Development of non-local learning

Alice Zhang (az1517@nyu.edu)
NYU Department of Psychology, 6 Washington Place
New York, New York USA

Kate Nussenbaum (katenuss@nyu.edu)
NYU Department of Psychology, 6 Washington Place
New York, New York USA

Catherine Hartley (cate@nyu.edu)
NYU Department of Psychology, 6 Washington Place
New York, New York USA

Abstract

Across the lifespan, people are faced with sequential decision-making problems that require integrating information from separate learning experiences. While 'offline' integration of past experiences during replay has been shown to support this kind of non-local learning in adults, it is unknown whether children and adolescents are similarly able to update their behavior based on new information about the downstream consequences of their choices. We adapted a two-step reward revaluation task for use in an online developmental study, and found that children, adolescents, and adults demonstrated similar replanning behavior. We also examined the relation between replanning behavior and memory. Contrary to our hypothesis that 'offline' replay would lead to both increased replanning and enhanced memory, we found that replanning was associated with worse memory for trial-specific stimuli.

Keywords: development; decision-making; revaluation; replay; memory

Introduction

Children and adults alike face multi-step decision making problems in their daily lives. Non-local learning, or integrating information from separate learning episodes, is crucial for solving these kinds of problems. For example, if you often visit a candy store because it has the best chocolate, developing a chocolate allergy should motivate you to stop going to that store even if you did not have a negative experience there.

Prior work in adults has found that 'offline' replay during rest supports the ability integrate non-local experiences and make good decisions in multi-step planning problems (Momennejad, Otto, Daw, & Norman, 2018; Liu, Mattar, Behrens, Daw, & Dolan, 2021). Additionally, prior developmental work has shown that multi-step planning improves from childhood to adulthood (Decker, Otto, Daw, & Hartley, 2016). Here, we asked how non-local learning in a reward revaluation task changes across development. We hypothesized that even younger children would be able to integrate separate experiences and update their behavior successfully given the opportunity for 'offline' replay during a rest period.

We had an additional exploratory hypothesis about the relationship between non-local learning and memory. While we did not measure neural replay in this behavioral study, we predicted that stimulus reactivation during rest might enhance *both* memory (Tambini & Davachi, 2019) and non-local learning, such that we would observe a positive correlation between replanning behavior and stimulus memory.

Methods

Participants

101 participants between the ages of 7 and 23 years old (n = 29 children (7 - 11 years old), n = 36 adolescents (12 - 17 years old), n = 36 adults (18 - 23 years old) completed the online study.

Task

We adapted a two-step reward revaluation task (Momennejad et al., 2018) that was previously used in adults for use in an online, developmental sample (Fig 1). In the task, participants explored two environments in search for treasure. In each block of the task, participants completed four steps: Training, Relearning, Rest, and Test.

In Training, participants made two-stage decisions to earn treasure, first going up or down in the environment to find an animal, and then going left or right to select a chest. Chests contained between 5 and 50 pieces of treasure, and their values remained constant during Training. In Relearning, participants did not make first-stage choices, but were instead presented with each animal several times and asked to select a chest. Critically, in the Revaluation condition, rewards changed in Relearning so that the best chest belonged to a new animal. In the Control condition, rewards remained unchanged. During Rest, participants completed a non-cognitively demanding task for 60 seconds. Finally, at Test, they were asked to make first-stage choices without feedback. Our main measure of interest was whether in the Revaluation condition, participants were able to integrate new values from Relearning with their knowledge of the task's structure to navigate to the animal with the best chest.

To assess the link between memory and replanning, we selected images from two scene categories (ocean and canyon) that were matched for memorability (Lu, Xu, Yang, & Wang, 2020). Each first-stage choice during Training was paired with a trial-specific image belonging to the scene category for that block. After completing both blocks of the revaluation task, participants completed a recognition memory test that included both old and new scene images.

Results

Learning

We first asked whether participants learned to make optimal two-stage choices over the course of the Training phase by examining performance measured as the amount of treasure earned divided by the maximum amount of treasure available for each trial. Performance increased as a function of increasing within-block trial number ($F(225) = 5.3, p < .001$). Older participants performed better ($F(106) = 17.4, p < .001$) and learned faster ($F(376) = 4.1, p = 0.02$) relative to younger participants. As expected, we did not find a significant effect of block condition (Revaluation vs. Control) ($F(226) = 0.2, p = .67$), as these conditions were identical in the Training phase. Participants did perform better in the second block compared to the first block of the task ($F(225) = 5.3, p = .02$).

We also examined performance in the last ten Training trials of each block, as we use these trials to assess replanning behavior. We found no significant effect of age on performance in the last ten trials of Training ($F(97) = 2.2, p = 0.14$), but we did observe an age x block order interaction, such that younger participants performed worse in the second block ($F(97) = 6.5, p = .01$).

