (.05), $\chi^2(1) = 37.6$, $p < .001$, increasingly so across trials (age x trial interaction: OR = 1.1 (.02), $\chi^2(1) = 33.8$, $p < .001$). Across age, performance was better in category-predictive relative to exemplar-predictive blocks, OR = 1.7 (.05), $\chi^2(1) = 195.2$, $p < .001$, suggesting that participants leveraged categorical information to guide their choices. Further, the effect of block condition varied by age — older participants demonstrated stronger benefits from the ability to exploit categorical information, OR = 1.1 (.03), $\chi^2(1) = 10.1$, $p = .001$. Performance also improved across blocks of the task, OR = 1.3 (.04), $\chi^2(1) = 61.0$, $p < .001$ (see Supplement for full details).

Generalization of learned category values to novel stimuli

Throughout each block of the learning task, participants encountered novel stimuli that they had never seen before. In exemplar-predictive blocks, the value of each stimulus was determined independently, meaning participants could not infer the value of unseen stimuli based on their previous experiences. In category-predictive blocks, however, participants could respond optimally to novel exemplars by generalizing learned category values. Indeed, in category-predictive blocks, participants responded correctly to novel stimuli at well-above-chance levels (Figure 2B), indicating successful generalization. Participants made more correct responses to novel stimuli in category- relative to exemplar-predictive blocks, OR = 1.7 (.05), $\chi^2(1) = 215.1$, $p < .001$, an effect that grew increasingly strong as participants encountered more stimuli from each category (block condition x category repetition interaction: OR = 1.3 (.03), $\chi^2(1) = 164.2$, $p < .001$). In addition, the effect of block condition on correct responses grew stronger with increasing age, OR = 1.1 (.03), $\chi^2(1) = 8.1$, $p = .004$, indicating more effective generalization in category-predictive blocks in older participants. Generalization also strengthened across blocks of the task (block condition x block number interaction: OR = 1.1 (.02), $\chi^2(1) = 16.0$, $p < .001$; see Supplement for full details).

Though successful generalization was not possible in the exemplar-predictive condition, participants may have nonetheless attempted to generalize learned stimulus values to other, within-category exemplars, particularly within the first few trials of each block. Indeed, at the beginning of blocks across both conditions, participants tended to repeat rewarded 'approach/avoid' responses and switch unrewarded responses upon their subsequent encounter with a *different* stimulus from the same broader category. On average, in the first 10 trials within each block, participants demonstrated this category "win-stay-lose-shift" behavior in both category-predictive *and* exemplar-predictive blocks (mean proportion WSLS: category:

.60 (SE = .01), exemplar: .59 (SE = .01); Figure 2C), indicating that they began each block with a propensity to use categorical information to guide choice. But across trials, WSLS behavior increased in category-predictive blocks, where it was an effective choice strategy, and decreased in exemplar-predictive blocks, where it was maladaptive (trial x block condition effect: OR = 1.3 (.02), $\chi^2(1)$ = 355.7, p < .001). Older participants demonstrated the strongest divergence of WSLS behavior across conditions (trial x block condition x age effect: OR = 1.08 (.02), $\chi^2(1)$ = 23.6, p < .001). WSLS behavior also diverged across block conditions more strongly in later blocks of the task (block condition x block number interaction: OR = 1.2 (.02), $\chi^2(1)$ = 99.8, p < .001; see Supplement for full details).

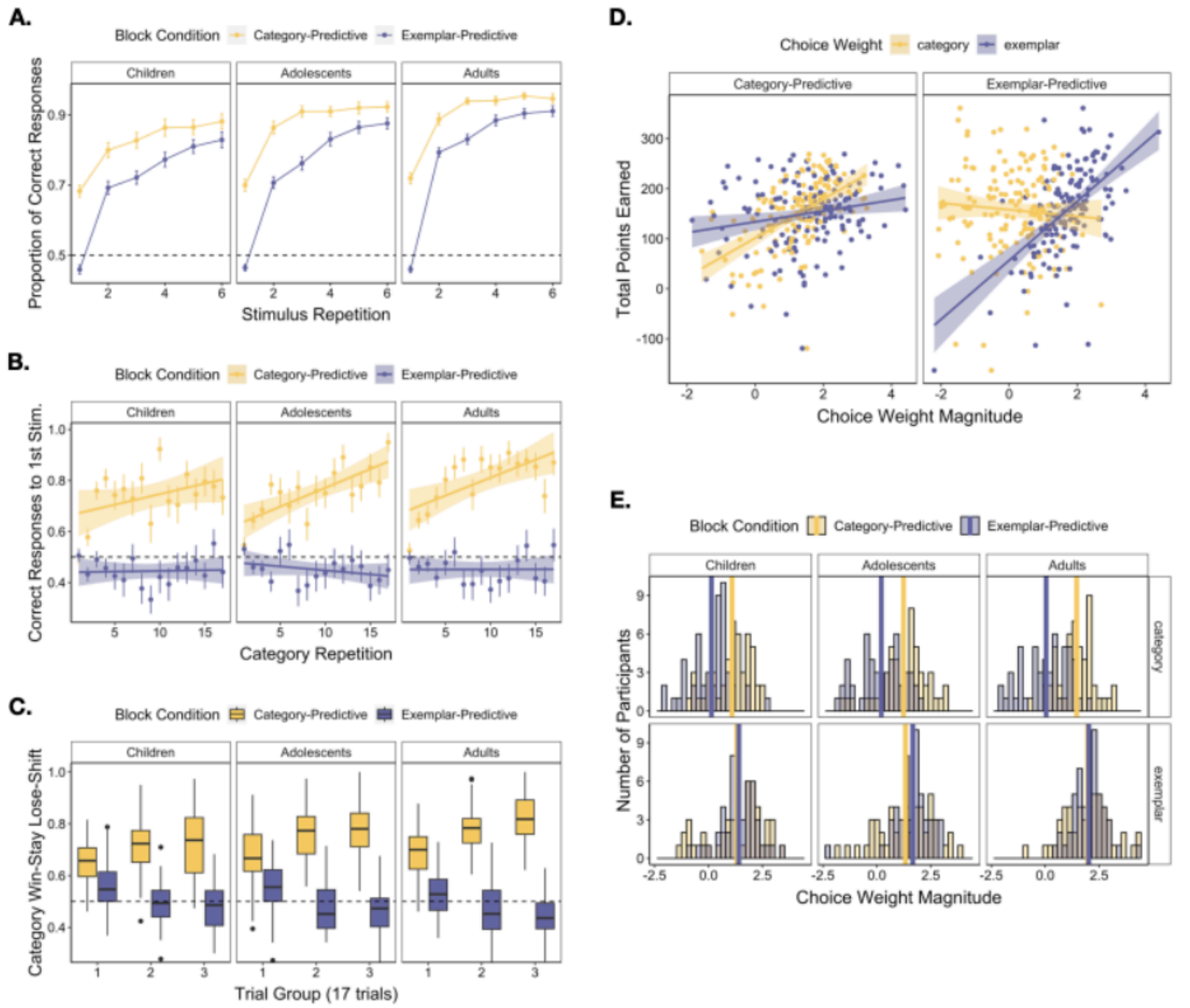


Figure 2. Participants across age flexibly adjusted the specificity of information used for learning. Panels A - C depict participant responses in the learning task, while C and D show parameter estimates derived from the best-fitting computational model of reinforcement learning. A) Over the course of each block, children and adults learned to make more optimal responses to stimuli in both the category-predictive and exemplar-predictive conditions, though performance was better in category-predictive relative to exemplar-predictive blocks. Points depict age group means and error bars show standard errors of participant means. B) In the category-predictive condition,

children and adults increasingly generalized learned category responses to respond optimally to novel stimuli. Points depict age group means and error bars show standard errors of participant means. The lines show the best-fitting linear regressions through the points, with the shaded region depicting 95% confidence intervals. C) Category "win-stay lose-shift" behavior increased across trials in category-predictive blocks and decreased across trials in exemplar-predictive blocks, increasingly so with age. The boxplot hinges show the first and third quartiles of the data averaged within each trial group; whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians. D) In the category-predictive block condition, participants with higher category-level choice weights and higher exemplar-level choice weights earned more points. In the exemplar-predictive block condition, participants with higher exemplar-level choice weights earned more points. Points show individual participants' total points summed across the three blocks within each condition; lines show the best-fitting linear regressions through the points, with the shaded region depicting 95% confidence intervals. E) Participants across age demonstrated higher category-level choice weights in category-predictive blocks, indicating that they increased the weight they placed on category-level information during decision-making when doing so was useful. The height of the shaded bars show the number of participants whose category (top row) and exemplar (bottom row) choice weight magnitudes fell within each bin in each block condition. The thick colored lines show the average choice weight magnitudes within each age group and block condition.

Flexibility in the specificity of learning representations

Taken together, our learning data suggest that participants across age could use both categorical and exemplar-level information to learn to respond optimally to each stimulus. To what extent did participants flexibly shift the extent to which they weighted categorical versus exemplar-level information when making decisions across block conditions? To address this central question, we fit our data with variants of a reinforcement-learning model that differentially weighted information across levels of abstraction during choice (see Methods). Briefly, all model variants assumed that participants tracked the value of approaching each stimulus at both the categorical and exemplar level, such that on every trial, they incrementally updated one of three categorical value estimates and one of fifteen exemplar-level value estimates based on the reward feedback they received. At choice, these value estimates were converted to choice probabilities via a softmax function with inverse temperature parameters (which we will refer to as 'choice weights') that determined the extent to which decisions were guided by categorical and exemplar-level value estimates. We fit variants of the model with a single choice weight (in which equal weight was placed on categorical and exemplar-level value estimates), two choice weights (in which the weights placed on categorical and exemplar-level value estimates differed) and four choice weights (in which the weights placed on categorical and exemplar-level value estimates differed *and* varied across block conditions). We used a Bayesian model-fitting and selection procedure (see Methods) to determine the best-fitting model at the group level. Relative to models with one and two choice weights, the four-choice-weight model had an exceedance probability of 1, indicating that it was the most frequent, best-fitting model across participants.

Choice weights derived from the best-fitting, four-choice-weight model related to task performance. Participants with higher category choice weights earned more points in category-predictive blocks, b = 38.8, SE = 4.8, $t(149) = 8.0$, $p < .001$, but not exemplar-predictive blocks, b = -6.6, SE = 6.9, $t(149) = -1.0$, $p = .339$ (Figure 2D). Participants with higher exemplar choice weights earned more points in both exemplar-predictive (b = 58.9, SE = 6.9, $t(149) = 8.5$, $p < .001$) and category-predictive blocks (b = 10.8, SE = 4.4, $t(149) = 2.5$, $p = .014$; Figure 2D).

Participants' category and exemplar choice weights varied across block conditions, indicating that they shifted the specificity of the representations used to guide choice in accordance with the reward structure of the learning environment (block condition x choice weight interaction effect, $\beta = .32$ (SE = .04), $F(1, 447) = 76.5$, $p < .001$; Figure 2E). Post-hoc analyses in which we separately examined category and exemplar choice weights indicated that participants had higher category choice weights in category- versus exemplar-predictive choice blocks ($\beta = .57$ (.05), $F(1, 150) = 125.1$, $p < .001$). Exemplar choice weights, however, did not significantly vary across block conditions ($\beta = -.08$ (.04), $F(1, 150) = 3.6$, $p = .061$), potentially reflecting the fact that exemplar-level information could be used to effectively gain reward across *both* block conditions. Though we had hypothesized that the flexible weighting of representations at different levels of abstraction would increase across development, we did not observe evidence for an age-varying block condition by choice weight interaction effect (age x block condition x choice weight: $\beta = .06$ (.04), $F(1, 447) = 2.33$, $p = .127$); participants across age shifted the weight they placed on more specific versus more general representations in accordance with the reward structure of the environment (Figure 2E). We did find that older participants demonstrated higher values of choice weights overall, $\beta = .18$ (.05), $F(1, 447) = 156.3$, $p < .001$, in line with prior findings suggesting an age-related decrease in choice stochasticity (Nussenbaum & Hartley, 2019).

An influence of the learning context on memory

Our learning data indicate that the reward statistics of the task environment influenced the specificity of the representations used for value-based choice. Did environmental reward statistics similarly influence the specificity of memory? To address this question, we analyzed data from a test of incidental memory, which was administered online one week after the initial reinforcement-learning task session. Importantly, the memory test included novel *exemplar* foils, which were drawn from the categories participants saw during learning (e.g., novel cows, horses, and goats; Figure 1A) and novel *category* foils, which were drawn from categories from each stimulus set that were not presented (e.g., sheep; Figure 1A). From these two classes of foil images, we constructed categorical and exemplar-level receiver operating characteristic (ROC) curves for each participant by examining their hit rates (i.e., responses to old images) and their false alarm rates (i.e., responses to foils) at each memory response level (1 - 4, 'definitely new', 'maybe new', 'maybe old', 'definitely old'; Figure 1D). We then computed the area under each of these curves (AUC; (Brady et al., 2022), to derive two measures of memory: category memory, which reflected the discrimination of old images from novel category foils, and exemplar memory, which reflected the discrimination of old images from novel exemplars drawn from the same categories they had seen during learning. In addition, we analyzed memory separately for the images (and foils) from category-predictive and exemplar-predictive blocks of the task, to derive measures of category and exemplar memory performance for each participant in each block condition.

We hypothesized that the reward statistics of the learning environment in which the stimuli were originally encountered, as determined by block condition, would affect memory at both levels of specificity. In exemplar-predictive blocks, participants may have paid more attention to the individuating features of each stimulus, encoding them more strongly and potentially preserving those useful details in memory over the week-long delay period. This in turn may have enhanced their ability to distinguish old items from *all* new foils, leading to better memory for stimuli encountered in exemplar-predictive versus category-predictive blocks of the task. In addition, we hypothesized that enhanced representations of the granular features of each stimulus would provide a particular benefit for discriminating old items from novel *exemplar-level* foils, boosting exemplar memory. Thus, we expected to observe both a main effect of block condition and a block condition x specificity interaction effect on memory.

Across task blocks, participants demonstrated better category versus exemplar memory, reflecting the increased difficulty of discriminating old items from novel, within-category exemplars, $\beta = .060$ (.002), $F(1, 451.2) = 605.5$, $p < .001$ (Figure 3A). In line with our first prediction, we observed a main effect of block condition on memory, such that participants were better able to distinguish old and new stimuli from exemplar-predictive versus category-predictive blocks, $\beta = -.008$ (.002), $F(1, 451.2) = 12.0$, $p < .001$ (Figure 3A). In contrast to our second prediction, however, we did not observe a significant block condition x specificity interaction effect, $\beta = .001$ (.002), $F(1, 451.2) = .27$, $p = .602$. Participants demonstrated a similar enhancement of exemplar *and* category memory for stimuli encountered in exemplar-predictive blocks — the reward statistics of the learning environment shaped overall memory, but not memory specificity per se. We additionally observed that overall memory performance improved with age, $\beta = .016$ (.007), $F(1,149.3) = 5.3$, $p = .023$, though the influence of block condition on memory did not significantly vary across development ($p = .085$).

Individual differences in learning influence how reward shapes mnemonic specificity

While our preceding memory analyses take into account the specificity of the representations that were *useful* for learning, they do not take into account the extent to which representations were actually *used* to guide choice. We expected the environment to influence memory via its effects on value-guided learning, meaning that we expected to see the largest influence of block condition on mnemonic specificity for participants who effectively learned the task's reward statistics. To test this prediction, we re-ran our memory accuracy model, but included participants' total number of points earned within each block condition as an interacting fixed effect. Here, we found that participants who earned the most points during learning demonstrated better memory across levels of specificity, $\beta = .009$ (.004), $F(1, 579.2) =$

4.2, p = .04. Critically, however, this benefit was particularly pronounced for exemplar-level information encountered in exemplar-predictive blocks, as evidenced by a points x block condition x specificity interaction, β = .007 (.003), $F(1, 437.9) = 5.6$, p = .018 (Figure 3B). Participants who demonstrated the best learning of the task's reward statistics also showed the strongest influence of the learning context on subsequent memory specificity. The influence of learning performance on memory did not significantly vary with age (ps > .08).

The influence of reinforcement-learning computations on memory increased with age

Next, we asked how individual differences in the representations used for choice influenced the effects of the learning environment on mnemonic specificity. Our analysis of model-derived choice weights revealed heterogeneity in the extent to which participants weighted exemplar-level information (Figure 2D). This heterogeneity may be reflected in subsequent memory specificity, with participants who relied on more specific representations during learning showing enhanced exemplar memory and participants who relied on more general representations during learning showing enhanced category memory.

We first examined how exemplar-level choice weights in each block condition influenced memory by adding them as an interacting fixed effect in our memory model (Supplementary Table 1). We observed a strong effect of choice weight magnitude on memory, with participants with higher exemplar choice weights exhibiting better memory performance at both levels of specificity, β = .018 (.005), $F(1, 577.2) = 14.6$, p < .001 (Figure 3C). We also observed a choice weight x block condition interaction effect, β = -.010 (.003), $F(1, 497.1)$ = 9.6, p = .002, such that the influence of exemplar-level choice weights on memory was greater in exemplar-predictive blocks. The influence of learning on memory varied with age, as evidenced by a choice weight magnitude x age interaction effect, β = .016 (.006), $F(1, 534.7)$ = 8.3, p = .004, and a choice weight magnitude x age x block condition interaction, β = -.008 (.003), $F(1, 500.2) = 5.1$, p = .025. Older participants demonstrated a stronger effect of exemplar choice weight magnitude on memory, particularly in the exemplar-predictive blocks (Figure 3C).

We observed a different pattern of results when we examined how category choice weights influenced memory (Supplementary Table 2; Figure 3C). Here, we found that participants who weighted category-level information most strongly demonstrated better category memory but worse exemplar memory (choice weight magnitude x specificity interaction effect: β = .007 (.003), $F(1, 443.1) = 5.9$, p = .015; Figure 3C). No other effects or interactions reached significance (ps > .053).

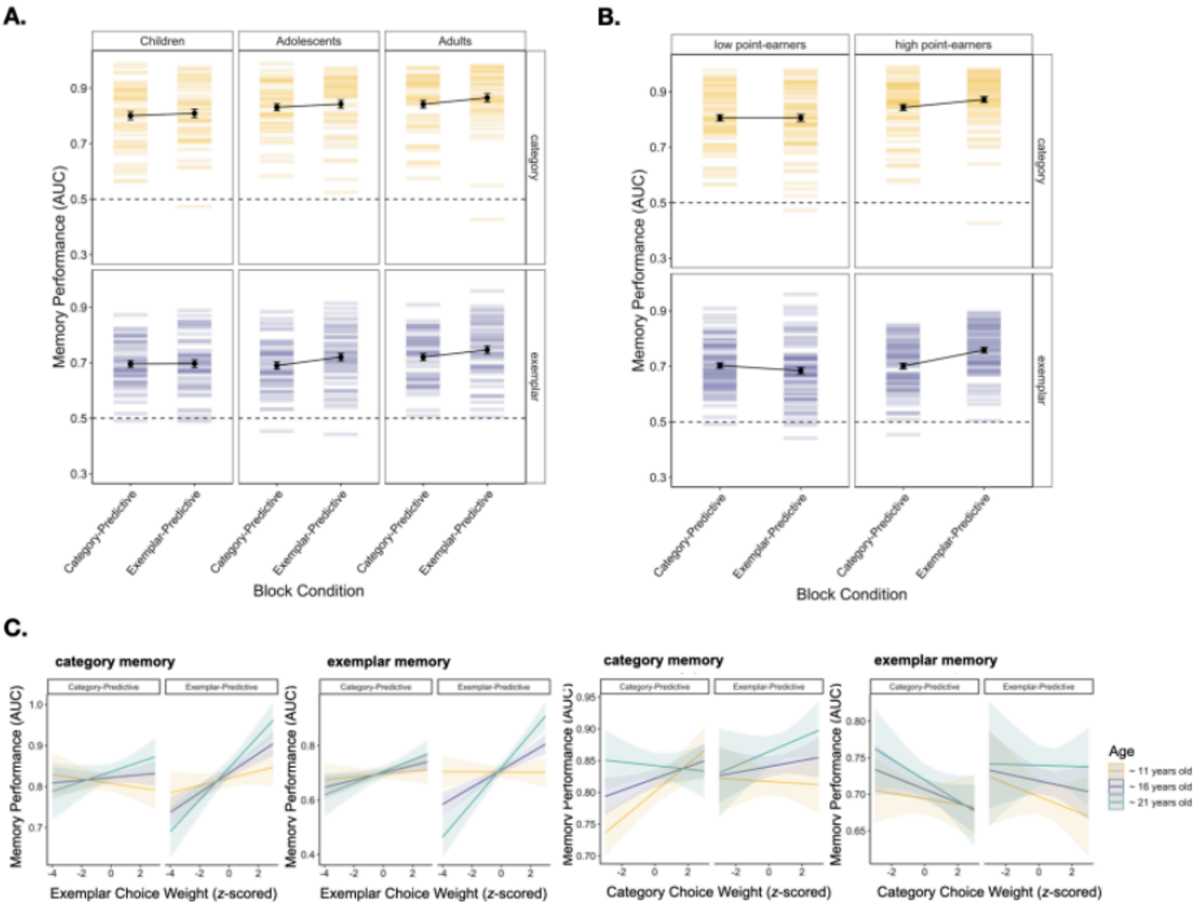


Figure 3. The learning context influenced memory across age. A) Participants demonstrated better memory for category-level versus exemplar-level information, as well as for stimuli from the exemplar-predictive versus category-predictive blocks of the task. Memory at both levels of specificity also improved with increasing age. B) Participants who earned the most points in the exemplar-predictive blocks also demonstrated better memory for exemplar-level information encountered in those blocks. Participants are binned into performance groups based on the number of points earned in each block condition for visualization purposes only. In panels A and B, thin colored lines show individual participants' category (top row) and exemplar (bottom row) memory performance, as indexed by AUC, within each block condition. The black points indicate group means, with the error bars depicting standard errors of participant means. C) Participants who weighted exemplar-level information most strongly demonstrated the best exemplar memory. This effect was stronger in the exemplar-predictive relative to the category-predictive condition, and increased with age. Participants who weighted category-level information most strongly demonstrated better category memory but worse exemplar memory. The plots depict marginal effects from linear-mixed effects models examining the effects of age, block condition, specificity (exemplar and category), choice weight magnitude (exemplar or category), and their interactions on memory performance, as indexed by AUC. Age was analyzed continuously; the plots show the predicted performance of participants at three different ages (the mean age of the sample, +/- 1 SD).

Together, these results support our hypothesis that the statistics of the learning environment influenced memory through their effects on the representations that were used to guide value-based choice. Participants who used exemplar-level representations to the greatest extent during learning also demonstrated the best memory for the exemplars they encountered, particularly in the environment in which specific representations were most useful. Critically, it was *not* the case that participants who were ‘better’ at learning were also better at memory across the board — in category-predictive blocks, higher category choice weights led to better learning performance but *worse* memory for exemplars. The strength of the relation between learning and memory varied across development; the extent to which older participants weighted exemplar-level information during choice more strongly related to their subsequent category and exemplar memory one week later.

Experiment 2

In Experiment 1, we found that people across age adapted the extent to which they weighted exemplar-level versus categorical representations when learning to make good choices, and that individual differences in the specificity of the representations used to guide choice were reflected in subsequent memory. Somewhat unexpectedly, we also found that the strength of the relation between the specificity of the representations used for value-based choice and memory increased with age. In Experiment 2, we aimed to replicate and extend these findings.

Experiment 2 followed the same general structure as Experiment 1, but the reinforcement-learning task differed in several ways (Figure 4). Our Experiment 1 design did not penalize the use of exemplar-level information in category-predictive blocks — the reward statistics of the task meant that in category-predictive blocks, exemplar-level information could still be used to guide optimal decision-making. This may explain why we did not observe shifts in exemplar-level choices weights across conditions, and why we observed global memory enhancements, rather than specificity enhancements, for stimuli encountered in exemplar-predictive blocks. Unlike in Experiment 1, in real-world environments, one advantage to using more abstract representations to guide choice is that they are more robust to stochasticity or ‘noise’ — a single aberrant experience will shift value representations of broader categories to a lesser degree, for a shorter period of time, because one will more rapidly accrue additional experiences with other category members. Further, using more abstract representations is less computationally demanding and requires learning a much smaller set of stimulus-action values. In our Experiment 1 task, exemplar-level reward distributions were not very noisy, and the computational demands of tracking individual exemplars may not have been sufficiently costly for participants to ignore or downweight exemplar-level representations during decision making. Thus, in Experiment 2, we changed the reinforcement-learning task to a) induce more ‘noise’ in reward distributions by making outcomes binary, and b) make tracking exemplar-level information more computationally demanding by having participants select between three actions on every trial (Figure 5C). In addition, because the age effects we observed in Experiment 1 were monotonic, we included only children (n = 34; ages 8 - 12 years) and adults (n = 39; ages 18 - 25 years), between whom we expected to see the largest performance differences.

Figure 4. Experiment 2 task design. A) Each block of the reinforcement-learning task included nine unique stimuli, which comprised three exemplars each drawn from three broader categories. Each stimulus set also included an additional stimulus category with five novel stimuli, as well as two additional novel exemplars per sampled category. B) In the category-predictive condition, rewards on every trial were sampled from Bernoulli distributions with win probabilities determined by the stimulus categories. In the exemplar-predictive condition, rewards on every trial were sampled from Bernoulli distributions with win probabilities determined by the individual exemplars. The optimal action (depicted by the shaded color) resulted in wins on 90% of trials and losses on 10% of trials. The two other actions resulted in wins on 10% of trials and losses on 90% of trials. C.) On every trial of the reinforcement-learning task, participants saw a stimulus and three choice options. After selecting an option, they viewed the outcome of their choice: either a win (+1 point) or a loss (-1 point).

Replication of Experiment 1 learning results

As in Experiment 1, participants made increasingly correct responses across trials (*p*s < .001; Figure 5A), with increasing age (p = .001), and in the category-predictive relative to the exemplar-predictive condition (p < .001). Older participants continued to demonstrate larger benefits from being able to use categorical information to guide choice, OR = .89 (.05), $\chi^2(1)$ = 4.0, p = .045.

Increasingly with age, participants used category values to guide their responses to novel stimuli, demonstrating generalization of correct responses to novel stimuli from previously encountered categories in the category-predictive block (Main effect of category repetition: OR = 1.12 (.04), $\chi^2(1)$ = 10.0, p = .002; category repetition x block condition interaction: OR = 1.26 (.05), $\chi^2(1)$ = 43.0, p < .001; age group x block condition interaction: OR = .89 (.04), $\chi^2(1)$ = 6.8, p = .009; Figure 5B). As in Experiment 1, participants also demonstrated increasing category win-stay lose-shift (WSLS) behavior in the category-predictive blocks and decreasing category WSLS behavior in the exemplar-predictive blocks (trial x block condition interaction effect: OR = 1.14 (.02), $\chi^2(1)$ = 49.0, p < .001), an effect that was stronger in adults than children (block condition x age group interaction: OR = .90 (.03), $\chi^2(1)$ = 10.0, p = .006).

When we fit reinforcement-learning models to the Experiment 2 choice data, the best-fitting model again included four choice weights, reflecting differences in the weighting of categorical and exemplar-level representations across block conditions. As in Experiment 1, choice weights related to task performance: Category choice weights positively related to the number of points participants earned in category-predictive blocks (b = 39.8, SE = 4.7, $t(71)$ = 8.5, p < .001) but not exemplar-predictive blocks (b = -6.8., SE = 3.7, $t(71)$ = -1.8, p = .069).

Exemplar choice weights positively related to the number of points participants earned in both exemplar-predictive (b = 19.2, SE = 2.6, $t(71) = 7.5$, $p < .001$) and category-predictive blocks (b = 16.5, SE = 6.4, $t(71) = 2.6$, $p = .011$).

Participants flexibly adapted the extent to which they weighted categorical versus exemplar-level representations across conditions (block condition x abstraction interaction effect: β = .14 (.05), $F(1, 213) = 6.14$, $p = .014$; Figure 5C). Here, we expected that by making exemplar-level information less useful in category-predictive blocks, we might observe changes in *both* category and exemplar choice weights across block conditions. We found, however, that as in Experiment 1, changes in the weighting of representations across blocks were still largely driven by changes in the extent to which participants weighted *categorical* representations (β = .18 (.06), $F(1, 72) = 7.8$, $p = .007$) rather than the extent to which they weighted exemplar-level representations (β = -.10 (.06), $F(1, 72) = 2.6$, $p = .112$).

Replication of Experiment 1 memory results

Replicating the results of Experiment 1, participants demonstrated better memory for stimuli encountered in exemplar-predictive relative to category-predictive learning blocks, β = -.018 (.004), $F(1, 213) = 20.0$, $p < .001$ (Figure 5D). Here, we also observed marginal support for our initial hypothesis: that *exemplar* memory would be specifically enhanced in exemplar-predictive blocks (block condition x abstraction level interaction: β = .008 (.004), $F(1, 213) = 3.74$, $p = .054$; Figure 6B). This effect did not interact with age group ($p = .439$). When we added participants' total number of points earned within each block condition as an interacting fixed effect, we found that participants who earned the most points demonstrated the best memory, β = .040 (.010), $F(1, 254.5) = 14.9$, $p < .001$, and that as in Experiment 1, this effect was strongest for exemplar-level memory for stimuli encountered in exemplar-predictive blocks, β = .014 (.007), $F(1, 192.4) = 4.1$, $p = .043$ (Figure 5E).

Individual differences in the extent to which participants weighted exemplar-level representations during learning, as indexed by exemplar choice weights, also robustly related to memory, β = .035 (.008), $F(1, 265.4) = 19.1$, $p < .001$ (Figure 5F; Supplementary Table 3). As in Experiment 1, the influence of learning on memory strengthened with age (choice weight magnitude x age group interaction effect: β = -.016 (.008), $F(1, 265.4) = 4.1$, $p = .045$). While this effect was strongest in exemplar-predictive blocks in Experiment 1, here, we did not observe a significant choice weight magnitude x age group x block condition interaction ($p = .181$); older participants demonstrated a stronger effect of exemplar choice weight magnitude on memory across conditions.

Replicating our Experiment 1 findings, we did not observe a main effect of category choice weight on memory ($p = .264$; Supplementary Table 4), but rather a choice weight magnitude x specificity interaction (β = .011 (.004), $F(1, 202.7) = 7.8$, $p = .006$; Figure 5F), indicating that participants who weighted category-level representations most strongly demonstrated better category memory, but not better exemplar memory. Here, we additionally observed an age group x choice weight magnitude x block condition interaction effect, β = .012

(.005), $F(1, 212.9) = 6.5$, $p = .012$ (Figure 5F), such that children demonstrated a more positive influence of category choice weight magnitude on memory in the category-predictive blocks. As in Experiment 1, while heavily weighting exemplar-level representations during learning enhanced memory *across* levels of specificity, heavily weighting *categorical* representations during learning only boosted category memory.

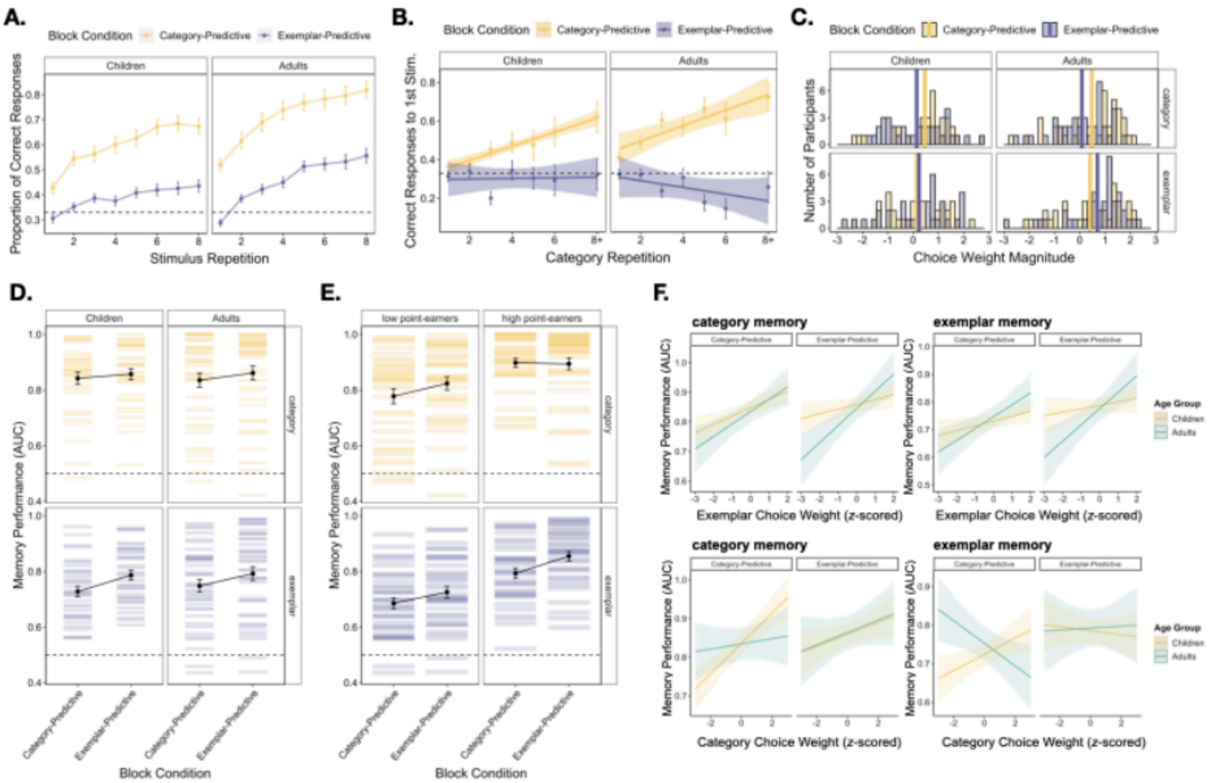


Figure 5. Results from Experiment 2 replicated key findings from Experiment 1. A) Over the course of each block, children and adults learned to make more optimal responses to stimuli in both the category-predictive and exemplar-predictive conditions, though performance was better in category-predictive relative to exemplar-predictive blocks. Points depict age group means and error bars show standard errors of participant means. B) In the category-predictive condition, participants increasingly generalized learned category responses to respond optimally to novel stimuli. Points depict age group means and error bars show standard errors of participant means. The lines show the best-fitting linear regressions through the points, with the shaded region depicting 95% confidence intervals. C) Participants across age groups demonstrated higher category-level choice weights in category-predictive blocks, indicating that they increased the weight they placed on category-level information during decision-making when doing so was useful. The height of the shaded bars show the number of participants whose category (top row) and exemplar (bottom row) choice weight magnitudes fell within each bin in each block condition. The thick colored lines show the average choice weight magnitudes within each age group and block condition. D) Participants demonstrated better memory for stimuli from the exemplar-predictive versus category-predictive blocks of the task. E) Participants who earned the most points in the exemplar-predictive blocks also demonstrated better memory for exemplar-level information encountered in those blocks. Participants are binned into performance groups based on the number of points earned in each block condition for visualization purposes only. In panels D and E, thin colored lines show individual participants' category (top row) and exemplar (bottom row) memory performance, as indexed by AUC, within each block condition. The black points indicate group means, with the error bars depicting standard errors of participant means. F) Participants who weighted exemplar-level information most strongly during learning

also demonstrated better category and better exemplar memory. The strength of this relation between learning and memory increased with increasing age. Participants who weighted category-level information most strongly demonstrated better category memory but not better exemplar memory. The plots depict marginal effects from linear-mixed effects models examining the effects of age group, block condition, specificity (exemplar and category), choice weight magnitude (exemplar or category), and their interactions on memory performance, as indexed by AUC.

In Experiment 2, we made learning the optimal responses to individual exemplars in the reinforcement-learning task more costly by making reward values binary and presenting three choice options on every trial. Despite these differences from the task used in Experiment 1, we continued to observe adaptive flexibility in participants' weighting of representations at different levels of abstraction, as well as reflections of learning weights in subsequent memory specificity. Critically, we also replicated our finding that the coupling between reinforcement learning and mnemonic specificity increased with age: Individual differences in the extent to which people weighted exemplar-level representations during learning were more tightly linked to individual differences in memory in adults versus children.

Discussion

Across two developmental studies, we examined how the specificity of the representations used for value-guided learning and memory are shaped by the statistics of the environment. We found that from childhood to early adulthood, participants adapted their learning representations to match the level of abstraction most useful for guiding behavior across environments. Originally, we hypothesized that more specific information would be preserved in memory only when it was useful for adaptive choice. We found, however, that specific information was remembered not when it was *useful*, but rather when it was *used*: Using specific representations to guide reward learning improved memory for both category and exemplar-level information. Using broader, categorical representations for learning boosted category memory only, and in some cases, even impaired exemplar memory. Moreover, the strength of the relation between learning and memory increased with age, such that relative to children, adults demonstrated a stronger influence of the specificity of their reinforcement-learning representations on subsequent memory. These findings suggest that the environment shapes memory specificity through its influence on reward learning, with the strength of the coupling between learning and memory increasing across development.

Our experiments revealed early-emerging flexibility in the specificity of value-guided learning. One challenge for learning within complex environments is determining which stimulus dimensions are relevant for choice (Radulescu et al., 2019). In our learning task, there were no explicit cues that signaled whether idiosyncratic exemplar features or more general stimulus categories determined reward contingencies; instead, participants had to learn through experience the specificity of the representations that could most effectively guide choice. We expected that over the course of each block, reciprocal interactions between attention and reinforcement learning would increasingly cause participants to attend to either the shared or individuating features of stimuli within a category (Leong et al., 2017; Mack et al., 2020; Niv et al., 2015). We found that participants across age demonstrated adaptive up- and down-weighting of categorical information based on the environment's reward structure. In contrast to prior work (Schiele et al., 2016), we did not observe evidence for greater generality in children's learning — older participants demonstrated stronger generalization of learned categorical responses to novel stimuli. Moreover, younger participants had lower exemplar choice weights *and* lower category choice weights, indicating that their poorer learning performance was driven by greater choice stochasticity (Eckstein et al., 2022; Giron et al., 2023; Nussenbaum & Hartley, 2019) rather than a bias toward overgenerality. Our work builds on prior research demonstrating that adults can learn reward contingencies across multiple levels of abstraction (Eckstein & Collins, 2020; M. J. Frank & Badre, 2012); here, we extend these findings and show that across development, in accordance with the predictions of theoretical models (Santoro et al., 2016), individuals can flexibly arbitrate between more specific and more general representations to guide behavior.

Here we extended past work showing that reward learning shapes memory (Davidow et al., 2016; Jang et al., 2019; Rosenbaum et al., 2022; Rouhani et al., 2018; Rouhani & Niv, 2021), demonstrating that memory reflects the level of abstraction of reward-learning computations. When participants used specific representations for choice, they preserved more detailed information in memory, whereas when they used more abstract representations, they demonstrated better generalization but poorer memory for individual exemplars. This tight link between between learning and mnemonic specificity aligns with the predictions of models of categorization; exemplar-based models (Kruschke, 1992; Medin & Schaffer, 1978; Nosofsky, 1986) posit that memories for individual exemplars facilitate inferences about novel instances, whereas prototype models (Rosch, 1975; Rosch et al., 1976) suggest that individuals store and use more abstracted features to represent meaningful groupings of the world. More recent category-learning models posit adaptive flexibility in representations, such that individuating features are represented only when needed for successful classification and inference (Love et al., 2004). A key property of all these models is that the way in which the world is parsed directly influences the specificity of the representations that are stored in memory over time. Merging multiple conceptual frameworks that propose mechanistic links between learning computations and memory (Gershman & Daw, 2017; Love & Gureckis, 2007), our work demonstrates that across development, reward shapes the granularity with which the world is partitioned, and in turn, the specificity of the information preserved in memory.

Moreover, we found that the strength of the relation between learning and memory specificity increased across development, which may be due to age-related increases in the influence of goals on feature-based selective attention at older ages (Plude et al., 1994; Wharton-Shukster & Finn, 2019). Adults may have learned through experience to attend to the information most useful for guiding choice, such that their exemplar choice weight magnitudes reflected the extent to which they both used *and* attended to exemplar-level information. Children, however, may have still *attended* to individuating features of stimuli even after learning that such features were irrelevant for decision-making. Indeed, prior research has suggested that relative to adults, children demonstrate broader patterns of attention, such that they attend to and learn about information that is irrelevant for the task at hand (Blanco & Sloutsky, 2019; Deng & Sloutsky, 2016; S. M. Frank et al., 2021; Plebanek & Sloutsky, 2017; Sloutsky & Fisher, 2004; Tandoc et al., 2022). This greater breadth of attention also influences memory, with children demonstrating better subsequent memory for information that they were *not* cued to attend to during learning (Deng & Sloutsky, 2016; Plebanek & Sloutsky, 2017; Sloutsky & Fisher, 2004). In our task, children may have shown a greater dissociation between the representations used for choice and their allocation of selective attention during learning. Their learning representations may therefore relate less strongly to the specificity of their subsequent memory. Future work can more directly test hypotheses about age-related change in attention during learning by using stimuli with spatially segregated features (e.g., (Leong et al., 2017) and measuring how differences in patterns of visual gaze during learning relate to subsequent memory specificity.

In our task, participants completed the memory test after a one-week delay. Prior research has suggested that both the influence of reward statistics on memory (Murayama & Kitagami, 2014; Murty et al., 2017; Patil et al., 2017) and individual differences in memory specificity (Callaghan et al., 2021), may strengthen as the delay between encoding and retrieval increases. The strengthening of these effects over time suggests that post-encoding consolidation processes play an important role in mnemonic specificity. Models of systems consolidation suggest that over time, memories may increasingly reflect generalized knowledge extracted from commonalities across multiple reactivated episodes (Dudai, 2012; Squire et al., 2015). In our experiments, it may be the case that age differences in the influence of reinforcement learning on memory were partially driven by age-related change in consolidation in the week between the learning and memory tasks (Callaghan et al., 2021; Cohen et al., 2022). It may be the case that in adults, useful information is more strongly prioritized during consolidation, such that representations used to guide decision making are 'replayed' or reactivated (Liu et al., 2021; Mattar & Daw, 2018; Murty et al., 2017; Sterpenich et al., 2021) to a greater extent than in children. Because all participants in our experiments completed the memory task after a week-long delay, we cannot determine how encoding versus consolidation mechanisms may have differentially contributed to memory for specific versus more general information across age. Future studies can test memory at different delay periods to examine the influence of consolidation time on the development of adaptive mnemonic specificity.

The ability to flexibly adjust the specificity of value learning and episodic memory is critical for building adaptive mental models of the environment across the lifespan. Here, we demonstrate that children, adolescents, and adults can dynamically adapt the relative weight they place on more specific versus more general information during reward learning. Further, we show that across development, the specificity of memory is increasingly shaped by the specificity of learning computations. The coupling of early-emerging flexibility in learning and a more protracted developmental timecourse of the influence of learning on memory may be adaptive (Nussenbaum, Prentis, et al., 2020; Nussenbaum & Hartley, 2021). Memory that is less constrained by beliefs about the usefulness of information may promote the acquisition of broad knowledge of the world, whiled protecting against adverse consequences of learning representations at ineffective levels of abstraction. Across development, individuals' adaptive parsing of the world's structure increasingly shapes memory, and these lasting traces may guide adaptive behavior over increasingly long timescales.

Methods

Experiment 1

Participants

A priori, we determined a target sample size of 150 participants based on our prior studies of learning across age-continuous samples of children, adolescents, and adults (Nussenbaum et al., 2022; Nussenbaum, Scheuplein, Phaneuf, et al., 2020). One hundred and fifty-one participants aged 8 - 25 years completed the two-part online study and were included in all analyses. An additional 24 participants completed both parts of the study but were excluded from all analyses for: (a) interacting with their browser window (minimized, maximized, or clicked outside the window) more than 20 times throughout either the learning or memory task (n = 15), (b) failing to respond on more than 10% of the 306 learning trials (> 30 trials) or 10% of the 192 memory trials (> 20 trials) (n = 7), or (c) responding in less than 100 ms on more than 10% of learning or memory trials (n = 2). In addition, one additional participant was excluded due to a glitch that prevented data from being saved. Participants were compensated with a $20 Amazon gift card for completing both parts of the study. They also received a bonus that ranged from $0 to $5 depending on their performance in the learning task.

The 151 participants included in the final sample comprised n = 50 children (8.0 - 13.0 years; Mean age = 10.4 years, n = 24 females), n = 50 adolescents (13.1 - 17.8 years; Mean age = 15.4 years, n = 28 females), and n = 51 adults (18.2 - 25.9 years, Mean age = 21.8 years, n = 31 females). All participants reported normal or corrected-to-normal vision and no diagnosed psychiatric or learning disorders. 57.6% of participants were White, 22.5% were Asian, 10.6% were Black and 9.3% were two or more races. In addition, 10.6% of participants were Hispanic. We include a more detailed description of participant demographics in the Supplement.

As with our previous online studies (Nussenbaum et al., 2022; Nussenbaum, Scheuplein, & Phaneuf, 2020), participants were primarily recruited from ads on Facebook and Instagram, as well as via word-of-mouth, science fairs, events, and fliers distributed around New York University. Prior to entering our participant database and being eligible to complete the online study, all potential participants completed a 5-minute zoom call with a researcher. During this zoom call, all participants (and a parent or guardian, if the participant was under 18 years of age) were required to be on camera and confirm the full name and date of birth they provided when they signed up for our database. Adult participants and parents of child and adolescent participants were further required to show photo identification.

Experimental procedure

Participants completed three experimental tasks across two sessions. All tasks were coded in jsPsych (de Leeuw, 2015) and hosted on Pavlovia. In the first session, participants completed a reinforcement-learning task, which took approximately 40 minutes. In the second session, participants completed a test of recognition memory, which took approximately 15 minutes. Participants who completed the learning task during the first session were invited to complete the second session six days later and had five days to complete it (e.g., if a participant completed the first session on a Wednesday, they would be invited to participate in the second session on Tuesday, and would have until the following Saturday to complete it). On average, participants completed the second session 7.1 days after completing the second session.

Reinforcement-learning task

To examine how participants used categorical and exemplar-level representations to guide learning, we developed a value-based learning task in which participants had to choose whether to 'approach' or 'avoid' a stimulus on every trial. If participants chose to 'approach' the stimulus, they would win or lose points depending on its value. If they chose to 'avoid' the stimulus, they would not win or lose any points, but they were provided with full counterfactual information, meaning they would see how many points they *would have* won or lost had they chosen to approach the stimulus.

The task comprised six blocks, each with its own stimulus set (Table 1). The stimulus set assigned to each of the six blocks was randomized for each participant. The six stimulus sets included 32 unique images, divided into four broader categories (Figure 1A). The broader categories were selected to be familiar to children as young as 8 years old (Figure 1E). All stimulus images were taken from Google images and edited such that they showed a single item on a white, square, uniformly sized background. The instructions for each task block followed the same format but varied depending on the stimulus set. For example, in the 'Pets' block, participants were instructed that petting animals would sometimes make them happy, causing them to win points, and sometimes make them angry, causing them to lose points, whereas in the 'Vehicles' block, participants were instructed that taking their friend for a ride in some vehicles would make them thrilled and other vehicles would make them upset.

Table 1: Stimulus sets used across blocks

Block	Stimulus categories
farm animals	cows, goats, horses, sheep
fruit	apples, bananas, oranges, strawberries
furniture	beds, chairs, sofas, tables
pets	cats, dogs, rabbits, rodents
plants	bushes, cacti, flowers, trees
vehicles	boats, planes, trains, trucks

In each block of the reinforcement-learning task, participants saw 15 unique images. For each participant, five images were randomly selected from three of the four categories in each block to serve as learning stimuli. Within each category, two images repeated five times, one image repeated three times, and two images were only shown once during learning. The order of image presentation was randomized within each block for each participant.

Critically, participants completed three blocks in the *category-predictive* condition and three blocks in the *exemplar-predictive* condition. In the category-predictive condition, stimulus values were sampled from Gaussian distributions (SD = 1.5) on every trial, where the mean of the distribution was determined by stimulus category. One category was randomly determined to be ‘good’ such that the mean of its reward distribution was between 3 and 6; one category was randomly determined to be ‘neutral’ such that the mean of its reward distribution was zero (though zero was never actually presented as an outcome); and one category was randomly determined to be ‘bad’ such that the mean of its reward distribution was between -6 and -3. Values were rounded to the nearest non-zero integer. Values were sampled from these distributions anew on every trial, meaning the reward associated with approaching the same stimulus might differ across repetitions.

In the exemplar-predictive condition, each stimulus was pseudo-randomly assigned a deterministic reward value between -9 and 9. To ensure that categorical information could not be used to effectively guide choice, one stimulus within each category was assigned a value between -9 and -6, one was assigned a value between -5 and -3, one was assigned a value between -2 and 2, one was assigned a value between 3 and 5, and one was assigned a value between 6 and 9. In addition, within each block, no two stimuli were assigned the same value, and no stimulus was assigned a value of zero. This meant that all broader categories included two or three stimuli that should be avoided, and two or three stimuli that should be approached.

The condition of the first block was counterbalanced across participants within each age group, such that roughly half of children, adolescents, and adults experienced a category-predictive block first, and the other half experienced an exemplar-predictive block first. For each participant, the first two blocks of the task were always different conditions. The latter four blocks included two additional exemplar-predictive blocks and two additional category-predictive blocks, in a random order.

To ensure participants had equal exposure to all stimuli, all trials lasted 3 seconds, regardless of how quickly participants made their response. Within the 3-second time limit, participants made their 'approach' or 'avoid' selection by pressing 1 or 0 on a standard keyboard, respectively. After making their selection, participants saw their choice highlighted for 500 ms, and then the outcome of their choice for the remainder of the trial (Figure 1B). For 'approach' decisions, winning outcomes were displayed in green text and losses were displayed in red text. For 'avoid' decisions, the points that the participant *would have* won or lost were always displayed in gray text, inside a red or green box. The colors of the boxes corresponded to whether they made an optimal or suboptimal choice on that trial. *Missed* wins were displayed in red boxes and avoided losses were displayed in green boxes — this color cue was intended to help participants across age with counterfactual learning. In addition, the choice screen displayed coins in each of its corners, which were animated depending on the choice outcome: The coins would bounce for wins, fall off the screen for losses, and become grayed out for 'avoid' decisions. Trials were separated by a 500 ms inter-trial interval in which no stimuli appeared on the screen. Participants lost 5 points each time they failed to respond within the 3-second time limit.

Prior to completing the real trials of the learning task, participants completed an extensive tutorial, which included child-friendly instructions that were both written on the screen and read aloud via audio recordings. Participants were unable to advance past each instruction page until the audio recording finished playing. The tutorial also included a short practice block in which participants had to approach or avoid different pieces of sports equipment. Participants completed twelve practice trials. Stimuli on each trial were sampled (with replacement) from 8 images across two categories (balls and rackets), and their values were randomly sampled on each trial from -9 to 9, with replacement. In this way, the reward structure of the practice block did not align with either the category-predictive or exemplar-predictive condition, but still allowed participants to learn the mechanics of the learning task. After the tutorial, participants answered three True/False comprehension questions about the task. If they answered a question incorrectly, they would see (and hear) the correct answer with an explanation, and have to try to answer the same question again. On average, participants answered all three comprehension questions correctly in 3.05 attempts (Mean number of attempts: Children: 3.06; Adolescents: 3.08; Adults: 3.00). There was not a significant effect of age on the number of question attempts required ($b = -.005$, $SE = .003$, $p = .15$).

Memory test

In the second experimental session, which took place between 6 and 10 days after the first (Mean delay = 7.1 days, SD = 1.3 days) participants completed a test of recognition memory (see Supplement for an analysis of the effects of delay duration on memory). On each trial, participants saw an image and had to determine whether it was Definitely New, Maybe New, Maybe Old, or Definitely Old, by pressing the 1, 2, 3, and 4 keys on their keyboard, respectively (Figure 1D). Participants had 10 seconds to make each response. They did not receive any feedback.

The memory test comprised 192 trials, which included all 32 images from each of the six stimulus sets. This meant that for each of the six stimulus sets, participants saw the 15 old images used during the learning task, 9 new exemplars from the three presented categories, and 8 new images from a fourth category that was not presented during learning. All images from all six stimulus sets were intermixed and presented in a random order at test. As with the learning task, participants completed a child-friendly tutorial and several practice trials prior to beginning the real memory test.

Analysis Approach

Learning

To examine participant performance during the learning task, we coded a “correct response” variable as 1 if participants chose to approach stimuli with positive values and avoid stimuli with negative values, and 0 if they chose to avoid stimuli with positive values and approach stimuli with negative values. For analyses of correct responses, we excluded trials involving stimuli from the neutral category in the category-predictive condition.

Memory

To examine memory performance, we used participants’ memory confidence ratings to construct receiver operating characteristic curves for each participant by computing the proportion of old and new images that they responded to at or below each confidence level (ranging from 1, definitely new, to 4, definitely old). We constructed four separate curves for each participant: one for each block condition (category-predictive, exemplar-predictive) at each level of stimulus abstraction (category, exemplar). We analyzed levels of stimulus abstraction separately to probe the *specificity* of memory representations — we aimed to examine, for example, whether participants remembered that they had seen images from broader categories (e.g., cows) or whether they had seen specific exemplars (e.g., a specific black-and-white cow). The same sets of “old” images were included in the analyses across both levels of abstraction, but the novel foils that were included in the computation varied: Novel category foils were used to construct the category ROC curves and novel exemplar foils were used to construct the exemplar ROC curves (Table 2). In this way, we could examine whether participants could distinguish old from new categories (e.g., cows vs. sheep) separately from whether they could distinguish old from new exemplars (e.g., an old black-and-white cow from a new brown cow). We then used the ‘pROC’ R package (Robin et al., 2011) to compute the area under each of these curves (AUC), as our measure of memory performance. AUC is a theory-neutral metric of memory performance that avoids the incorrect "all-or-none" (i.e., remembered or forgotten) assumptions that are inherent to measures like corrected recognition, and is instead sensitive to graded confidence levels (Brady et al., 2022). AUC values of 1 indicate perfect memory, while values of .5 indicate chance-level performance.

In addition to the memory analyses described in the results, we also examined how memory varied as a function of the delay (in days) between the learning and memory tasks, and as a function of the number of times each stimulus repeated during the learning task. While we observed effects of both delay (i.e., worse memory with increasing delays) and stimulus repetition (i.e., better memory with increasing repetitions), these effects did not interact with any of our predictors of interest. As such, for simplicity, we collapsed across these variables in the models described in the main text, and report delay and repetition analyses in the Supplement.

Table 2: Stimuli used in memory analyses

Block condition	Memory specificity	Old images	New images
Category-predictive	Category	45 old images from category-predictive learning blocks	24 **novel category foils** from stimulus sets used in category-predictive learning blocks
	Exemplar		27 **novel exemplar foils** from stimulus sets used in category-predictive learning blocks
Exemplar-predictive	Category	45 old images from exemplar-predictive learning blocks	24 **novel category foils** from stimulus sets used in exemplar-predictive learning blocks
	Exemplar		27 **novel exemplar foils** from stimulus sets used in exemplar-predictive learning blocks

Mixed-effects models

We used the 'afex' package for R (Singmann et al., 2020) to fit mixed-effects models to our data. All continuous variables were *z*-scored prior to their inclusion in the models. Models included random intercepts for each participant and random slopes across fixed effects and their interactions for each participant. When models failed to converge, we pruned interactions between random slopes and then random slopes themselves (Barr et al., 2013). For logistic mixed-effects models, we assessed the significance of fixed effects with likelihood ratio tests. For linear mixed-effects models, we assessed the significance of fixed effects with *F* tests using Satterthwaite approximation to estimate the degrees of freedom.

Reinforcement-learning models

Model-fitting and comparison methods. To test how participants learned and used categorical and exemplar-level information to choose whether to approach or avoid each stimulus, we fit our data with variants of a temporal difference reinforcement learning model (Sutton et al., 1998). Due to the number of model variants we considered, we performed model comparison and selection in stages, which we describe below. Model-fitting and comparison were conducted using the computational and behavioral modeling (cbm) package (Piray et al., 2019) within Matlab 2020b (The MathWorks Inc., 2020). Because the cbm package relies on normally distributed parameters, within each model, we exponentiated choice weight parameters to ensure they were positive, and transformed learning rate parameters to be between 0 and 1, using sigmoidal functions. We first fit all models to each participant's choice data individually. For first-level fitting, we used common, relatively uninformative priors for all model parameters: Normal(mean = 0, variance = 6.25). We also scaled all reward outcomes to be between -1 and 1 by dividing by the maximum absolute reward value that participants experienced (11). We similarly scaled initial Q values within each model by dividing them by 10.

These first-level fits were then fed into a second-level fitting and model comparison algorithm. The second-level fitting procedure performs simultaneous hierarchical parameter estimation and Bayesian model comparison, in which each participant is treated as a random effect (i.e., different participants may be best fit by different models). We determined the best-fitting model at the group level by examining exceedance probabilities (XP), which reflect the probability that a given model is the most frequent best-fitting model for a group of participants (Rigoux et al., 2014).

First-stage model comparison: Choice weights. The baseline model assumed that participants tracked the overall value of each category of stimuli (three value estimates per block) as well as the value of each individual exemplar (nine value estimates per block). On every trial, the probability that a participant would approach the stimulus was determined via a softmax function with choice weight parameters (inverse temperatures; β_c, β_e) that scaled the category- and exemplar-level value estimates ($Q(c)$, $Q(e)$):

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

The choice weights thus govern the extent to which category-level and exemplar-level information influence choices, with higher weights indicating choices that are more driven by the category- and exemplar-level value estimates.

After choosing to approach or avoid each stimulus, participants update their category-level and exemplar-level value estimates such that:

$$Q(c)_{t+1} = Q(c)_t + \alpha * (r - Q(c)_t)$$
$$Q(e)_{t+1} = Q(e)_t + \alpha * (r - Q(e)_t)$$

where r is the reward for approaching the stimulus on that trial and α is a participant-specific learning rate, governing the extent to which recent rewards influence value estimates.

We tested variants of this model with one choice weight, two, and four choice weights, (allowing them to vary across levels of abstraction and allowing them to vary across *both* levels of abstraction and block conditions, respectively). At the group level, the best-fitting model included four choice weights (i.e., 'fourB', model frequency = .78; exceedance probability (XP) = 1).

Second-stage model comparison: Initial Q values. Next, we assessed whether model fit could be further improved by allowing initial stimulus value estimates to vary (rather than being fixed at 0). We tested three variants of the fourB model that allowed for 0, 1, or 2 initial Q values (across levels of abstraction). At the group-level, the best-fitting model included a single free parameter for initial Q values (i.e., 'fourB_oneQ', model frequency = .70, XP = 1).

Third-stage model comparison: Counterfactual learning. Finally, we assessed whether participants updated value estimates equivalently after approach and avoid decisions. We tested three variants of the fourB_oneQ model: the winning model from our prior comparisons, which included full counterfactual learning, as well as a model that allowed for separate learning rates after approach versus avoid decisions, and a model with *no* counterfactual learning, in which value estimates were not updated on 'avoid' trials. At the group-level, the best-fitting model included full counterfactual learning (fourB_oneQ, model frequency = .76, XP = 1).

Model recoverability. To ensure that models within each comparison set were distinguishable from one another, we conducted recoverability analyses in which we generated 100 simulated datasets of 151 simulated agents. Parameters for each simulated agent were drawn randomly from uniform distributions with minima and maxima determined by the minimum and maximum fitted values from the empirical data. In addition, for the model with two initial q values, we constrained their non-transformed values to be at least 1 apart. For the model with a separate counterfactual learning rate, we constrained the non-transformed values of the counterfactual learning rates to be greater than -2 (so that its transformed value would be distinguishable from 0) and to be at least 1 apart from the choice learning rate. We then fit each simulated dataset with all three models in each model comparison set, and fed these first-level fits through the same second-level fitting and model comparison algorithm that we used with our empirical data. We then examined the proportion of the 100 simulated experiments for which the model with the highest exceedance probability matched the true, generating model.

The model with four choice weights was highly distinguishable from the models with one or two choice weights (Figure 6A). Similarly, the model with full counterfactual learning was highly distinguishable from the model with a separate counterfactual learning rate and the model without counterfactual learning (Figure 6C). However, while distinguishable from the model that initialized value estimates at 0, the model with one initial q value was not distinguishable from the model with two initial q values (Figure 6B), indicating that with our experimental design, we could not measure whether participants initialized categorical stimulus values differently from exemplar stimulus values.

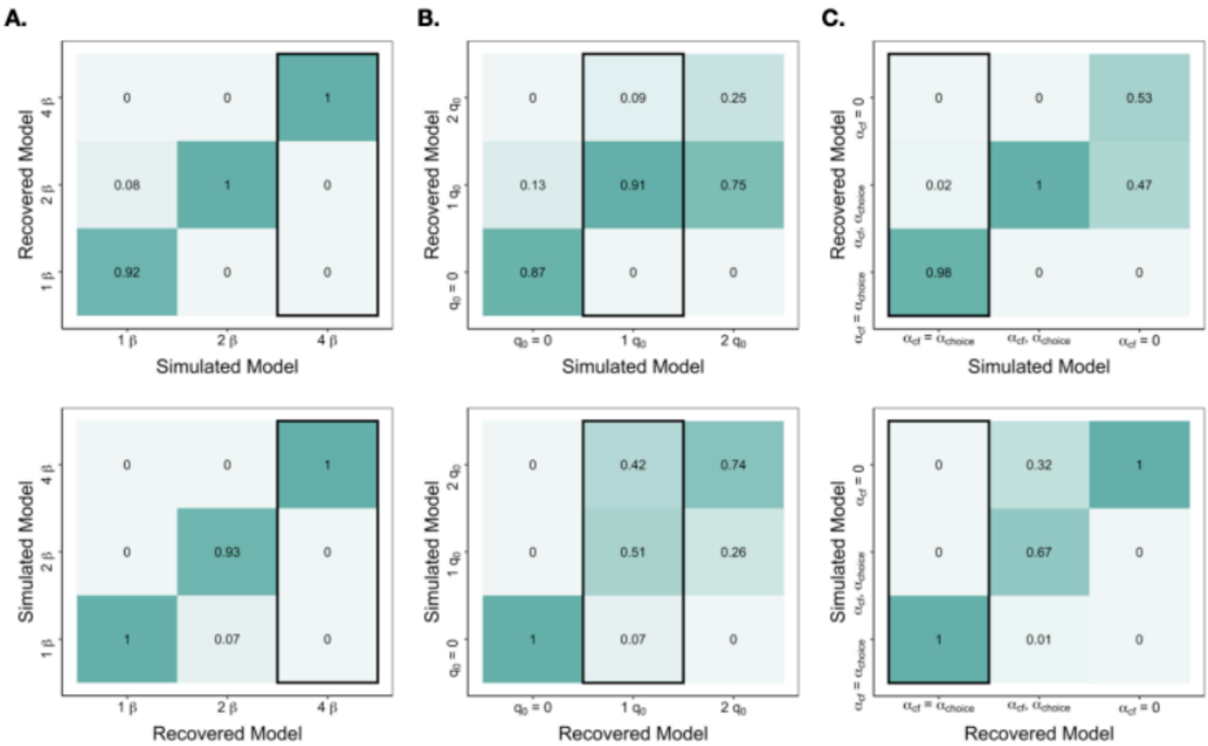


Figure 6. Experiment 1 model recoverability. For each model within each stage of model comparison, 100 simulated ‘experiments’ were conducted in which choice data was simulated from 151 agents, with parameters sampled from uniform distributions with ranges determined by the empirical fits. Data from each simulated experiment were then fit with each model within the comparison set. The top panels show confusion matrices, where the values within each tile represent the proportion of experiments for which each fitted model had the highest exceedance probability (top panels). The bottom panels show inversion matrices, where the values within each tile represent the proportion of experiments for which the fitted model had the highest exceedance probability that were generated by each of the models. Black lines outline the model that best-fit the empirical data within each comparison stage. A) Models with different numbers of choice weights were highly distinguishable from one another. B) Models in which exemplar and category values were initialized with either one or two free parameters were distinguishable from a model in which exemplar and category values were both initialized at 0. However, models with one or two initial values were not distinguishable from one another. C) The winning model, in which participants learned equivalently from experienced and counterfactual outcomes, was highly distinguishable from a model in which participants learned with separate learning rates for experienced and counterfactual outcomes, as well as a model in which participants did not learn from counterfactual outcomes.

Parameter estimates and parameter recoverability. We extracted parameter estimates from our best-fitting model (fourB_oneQ) to examine their relation with other variables of interest (e.g., age, memory performance). Because we were interested in individual differences in parameter estimates, we examined estimates from the first level of model-fitting, in which models were fit to individual participants' data using common, uninformative priors.

To ensure that parameter values were recoverable (Wilson & Collins, 2019), we simulated data from 15,100 participants (e.g., 100 simulations of 151-participant 'experiments'). For each simulated participant, we randomly sampled a task stimulus and reward sequence from one of our participants. We sampled parameter values from uniform distributions with minima and maxima determined by the minimum and maximum parameter estimates from fitted participant data. We then performed first-level model-fitting on these simulated datasets, and examined the correlation between simulated and recovered parameter values. Across all parameters, recoverability was high, with correlations ranging from .73 to .91 (Figure 7A). When exemplar choice weights were high (e.g., > 1), exemplar values dictated participants' choices and category choice weights were pulled toward the mean of the prior (0), such that they demonstrated poorer recoverability (Figure 7B).

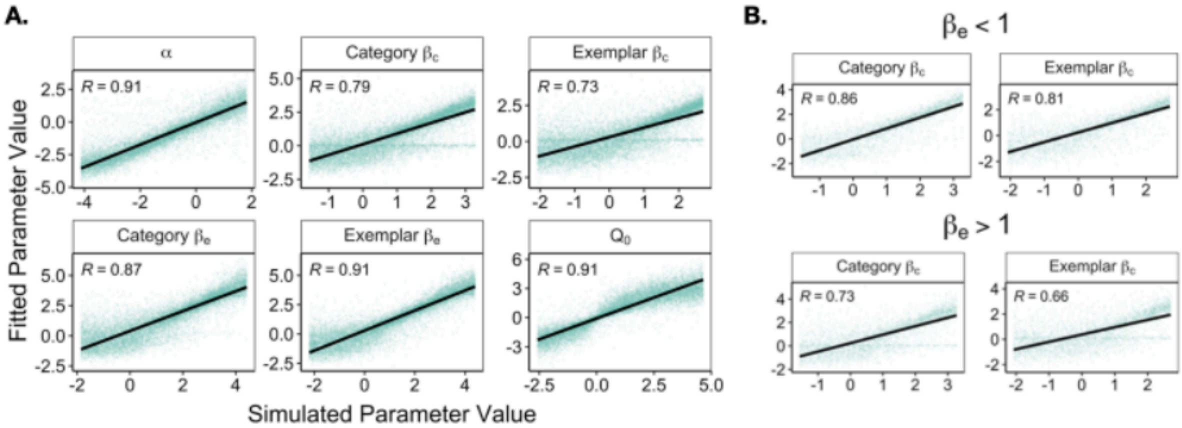


Figure 7. Experiment 1 parameter recovery for the fourB_oneQ model. A) Correlations between simulated and recovered parameter values for the fourB_oneQ model ranged from .73 to .91. Choice data were simulated for 15,100 agents, with parameter values sampled from uniform distributions with minima and maxima determined by the minimum and maximum parameter estimates from the fitted data. B) Exemplar choice weight values influenced category choice weight recoverability, such that category choice weights could be better recovered when exemplar choice weights were lower.

Finally, we also examined the extent to which model simulations recapitulated key aspects of our behavioral results. We describe these posterior predictive checks in the Supplement, but note here that simulations from our winning model can successfully reproduce important signatures of participant task performance.

Experiment 2

Participants

Seventy-three participants completed the two-part online study and were included in all analyses. An additional 13 participants completed both parts of the study but were excluded from all analyses for: (a) interacting with their browser window (minimized, maximized, or clicked outside the window) more than 20 times throughout either the learning or memory task ($n = 8$), (b) failing to respond on more than 10% of learning or memory trials ($n = 1$), or (c) responding in less than 100 ms on more than 10% of learning or memory trials ($n = 4$). In addition, three additional participants were excluded due to a glitch that prevented data from being saved. Participants were recruited, tested, and compensated as in Experiment 1, though base payment was increased to $23 because the learning task was slightly longer.

The 73 participants included in the final sample comprised $n = 34$ children (8.1 - 12.9 years; Mean age = 10.9 years, $n = 34$ females) and $n = 39$ adults (18.4 - 25.8 years, Mean age = 21.9 years, $n = 27$ females). 52% of participants were White, 32% were Asian, 2.7% were Black, 12% were two or more races, and 1.3% were Pacific Islander or Native Hawaiian. In addition, 9.3% of participants were Hispanic (see Supplement for more demographic details).

Experimental procedure

Reinforcement-learning task

The reinforcement-learning task used in Experiment 2 was similar to that used in Experiment 1, but participants had to select from three choice options on every trial (Figure 6C). Each stimulus was associated with an optimal choice that would usually cause the participant to win a point, and two suboptimal choices that would usually cause the participant to lose a point. As in Experiment 1, the task comprised six blocks, each with its own stimulus set. We used the same six stimulus sets as in Experiment 1, though they were modified to include fewer images. Here, the six stimulus sets each included 20 unique images, divided into four broader categories (Figure 4A). Three stimuli from three broader categories were used during each block of the learning task. The two unused stimuli from each of the three broader categories, and all five stimuli from the remaining, fourth category were used as novel exemplar and category images, respectively, in the subsequent test of recognition memory.

As in Experiment 1, the instructions for each task block varied slightly depending on the stimulus set. For example, in the ‘Farm Animals’ block, participants were instructed that they had to return animals to the barns where they lived. Returning animals to the correct barn would usually make them happy, causing participants to win a point, and returning animals to the incorrect barn would usually make them sad, causing participants to lose a point.

Participants completed three category-predictive blocks and three exemplar-predictive blocks, in a pseudorandom order (as in Experiment 1). In category-predictive blocks, the broader stimulus categories determined the stimulus-action reward probabilities, such that all members of a broader category were associated with the same optimal choice (Figure 4B). In exemplar-predictive blocks, one exemplar from each broader category was randomly paired with each of the three choice options, such that *within* a broader category, the optimal choices for all three exemplars differed. Selecting the optimal choice for a stimulus resulted in winning one point on 90% of trials and losing one point on 10% of trials; selecting either of the two other choices resulted in winning one point on 10% of trials and losing one point on 90% of trials. All stimuli repeated 8 times within each block, in a random order, such that each block comprised 72 trials. Block order was pseudorandomized for each participant as in Experiment 1.

All trials lasted 4 seconds, regardless of how quickly participants made their response. Within the 4-second time limit, participants made their choice selection by pressing 1, 2, or 3 on a standard keyboard. After making their selection, participants saw their choice highlighted for 500 ms, and then the outcome of their choice for the remainder of the trial (Figure 4C). Participants saw green checkmarks with a bouncing animation and "+1" if they won a point, and red X's with a swinging animation and "-1" if they lost a point. Trials were separated by a 500 ms inter-trial interval in which no stimuli appeared on the screen. The positions of the choice images were randomized on every trial.

As in Experiment 1, participants completed an extensive, child-friendly tutorial prior to beginning the reinforcement-learning task. After the tutorial, participants answered three True/False comprehension questions about the task. If they answered a question incorrectly, they would see (and hear) the correct answer with an explanation. On average, participants answered 2.7 comprehension questions correctly (Age group means: Children: 2.68; Adults: 2.71). There was no significant effect of age group on the number of questions answered correctly ($b = .04$, $SE = .11$, $p = .71$).

Memory test

The memory test was identical to that used in Experiment 1. Participants completed the memory test between 6 and 9 days after they completed the reinforcement-learning task (except for one adult who completed the memory test on day 5, and one child who completed it on day 10; Mean delay = 7.1 days; SD = 1.1 days; See Supplement for an analysis of the effects of delay duration on memory).

Analysis approach

Our analysis approach largely aligned with our approach in Experiment 1. Here, however, because we did not collect data from adolescents, rather than analyzing age as a continuous variable, we treated age group (children and adults) as a categorical variable.

Modeling approach. We fit the same computational models to our data, with several modifications to take into account the different reward structure of the task. First, rather than tracking three category values and 15 exemplar values, here, the models track 9 (3 categories x

3 choices) category-action values and 27 (9 exemplars x 3 choices) exemplar-action values. We re-coded the binary rewards as 0 and 1, and constrained initial Q values to fall in this range. In addition, in Experiment 1, counterfactual feedback was presented explicitly to participants on 'avoid' trials in which they did not experience gains or losses. Here, no counterfactual feedback was presented. Instead, our counterfactual learning models assumed that participants might infer that the unselected choice options would have yielded gains when the selected choice option yielded a loss and vice versa.

We followed the same model-fitting and stage-wise selection approach as in Experiment 1, and found that the best-fitting model also included four choice weights and one initial Q value. Here, rather than exhibiting equivalent learning from experienced and counterfactual outcomes, we found that participants' choices were best fit by a model (fourB_oneQ_CF) that included separate learning rates for experienced outcomes from selected options and inferred outcomes froml unselected choice options. Models (Figure 8) and parameter values from the winning model (Figure 9) both showed good recoverability, with correlations between simulated and fitted parameters ranging from .65 to .91. In addition, model simulations recapitulated key features of participant choice behavior (see posterior predictive checks in the Supplement).

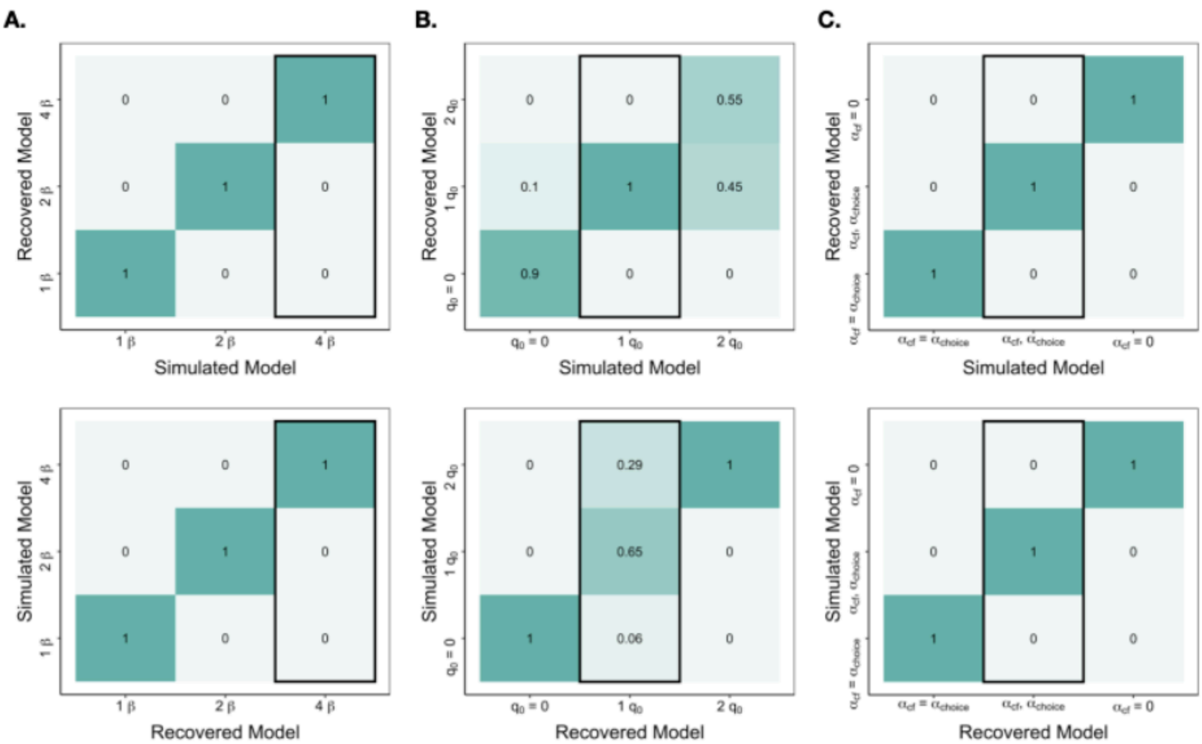


Figure 8. Experiment 2 model recoverability. For each model within each stage of model comparison, 100 simulated 'experiments' were conducted in which choice data was simulated from 73 agents, with parameters sampled from uniform distributions with ranges determined by the empirical fits. Data from each simulated experiment were then fit with each model within the comparison set. The top panels show confusion matrices, where the values within each tile represent the proportion of experiments for which each fitted model had the highest exceedance

probability (top panels). The bottom panels show inversion matrices, where the values within each tile represent the proportion of experiments for which the fitted model had the highest exceedance probability that were generated by each of the models. Black lines outline the model that best-fit the empirical data within each comparison stage. A) Models with different numbers of choice weights were highly distinguishable from one another. B) Models in which exemplar and category values were initialized with either one or two free parameters were moderately distinguishable from one another. C) Models with a single learning rate, separate learning rates for experienced and inferred counterfactual outcomes, and no counterfactual learning were highly distinguishable from one another.

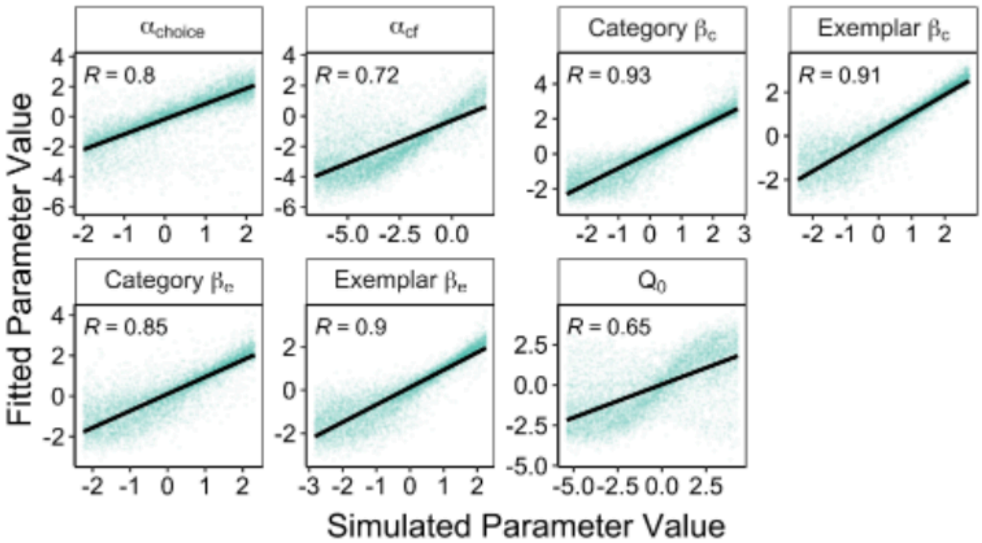


Figure 9. Experiment 2 parameter recoverability for the fourB_oneQ model. Correlations between simulated and recovered parameter values for the fourB_oneQ model ranged from .65 to .91. Choice data were simulated for 10,000 agents, with parameter values sampled from uniform distributions with minima and maxima determined by the minimum and maximum parameter estimates from the fitted data.

Data and code availability

All task code, anonymized data, and analysis code are available on the Open Science Framework: https://osf.io/8yjvr/

Acknowledgments

We thank Naiti Bhatt and Juan Velez for help with task design and stimulus preparation, Todd Gureckis for helpful discussions and feedback on an early draft, and Susan Benear and Nora Harhen for helpful feedback on the manuscript. This work was supported by the National Institute of Mental Health (R01 MH126183 to C.A.H., F31 MH129105 to K.N.), the American Psychological Association (Dissertation Research Award to K.N.), and the CV Starr Foundation Fellowship (to K.N.). This work was supported in part through the NYU IT High Performance Computing resources, services, and staff expertise.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). https://doi.org/10.1016/j.jml.2012.11.001

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

Blanco, N. J., & Sloutsky, V. M. (2019). Adaptive flexibility in category learning? Young children exhibit smaller costs of selective attention than adults. *Developmental Psychology*, *55*(10), 2060–2076. https://doi.org/10.1037/dev0000777

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*. https://doi.org/10.3758/s13423-022-02179-w

Calderon, C. B., De Loof, E., Ergo, K., Snoeck, A., Boehler, C. N., & Verguts, T. (2021). Signed Reward Prediction Errors in the Ventral Striatum Drive Episodic Memory. *Journal of Neuroscience*, *41*(8), 1716–1726. https://doi.org/10.1523/JNEUROSCI.1785-20.2020

Callaghan, B., Gasser, C., Silvers, J., VanTieghem, M., Choy, T., O'Sullivan, K., Tompary, A., Davachi, L., & Tottenham, N. (2021). Age-related increases in posterior hippocampal granularity are associated with remote detailed episodic memory in development. *Journal of Neuroscience*, *41*(8), 1738–1754. https://doi.org/10.1523/JNEUROSCI.1738-20.2020

Cohen, A. O., Glover, M. M., Shen, X., Phaneuf, C. V., Avallone, K. N., Davachi, L., & Hartley, C. A. (2022). Reward enhances memory via age-varying online and offline neural mechanisms across development. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.1820-21.2022

Cohen, A. O., Matese, N. G., Filimontseva, A., Shen, X., Shi, T. C., Livne, E., & Hartley, C. A. (2019). Aversive learning strengthens episodic memory in both adolescents and adults. *Learning & Memory*, *26*(7), 272–279. https://doi.org/10.1101/lm.048413.118

Davidow, J. Y., Foerde, K., Galván, A., & Shohamy, D. (2016). An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. *Neuron*, *92*(1), 93–99. https://doi.org/10.1016/j.neuron.2016.08.031

Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective & Behavioral Neuroscience*, *15*(2), 310–320. https://doi.org/10.3758/s13415-014-0332-5

de Leeuw, J. R. (2015). jsPsych: a JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Deng, W. S., & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62. https://doi.org/10.1016/j.cogpsych.2016.09.002

Dudai, Y. (2012). The Restless Engram: Consolidations Never End. *Annual Review of Neuroscience*, *35*(1), 227–247. https://doi.org/10.1146/annurev-neuro-062111-150500

Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: how humans generalize fear. *Trends in Cognitive Sciences*, *19*(2), 73–77. https://doi.org/10.1016/j.tics.2014.12.003

Eckstein, M. K., & Collins, A. G. E. (2020). Computational evidence for hierarchically structured reinforcement learning in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(47), 29381–29389.

Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). The interpretation of computational model parameters depends on the context. *eLife*, *11*(e75474). https://doi.org/10.7554/eLife.75474

Fivush, R., Hudson, J., & Nelson, K. (1984). Children's Long-Term Memory for a Novel Event: An Exploratory Study. *Merrill-Palmer Quarterly*, *30*(3), 303–316. http://www.jstor.org/stable/23086104

Frank, M. J., & Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cerebral Cortex* , *22*(3), 509–526. https://doi.org/10.1093/cercor/bhr114

Frank, S. M., Bründl, S., Frank, U. I., Sasaki, Y., Greenlee, M. W., & Watanabe, T. (2021). Fundamental Differences in Visual Perceptual Learning between Children and Adults. *Current Biology: CB*, *31*(2), 427–432.e5. https://doi.org/10.1016/j.cub.2020.10.047

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

Ghetti, S., & Fandakova, Y. (2020). Neural Development of Memory and Metamemory in Childhood and Adolescence: Toward an Integrative Model of the Development of Episodic Recollection. *Annual Review of Developmental Psychology*, *2*(1), 365–388.

Giron, A. P., Ciranka, S. K., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in learning resemble stochastic optimization. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-023-01662-1

Glenn, C. R., Klein, D. N., Lissek, S., Britton, J. C., Pine, D. S., & Hajcak, G. (2012). The development of fear learning and generalization in 8--13 year-olds. *Developmental Psychobiology*, *54*(7), 675–684. https://doi.org/10.1002/dev.20616

Hartley, C. A., Nussenbaum, K., & Cohen, A. O. (2021). Interactive Development of Adaptive Learning and Memory. *Annual Review of Developmental Psychology*, *3*, 59–85. https://doi.org/10.1146/annurev-devpsych-050620-030227

Jang, A. I., Nassar, M. R., Dillon, D. G., & Frank, M. J. (2019). Positive reward prediction errors during decision-making strengthen memory encoding. *Nature Human Behaviour*, *3*(7), 719–732. https://doi.org/10.1038/s41562-019-0597-3

Kalbe, F., & Schwabe, L. (2020). Beyond arousal: Prediction error related to aversive events promotes episodic memory formation. *Journal of Experimental Psychology. Learning,*

Memory, and Cognition, *46*(2), 234–246. https://doi.org/10.1037/xlm0000728
Keresztes, A., Bender, A. R., Bodammer, N. C., Lindenberger, U., Shing, Y. L., & Werkle-Bergner, M. (2017a). Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(34), 9212–9217. https://doi.org/10.1073/pnas.1710654114
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.
Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: parallel brain systems for item memory and category knowledge. *Science*, *262*(5140), 1747–1749. https://doi.org/10.1126/science.8259522
Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44. https://doi.org/10.1037/0033-295x.99.1.22
Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends in Cognitive Sciences*, *20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004
Lambert, F. R., Lavenex, P., & Lavenex, P. B. (2015). Improvement of allocentric spatial memory resolution in children from 2 to 4 years of age. *International Journal of Behavioral Development*, *39*(4), 318–331. https://doi.org/10.1177/0165025415584808
Lengyel, M., & Dayan, P. (2007). Hippocampal contributions to control: The third way. *Advances in Neural Information Processing Systems*, *20*.
Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. *Neuron*, *93*(2), 451–463. https://doi.org/10.1016/j.neuron.2016.12.040
Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, *218*, 104940.
Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x
Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D., & Dolan, R. J. (2021). Experience replay is associated with efficient nonlocal learning. *Science*, *372*, eabf1357. https://doi.org/10.1126/science.abf1357
Love, B. C. (2005). Environment and Goals Jointly Direct Category Acquisition. *Current Directions in Psychological Science*, *14*(4), 195–199. https://doi.org/10.1111/j.0963-7214.2005.00363.x
Love, B. C., & Gureckis, T. M. (2007). Models in search of a brain. *Cognitive, Affective & Behavioral Neuroscience*, *7*(2), 90–108. https://doi.org/10.3758/cabn.7.2.90
Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: a network model of category learning. *Psychological Review*, *111*(2), 309–332.

https://doi.org/10.1037/0033-295X.111.2.309
Mack, M. L., Preston, A. R., & Love, B. C. (2020). Ventromedial prefrontal cortex compression during concept learning. *Nature Communications*, *11*(1), 46. https://doi.org/10.1038/s41467-019-13930-8
Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*(11), 1609–1617. https://doi.org/10.1038/s41593-018-0232-z
McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). 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*(3), 419. https://doi.org/10.1037/0033-295X.102.3.419
Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238. https://doi.org/10.1037/0033-295X.85.3.207
Mednick, S. A., & Lehtinen, L. E. (1957). Stimulus generalization as a function of age in children. *Journal of Experimental Psychology*, *53*(3), 180–183. https://doi.org/10.1037/h0047497
Murayama, K., & Kitagami, S. (2014). Consolidation power of extrinsic rewards: reward cues enhance long-term memory for irrelevant past events. *Journal of Experimental Psychology. General*, *143*(1), 15–20. https://doi.org/10.1037/a0031992
Murty, V. P., Tompary, A., Adcock, R. A., & Davachi, L. (2017). Selectivity in postencoding connectivity with High-Level visual cortex is associated with reward-motivated memory. *Journal of Neuroscience*, *37*(3), 537–545. https://doi.org/10.1523/jneurosci.4032-15.2017
Nelson, K., & Gruendel, J. (1981). Generalized event representations: Basic building blocks of cognitive development. *Advances in Developmental Psychology*. https://doi.org/10.4324/9780203728185-4/generalized-event-representations-katherine-nelson-janice-gruendel
Ngo, C. T., Benear, S. L., Popal, H., Olson, I. R., & Newcombe, N. S. (2021). Contingency of semantic generalization on episodic specificity varies across development. *Current Biology: CB*, *31*(12), 2690–2697.e5.
Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2018). The ontogeny of relational memory and pattern separation. *Developmental Science*, *21*(2). https://doi.org/10.1111/desc.12556
Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2019). Gain-loss framing enhances mnemonic discrimination in preschoolers. *Child Development*, *90*(5), 1569–1578.
Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(21), 8145–8157. https://doi.org/10.1523/JNEUROSCI.2978-14.2015
Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology. General*, *115*(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39
Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What

insights can we draw from a decade of research? *Developmental Cognitive Neuroscience*, *40*, 100733. https://doi.org/10.1016/j.dcn.2019.100733
Nussenbaum, K., & Hartley, C. A. (2021). Developmental change in prefrontal cortex recruitment supports the emergence of value-guided memory. *Elife*, *10*.
Nussenbaum, K., Prentis, E., & Hartley, C. A. (2020). Memory's reflection of learned information value increases across development. *Journal of Experimental Psychology. General*, *149*(10), 1919–1934.
Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra: Psychology*, *6*(1).
Nussenbaum, K., Velez, J. A., Washington, B. T., Hamling, H. E., & Hartley, C. A. (2022). Flexibility in valenced reinforcement learning computations across development. *Child Development*, *93*(5), 1601–1615. https://doi.org/10.1111/cdev.13791
O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science*, *38*(6), 1229–1248. https://doi.org/10.1111/j.1551-6709.2011.01214.x
Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning & Memory* , *24*(1), 65–69. https://doi.org/10.1101/lm.042978.116
Piray, P., Dezfouli, A., Heskes, T., Frank, M. J., & Daw, N. D. (2019). Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Computational Biology*, *15*(6), e1007043. https://doi.org/10.1371/journal.pcbi.1007043
Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of Selective Attention: When Children Notice What Adults Miss. *Psychological Science*, *28*(6), 723–732. https://doi.org/10.1177/0956797617693005
Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: a life-span overview. *Acta Psychologica*, *86*(2-3), 227–272. https://doi.org/10.1016/0001-6918(94)90004-3
Price, D. W., & Goodman, G. S. (1990). Visiting the wizard: children's memory for a recurring event. *Child Development*, *61*(3), 664–680. https://www.ncbi.nlm.nih.gov/pubmed/2364742
Radulescu, A., Niv, Y., & Ballard, I. (2019). Holistic Reinforcement Learning: The Role of Structure and Attention. *Trends in Cognitive Sciences*, *23*(4), 278–292. https://doi.org/10.1016/j.tics.2019.01.010
Ramsaran, A. I., Schlichting, M. L., & Frankland, P. W. (2019). The ontogeny of memory persistence and specificity. *Developmental Cognitive Neuroscience*, *36*, 100591.
Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making*, *7*(3), 332–359.
Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature*

Neuroscience, *17*(7), 981–986. https://doi.org/10.1038/nn.3736
Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies - revisited. *NeuroImage*, *84*, 971–985. https://doi.org/10.1016/j.neuroimage.2013.08.065
Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. https://doi.org/10.1186/1471-2105-12-77
Rollins, L., & Cloude, E. B. (2018). Development of mnemonic discrimination during childhood. *Learning & Memory* , *25*(6), 294–297. https://doi.org/10.1101/lm.047142.117
Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology. General*, *104*(3), 192–233. https://doi.org/10.1037/0096-3445.104.3.192
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439.
Rosenbaum, G. M., Grassie, H. L., & Hartley, C. A. (2022). Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. *eLife*, *11*. https://doi.org/10.7554/eLife.64620
Rouhani, N., & Niv, Y. (2019). Depressive symptoms bias the prediction-error enhancement of memory towards negative events in reinforcement learning. *Psychopharmacology*, *236*(8), 2425–2435. https://doi.org/10.1007/s00213-019-05322-z
Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife*, *10*. https://doi.org/10.7554/eLife.61077
Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *44*(9), 1430–1443.
Santoro, A., Frankland, P. W., & Richards, B. A. (2016). Memory transformation enhances reinforcement learning in dynamic environments. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(48), 12228–12242.
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. https://doi.org/10.1002/dev.21393
Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for Rewards Like a Child Means Less Generalization and More Directed Exploration. *Psychological Science*, *30*(11), 1561–1572. https://doi.org/10.1177/0956797619863663
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–1323. https://doi.org/10.1126/science.3629243
Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *Afex: analysis of factorial experiments* (Version .27-2). https://CRAN.R-project.org/package=afex
Sloutsky, V. M., & Fisher, A. V. (2004). When development and learning decrease memory. Evidence against category-based induction in children. *Psychological Science*, *15*(8),

553–558. https://doi.org/10.1111/j.0956-7976.2004.00718.x

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory Consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Starita, F., Kroes, M. C. W., Davachi, L., Phelps, E. A., & Dunsmoor, J. E. (2019). Threat learning promotes generalization of episodic memory. *Journal of Experimental Psychology. General*, *148*(8), 1426–1434.

Sterpenich, V., van Schie, M. K. M., Catsiyannis, M., Ramyead, A., Perrig, S., Yang, H.-D., Van De Ville, D., & Schwartz, S. (2021). Reward biases spontaneous neural reactivation during sleep. *Nature Communications*, *12*(1), 4162. https://doi.org/10.1038/s41467-021-24357-5

Sutton, R. S., Barto, A. G., Co-Director Autonomous Learning Laboratory Andrew G Barto, & Bach, F. (1998). *Reinforcement Learning: An Introduction*. MIT Press. https://play.google.com/store/books/details?id=CAFR6IBF4xYC

Tandoc, M. C., Nadendla, B., Pham, T., & Finn, A. S. (2022). Dividing attention hurts learning in adults but not children. In *PsyArXiv*. https://osf.io/preprints/psyarxiv/aqw28/

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *The Behavioral and Brain Sciences*, *24*(4), 629–640; discussion 652–791. https://doi.org/10.1017/s0140525x01000061

The MathWorks Inc. (2020). *MATLAB version 9.9.0.1 (R2020b)*. The Mathworks Inc. https://www.mathworks.com

Tompary, A., Zhou, W., & Davachi, L. (2020). Schematic memories develop quickly, but are not expressed unless necessary. *Scientific Reports*, *10*(1), 16968.

Wharton-Shukster, E., & Finn, A. S. (2019). A Trade-Off in Learning Across Levels of Abstraction in Adults and Children. *Proceedings of the 41st Annual Conference of the Cognitive Science Society*.

Wilson, R. C., & Collins, A. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, 8:e49547. https://doi.org/10.7554/eLife.49547

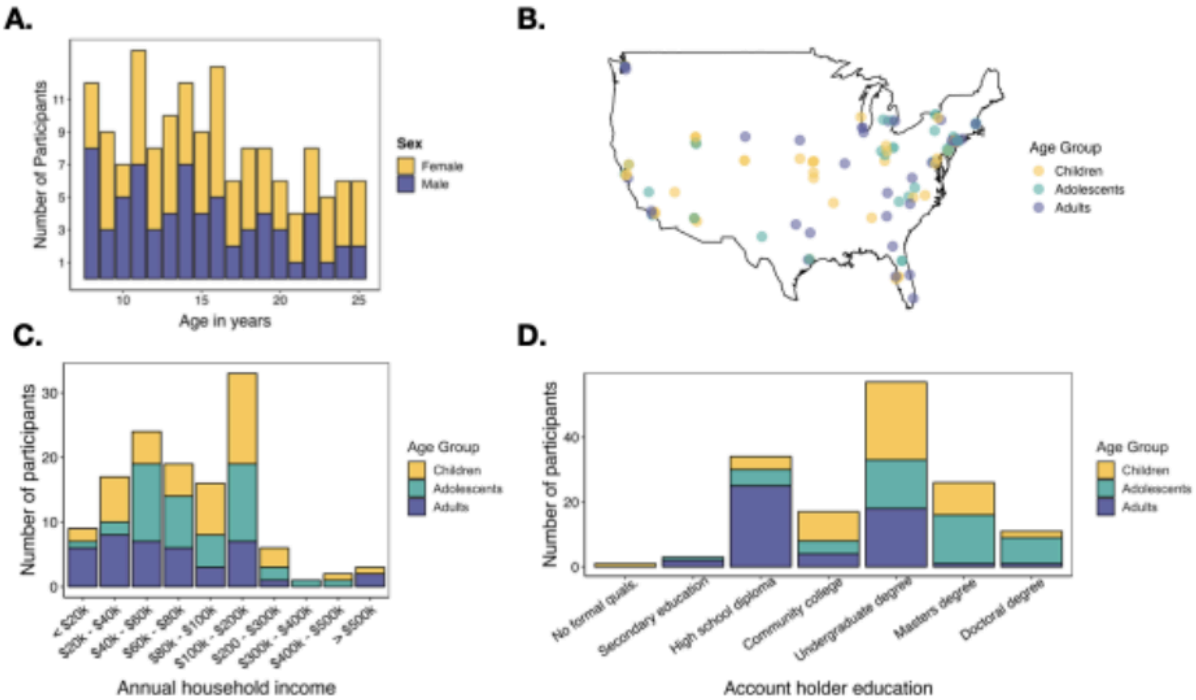
Wittmann, B. C., Dolan, R. J., & Düzel, E. (2011). Behavioral specifications of reward-associated long-term memory enhancement in humans. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(5), 296–300.

Zaki, S. R., & Nosofsky, R. M. (2001). A single-system interpretation of dissociations between recognition and categorization in a task involving object-like stimuli. *Cognitive, Affective & Behavioral Neuroscience*, *1*(4), 344–359. https://doi.org/10.3758/cabn.1.4.344

Supplementary Information

A. Supplementary information for Experiment 1

A1. Participant demographics



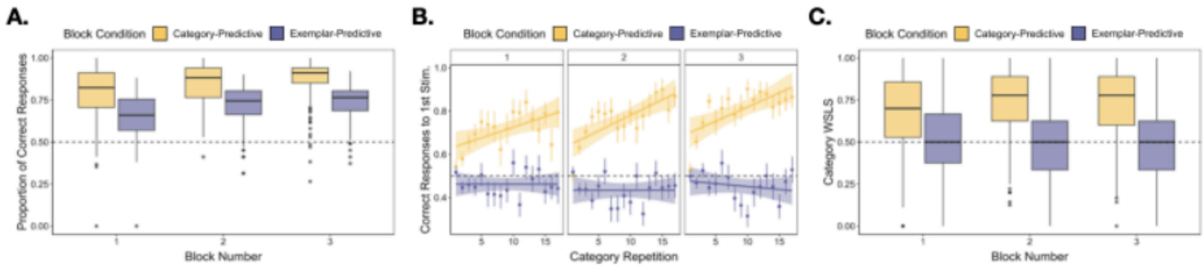
Supplementary Figure 1. Experiment 1 demographics. Distributions of participant A) age and sex, B) zip codes, C) annual household incomes, and D) (parental) levels of educational attainment. Educational attainment levels reflect those of adult participants or the parent/guardian who registered their child to participate in the study.

A2. Effects of block number on learning

As reported in the main text, participants demonstrated improvements in learning across blocks, indicating that they may have learned general task strategies to accelerate learning of new reward contingencies for new stimuli. Participants made more correct responses to stimuli in later task blocks within each condition, OR = 1.27 (.04), $\chi^2(1) = 61.0$, $p < .001$, particularly in exemplar-predictive blocks (block number x block condition interaction, OR = 1.07 (.02), $\chi^2(1) = 8.5$, $p = .004$; Supplementary Figure 2A). Participants also demonstrated faster learning in later blocks (block number x trial interaction, OR = 1.08 (.02), $\chi^2(1) = 19.2$, $p < .001$). The effects of block number on correct responses did not vary across age (*p*s > .35).

Participants also demonstrated stronger generalization of learned category values in later blocks of the task, such that they increasingly responded correctly to novel stimuli in category-predictive task blocks (main effect of block number: OR = 1.10 (.02), $\chi^2(1) = 15.8$, $p < .001$; block number x block condition interaction: OR = 1.10 (.02), $\chi^2(1) = 16.0$, $p < .001$; block number x block condition x category repetition interaction: OR = 1.7 (.03), $\chi^2(1) = 9.0$, $p = .003$; Supplementary Figure 2B). As with correct responses, the effects of block number on generalization similarly did not vary with age (*p*s > .15).

Category 'win-stay lose-shift' behavior also increased across the experiment. Participants demonstrated more WSLS behavior in later task blocks, particularly in category predictive blocks (main effect of block number: OR = 1.18 (.02), $\chi^2(1) = 59.2$, $p < .001$; block number x block condition interaction: OR = 1.17 (.02), $\chi^2(1) = 99.8$, $p < .001$; Supplementary Figure 2C). The effects of block number on WSLS behavior did not vary with age (*p*s > .16).



Supplementary Figure 2. Experiment 1 learning improvements across blocks. A) Participants made more correct responses during learning as a function of increasing, within-condition block number. B) They also demonstrated stronger generalization and C) stronger category win-stay lose-shift behavior in the category-predictive condition as the experiment progressed. In panels A and C, the boxplot hinges show the first and third quartiles of the data. whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians. In panel B, points depict within-condition block number means and error bars show standard errors. The lines show the best-fitting linear regressions through the points, with the shaded region depicting 95% confidence intervals.

A3. Full results from mixed-effects models examining relations between learning and memory

Supplementary Table 1: Effects of age, exemplar choice weight magnitude, specificity, block condition, and their interactions on memory performance

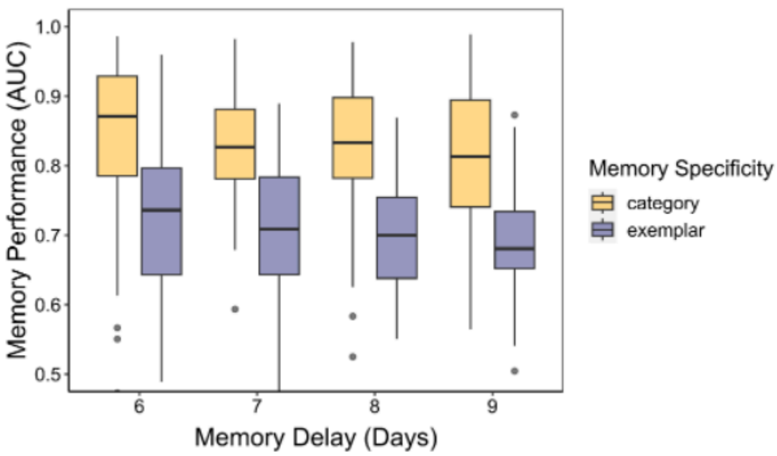
effect	**β**	**SE**	**df**	***F***	***p***
age	.0097	.0068	1, 152.46	2.06	.154
exemplar choice weight	**.0180**	**.0047**	**1, 577.16**	**14.59**	**<.001**
specificity	**.0610**	**.0025**	**1, 440.01**	**580.97**	**<.001**
block condition	**-.0052**	**.0026**	**1, 454.64**	**4.10**	**.043**
age x exemplar choice weight	**.0159**	**.0055**	**1, 534.66**	**8.34**	**.004**
age x specificity	.0038	.0025	1, 440.01	2.21	.138
exemplar choice weight x specificity	-.0044	.0027	1, 440.01	2.64	.105
age x block condition	-.0003	.0026	1, 456.20	0.01	.913
exemplar choice weight x block condition	**-.0097**	**.0031**	**1, 497.08**	**9.57**	**.002**
specificity x block condition	-.0001	.0025	1, 440.01	0.00	.958
age x exemplar choice weight	-.0041	.0030	1, 440.01	1.88	.171
age x exemplar choice weight x block condition	**-.0075**	**.0033**	**1, 500.15**	**5.05**	**.025**
age x specificity x block condition	.0011	.0025	1, 440.01	0.18	.671
exemplar choice weight x specificity x block condition	-.0006	.0027	1, 440.01	0.05	.830
age x exemplar choice weight x specificity x block condition	.0044	.0030	1, 440.01	2.19	.140

Supplementary Table 2: Effects of age, category choice weight magnitude, specificity, block condition, and their interactions on memory performance

effect	**β**	**SE**	**df**	***F***	***p***
age	**.0196**	**.0068**	**1, 172.10**	**7.41**	**.007**
category choice weight	.0000	.0043	1, 575.65	0.00	.991
specificity	**.0593**	**.0029**	**1, 443.07**	**424.1**	**.001**
block condition	**-.0076**	**.0034**	**1, 520.44**	**5.12**	**.024**
age x category choice weight	-.0017	.0042	1, 573.22	0.16	.687
age x specificity	.0025	.0029	1, 443.07	0.71	.400
category choice weight x specificity	**.0070**	**.0029**	**1, 443.07**	**5.93**	**.015**
age x block condition	-.0029	.0034	1, 519.75	0.70	.402
category choice weight x block condition	.0001	.0036	1, 510.64	0.00	.972
specificity x block condition	-.0019	.0029	1, 443.07	0.43	.511
age x category choice weight	-.0012	.0029	1, 443.07	0.17	.683
age x category choice weight x block condition	-.0070	.0036	1, 513.90	3.77	.053
age x specificity x block condition	.0014	.0029	1, 443.07	0.24	.627
category choice weight x specificity x block condition	.0022	.0029	1, 443.07	0.58	.448
age x category choice weight x specificity x block condition	-.0023	.0029	1, 443.07	0.63	.428

A4. Effects of delay on memory performance

Participants were invited to complete the memory test six days after participating in the reinforcement-learning task, and had until 10 days after the learning task to complete it. Participants who had not completed the memory test by Day 8 received reminder emails on Day 8 and Day 9. To analyze the influence of the delay between learning and memory on memory performance, we re-ran our analysis examining the effects of age group, block condition and level of abstraction on memory, with delay (in days) as an interacting fixed effect. We continued to observe main effects of block condition and level of abstraction (*p*s < .001). Here, we also observed a main effect of delay on memory performance, such that memory was worse at longer delays (β = -.015 (.007), $F(1, 147) = 4.71$, $p = .032$; Supplementary Figure 3), as well as a delay x age interaction, indicating the stronger effects of memory delay for younger participants (β = -.014 (.007), $F(1, 147) = 3.99$, $p = .048$).

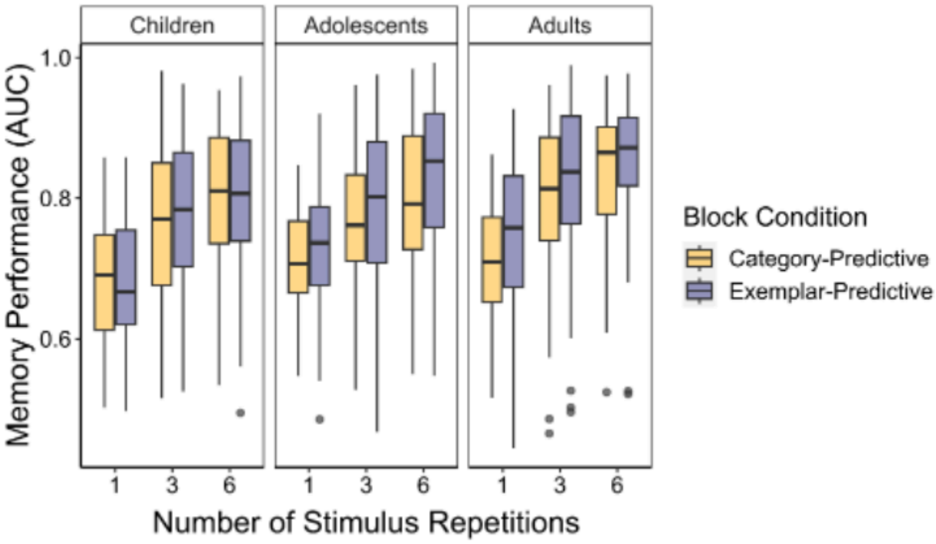


Supplementary Figure 3. Experiment 1 memory across levels of abstraction at different delays. Participants who completed the memory test at longer delays demonstrated worse memory for exemplar-level information than those who completed it at shorter delays. The boxplot hinges show the first and third quartiles of the data. whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians.

Because all participants were invited to complete the memory test after 6 days, there may be systematic differences between the participants who completed the test after a shorter versus longer delay. For example, participants who did not complete the memory test until Day 9 (after multiple reminder emails) may not have been as engaged in the online study as those who participated immediately after being invited on Day 6. Future work can isolate the influence of delay time on memory by systematically manipulating the delay at which participants are *invited* to complete the memory test.

A5. Effects of stimulus repetitions on memory

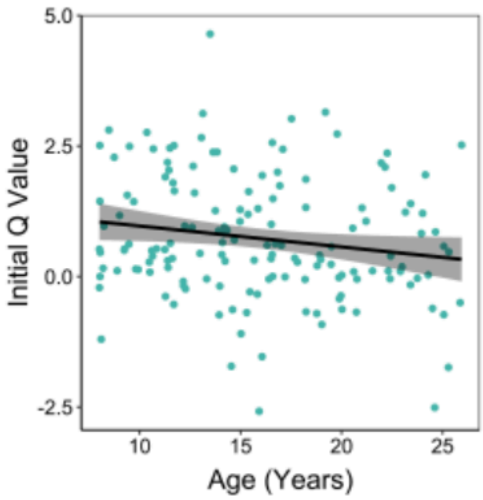
In the reinforcement-learning task, stimuli within each block repeated six times, three times, or once. In the memory analyses in the main text of the manuscript, we constructed ROC curves and computed AUC collapsing across all numbers of stimulus repetitions. However, we also analyzed how number of stimulus repetitions influenced memory performance by collapsing across levels of abstraction, constructing ROC curves and computing AUC for each number of stimulus repetitions within each block condition. We then examined how number of stimulus repetitions, block condition, and age influenced memory. As in our prior analyses, we continued to observe better memory performance in exemplar-predictive relative to category-predictive blocks, $\beta = -.01$ (.00), $F(1, 755.2) = 16.6$, $p < .001$, and better memory performance at older ages, $\beta = .01$ (.01), $F(1, 149.2) = 4.3$, $p = .039$ (Supplementary Figure 4). Here, we also observed a main effect of stimulus repetition number, such that participants demonstrated better memory for stimuli they saw more times during learning, $\beta = .04$ (.01), $F(1, 755.2) = 372.5$, $p = < .001$. No other effects or interactions were significant (*p*s > .059).



Supplementary Figure 4. Experiment 1 memory across stimulus repetitions. Participants demonstrated better memory for stimuli that repeated more often during learning. The boxplot hinges show the first and third quartiles of the data. whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians.

A6. Relations between age and model parameters

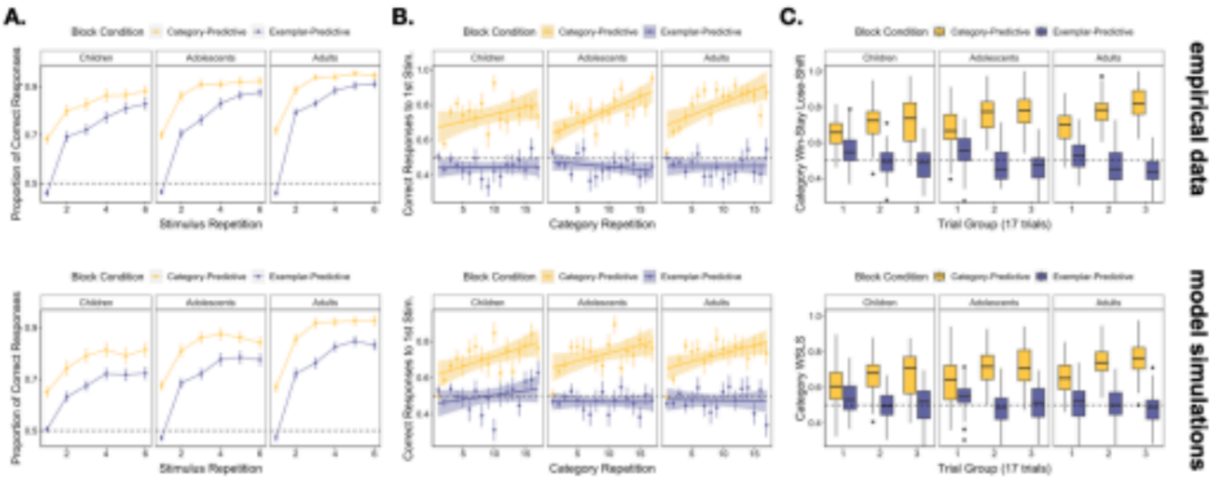
Beyond choice weights, we also examined how the other model-derived parameters varied across age. We did not observe evidence for age-related change in learning rates (b = .008 (.079), $t(149) = .10$, $p = .921$). We did observe a significant relation between age and initial q values (b = -.20 (.095), $t(149) = -2.11$, $p = .036$; Supplementary Figure 5). The majority of participants in our sample had positive initial q values, suggesting that they had an initial bias to approach novel stimuli; younger participants had higher initialization values, suggesting that they may have been more optimistic about the value of novel options.



Supplementary Figure 5. Experiment 1 initial q values by age. Participants across age tended to optimistically initialize the value of novel stimuli in the learning task. Initial q values decreased with age. Points represent individuals' model-derived initial q values. The line shows the best-fitting linear regression through the points, with the shaded region showing the 95% confidence interval.

A7. Posterior predictive checks

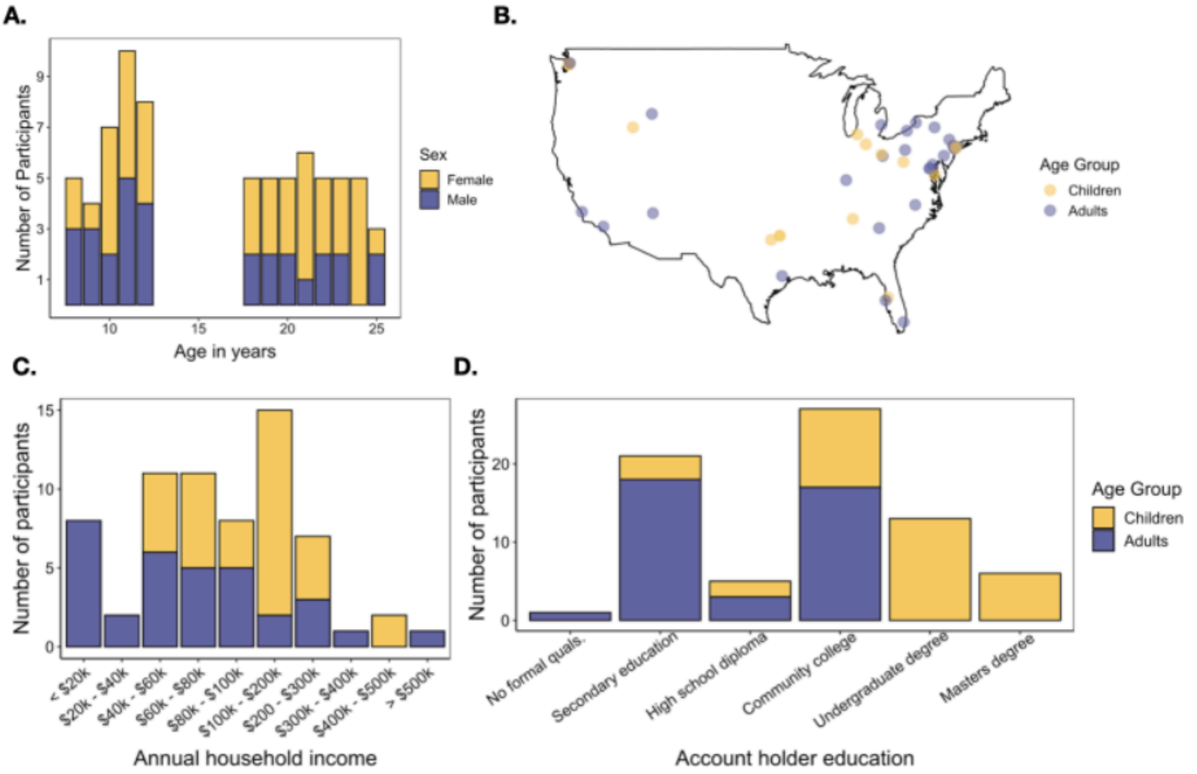
To examine the extent to which our winning reinforcement-learning model (fourB_oneQ) could recapitulate patterns of participant behavior on the task, we simulated data from 100 'versions' of each participant, using their best-fitting parameters and trial sequence. We then analyzed this dataset of 14,800 simulated participants in the same way that we analyzed our empirical data, and found that our model reproduced all key qualitative signatures of participants' learning behavior (Supplementary Figure 6).



Supplementary Figure 6. Experiment 1 posterior predictive checks. Data generated from the best-fitting model (fourB_oneQ) using participants' best-fitting parameters (bottom panels) captured many of the qualitative features of participants' real choice data (top panels). Data was first averaged over the 100 simulations run with each participant's trial sequence and parameter values (to derive average values for a single, simulated agent). A) Over the course of each block, both the empirical data and model simulations show that children and adults learned to make more optimal responses to stimuli in both block conditions, though performance was better in category-predictive relative to exemplar-predictive blocks. Points depict age group means and error bars show standard errors of participant (or agent) means. B) Both the empirical data and model simulations show that in the category-predictive condition, children and adults increasingly generalized learned category responses to respond optimally to novel stimuli. Points depict age group means and error bars show standard errors of participant means. The lines show the best-fitting linear regressions through the points, with the shaded region depicting 95% confidence intervals. C) Both the empirical data and model simulations show that category "win-stay lose-shift" behavior increased across trials in category-predictive blocks and decreased across trials in exemplar-predictive blocks, increasingly so with age. The boxplot hinges show the first and third quartiles of the data averaged within each trial group; whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians.

B. Supplementary information for Experiment 2

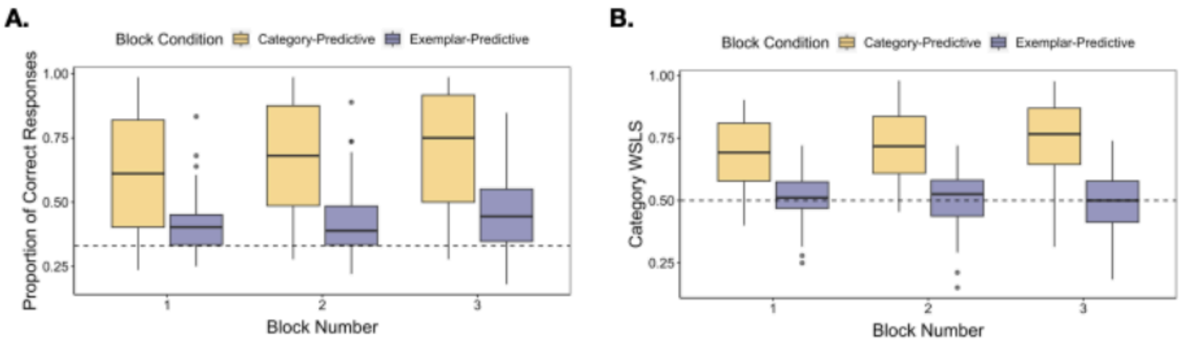
B1. Participant demographics



Supplementary Figure 7. Experiment 2 participant demographics. Distributions of participant A) age and sex, B) zip codes, C) annual household incomes, and D) (parental) levels of educational attainment. Educational attainment levels reflect those of adult participants or the parent/guardian who registered their child to participate in the study.

B2. Effects of block number on learning

As in Experiment 1, participants demonstrated improvements in learning across blocks, indicating that they may have learned general task strategies to accelerate learning of new reward contingencies for new stimuli. Participants made more correct responses to stimuli in later task blocks within each condition, OR = 1.15 (.04), $\chi^2(1) = 15.2$, $p < .001$; Supplementary Figure 8A). The effect of block number on correct responses did not vary across age group (p = .593). Category 'win-stay lose-shift' behavior also increased across the experiment. Participants demonstrated more WSLS behavior in later task blocks, particularly in category predictive blocks (main effect of block number: OR = 1.06 (.02), $\chi^2(1) = 7.69$, $p = .006$; block number x block condition interaction: OR = 1.07 (.02), $\chi^2(1) = 13.0$, $p < .001$; Supplementary Figure 8B). The effects of block number on WSLS behavior did not vary across age groups (ps > .30).



Supplementary Figure 8. Experiment 2 learning improvements across blocks. A) Participants made more correct responses during learning as a function of increasing, within-condition block number. B) They also demonstrated stronger category win-stay lose-shift behavior in the category-predictive condition as the experiment progressed. In both panels, the boxplot hinges show the first and third quartiles of the data; whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians.

B3. Full results from mixed-effects models examining relations between learning and memory

Supplementary Table 3: Effects of age, exemplar choice weight magnitude, specificity, block condition, and their interactions on memory performance

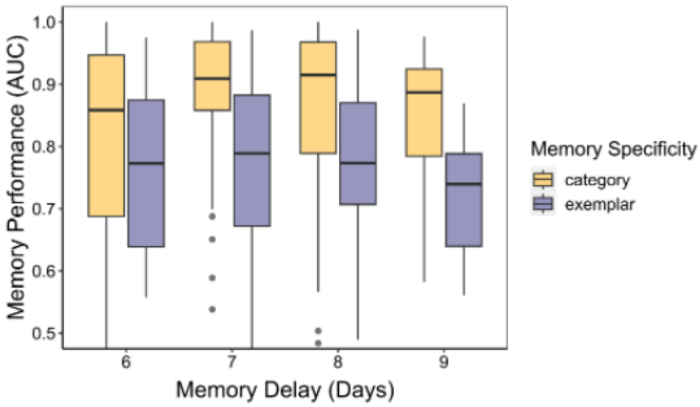
effect	β	SE	df	*F*	*p*
age group	.0027	.0123	1, 63.52	0.05	.824
exemplar choice weight	**.0345**	**.0079**	**1, 265.44**	**19.05**	**<.001**
specificity (exemplar, category)	**.0431**	**.0041**	**1, 197.27**	**108.04**	**<.001**
block condition (exemplar-predictive, category-predictive)	**-.0136**	**.0042**	**1, 201.20**	**10.48**	**.001**
age group x exemplar choice weight	**-.0159**	**.0079**	**1, 265.44**	**4.07**	**.045**
age group x specificity	.0039	.0041	1, 197.27	0.89	.347
exemplar choice weight x specificity	.0015	.0042	1, 197.27	0.13	.722
age group x block condition	-.0035	.0042	1, 201.20	0.70	.405
exemplar choice weight x block condition	-.0018	.0046	1, 209.60	0.16	.689
specificity x block condition	.0079	.0041	1, 197.27	3.66	.057
age group x exemplar choice weight	.0020	.0042	1, 197.27	0.24	.628
age group x exemplar choice weight x block condition	.0061	.0046	1, 209.60	1.81	.181
age group x specificity x block condition	.0034	.0041	1, 197.27	0.66	.418
exemplar choice weight x specificity x block condition	.0010	.0042	1, 197.27	0.05	.819
age group x exemplar choice weight x specificity x block condition	.0008	.0042	1, 197.27	0.04	.842

Supplementary Table 4: Effects of age group, category choice weight magnitude, specificity, block condition, and their interactions on memory performance

effect	**β**	**SE**	**df**	***F***	***p***
age group	-.0048	.0144	1, 68.22	0.11	.741
category choice weight	.0083	.0075	1, 270.95	1.25	.264
specificity (exemplar, category)	**.0423**	**.0040**	**1, 202.69**	**114.58**	**<.001**
block condition (exemplar-predictive, category-predictive)	**-.0192**	**.0041**	**1, 209.58**	**22.19**	**<.001**
age group x category choice weight	.0094	.0075	1, 270.95	1.58	.209
age group x specificity	.0040	.0040	1, 202.69	1.05	.308
category choice weight x specificity	**.0111**	**.0040**	**1, 202.69**	**7.75**	**.006**
age group x block condition	-.0019	.0041	1, 209.58	0.22	.643
category choice weight x block condition	.0011	.0045	1, 212.86	0.06	.801
specificity x block condition	.0061	.0040	1, 202.69	2.34	.127
age group x category choice weight	-.0013	.0040	1, 202.69	0.11	.741
age group x category choice weight x block condition	**.0115**	**.0045**	**1, 212.86**	**6.45**	**.012**
age group x specificity x block condition	.00035	.0040	1, 202.69	0.77	.380
category choice weight x specificity x block condition	.0025	.0040	1, 202.69	0.39	.531
age group x category choice weight x specificity x block condition	-.0031	.0040	1, 202.69	0.61	.435

B4. Effects of delay on memory performance

Participants were invited to complete the memory test six days after participating in the reinforcement-learning task, and had until 9 days after the learning task to complete it (one participant was invited on Day 5, and one completed the memory test on Day 10). Participants who had not completed the memory test by Day 8 received a reminder email on Day 8 and Day 9. To analyze the influence of the delay between learning and memory on memory performance, we re-ran our analysis examining the effects of age group, block condition and level of abstraction on memory, with delay (in days) as an interacting fixed effect. We continued to observe main effects of block condition and level of abstraction (*p*s < .001), as well as a marginal block condition x abstraction interaction (*p* = .052). Here, we also observed an abstraction x delay interaction effect, β = .0097 (.004), $F(1, 207) = 5.99$, $p = .015$ (Figure S9), such that exemplar-level memory was worse at longer delays. Thus, in line with prior research (Winocur et al., 2010), these data suggest that memory for detailed information may fade over time. As noted in our analysis of delay effects within Experiment 1, in our experimental design, memory delay may be confounded with other participant-level characteristics, and future work is needed to isolate the influence of delay time on the influence of reinforcement learning on memory across levels of abstraction.



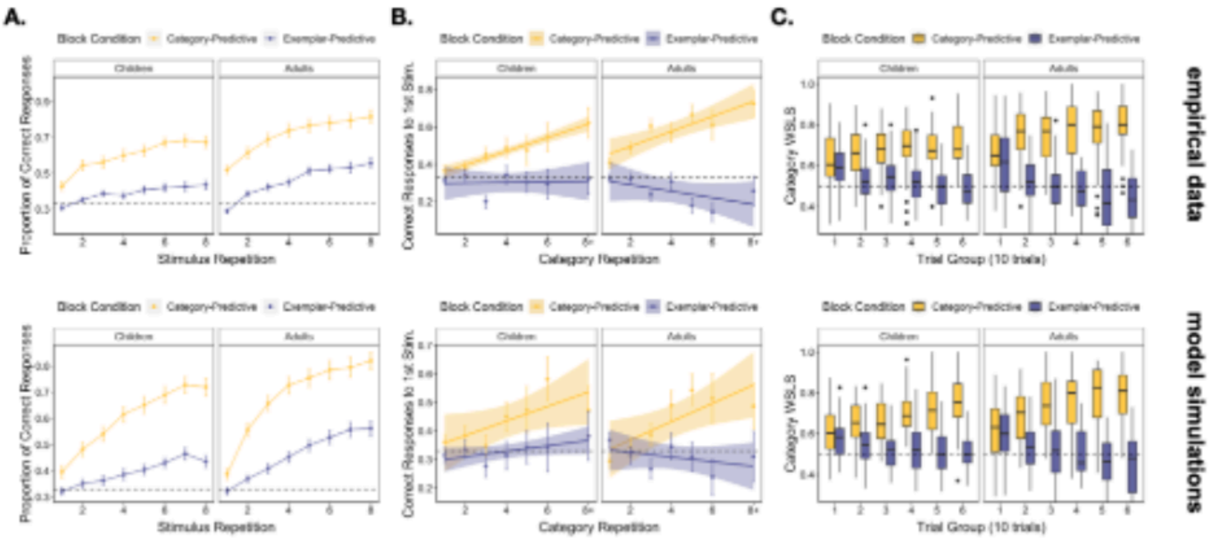
Supplementary Figure 9. Memory across levels of abstraction at different delays. Participants who completed the memory test at longer delays demonstrated worse memory for exemplar-level information than those who completed it at shorter delays. The boxplot hinges show the first and third quartiles of the data; whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians.

B5. Relations between age and model parameters

We also examined how the other model-derived parameters varied across age groups. We did not observe evidence for age-related change in learning rates or counterfactual learning rates (*ps* > .40). Here, we also did not find significant evidence for age-related changes in initial q values ($p = .684$).

B6. Posterior predictive checks

As in Experiment 1, to examine the extent to which our winning reinforcement-learning model (fourB_oneQ) could recapitulate patterns of participant behavior on the task, we simulated data from 100 'versions' of each participant, using their best-fitting parameters and trial sequence. We then analyzed this dataset of 7300 simulated participants in the same way that we analyzed our empirical data, and found that our model reproduced all key qualitative signatures of participants' learning behavior (Supplementary Figure 10).



Supplementary Figure 10. Data generated from the best-fitting model (fourB_oneQ) using participants' best-fitting parameters (bottom panels) captured many of the qualitative features of participants' real choice data (top panels). Data were first averaged over the 100 simulations run with each participant's trial sequence and parameter values and then averaged within each age group. A) Over the course of each block, both the empirical data and model simulations show that children and adults learned to make more optimal responses to stimuli in both block conditions, though performance was better in category-predictive relative to exemplar-predictive blocks. Points depict age group means and error bars show standard errors of participant (or agent) means. B) Both the empirical data and model simulations show that in the category-predictive condition, children and adults increasingly generalized learned category responses to respond optimally to novel stimuli. Points depict age group means and error bars show standard errors of participant means. The lines show the best-fitting linear regressions through the points, with the shaded region depicting 95% confidence intervals. C) Both the empirical data and model simulations show that category "win-stay lose-shift" behavior increased across trials in category-predictive blocks and decreased across trials in exemplar-predictive blocks, increasingly so with age. The boxplot hinges show the first and third quartiles of the data averaged within each trial group; whiskers extend to the values no further than 1.5 times the interquartile range; outlying points are plotted individually. Center lines show medians.