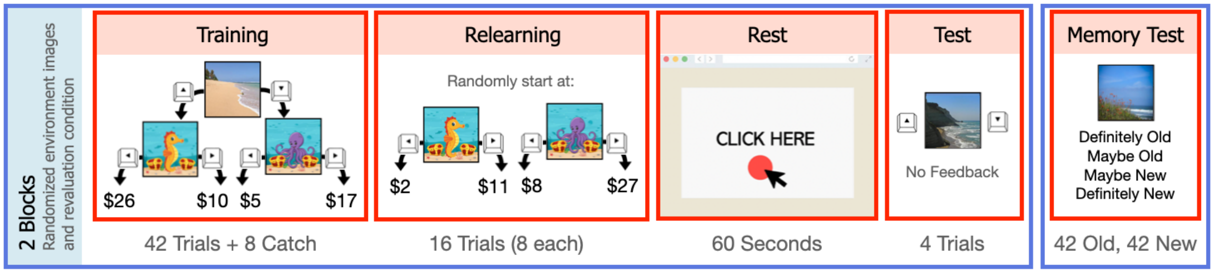

Figure 1: Task design. In each task block, participants made two-stage decisions to maximize reward, relearned second-stage reward values, completed an active rest task, and made first-stage decisions again without feedback. In the Revaluation block (shown here), reward values change in Relearning so that a new animal has the best treasure chest. In the Control block, second-stage rewards do not change. At the end, participants were asked to remember images associated with their first-stage choices.

Replanning

To assess non-local learning, we calculated a replanning score for each participant block that reflected the extent to which participants changed their first-stage choices from Training to Test. This replanning score is computed by taking the proportion of optimal first-stage choices at Test and subtracting the proportion of matching first-stage choices in the last 10 Training trials. In the Control condition, making the same first-stage choices during Test and Training was optimal, so participants who behaved optimally would have scores of 0. In the Revaluation condition, changing first-stage choices between Training and Test was optimal, and participants who behaved optimally would have scores of 1. As predicted, we found a significant effect of block condition (Revaluation vs. Control) on replanning score ($F(97) = 112.3, p < .001$), indicating that participants were more likely to switch their first-stage choice when they learned new second-stage reward values in the Relearning phase. We found no significant effect of age on replanning behavior ($F(97) = 0.2, p = .66$, Fig 2).

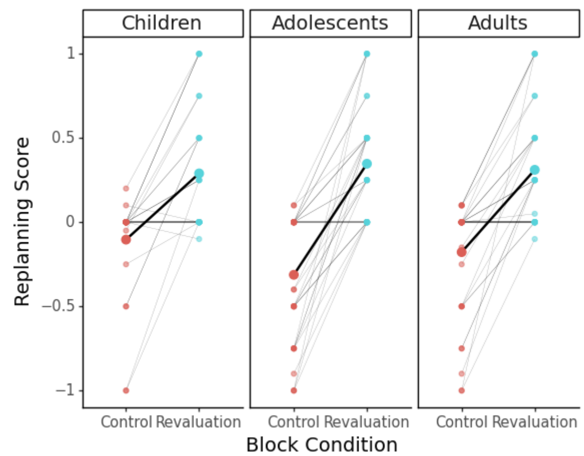

Figure 2: Replanning score (computed as the proportion of correct first-stage choices at Test minus the proportion of matching first-stage choices in the last ten trials of Training) across age groups and block conditions.

Memory

To assess the relation between memory for first-stage stimuli and replanning behavior, we computed a corrected recognition (CR) score for images from each block of the task for each participant. We originally hypothesized that memory would be enhanced in the Revaluation block due to increased replay of first-stage stimuli, and that this effect would be stronger for participants who showed greater replanning behavior. However, we observed a significant replanning score x block condition interaction effect on corrected recognition ($F(184) = 4.2, p = 0.04$) in the opposite direction. Increased replanning behavior in the Revaluation condition was associated with *reduced* corrected recognition for first-stage stimuli.

Discussion

We found that replanning behavior did not differ significantly across development in a two-step revaluation task. While previous work has shown that the ability to use 'online' model-based planning strengthens through childhood and adolescence (Decker et al., 2016), here, we found that younger children and adults demonstrated comparable integration of non-local experiences into learning.

Unexpectedly, we also found that increased replanning behavior in Revaluation blocks of our task was associated with worse memory for task stimuli. One possible explanation for this result is that offline replay preserved a more abstract category representation rather than the details of individual stimuli.

Going forward, we plan to run a version of the task without the Rest phase. Without the opportunity to 'replay' experiences and update value estimates during rest, children may show reduced non-local learning, whereas adolescents and adults may be better able to leverage 'online', model-based planning to use newly learned reward values at Test.

Acknowledgments

We thank the Hartley Lab for helpful feedback throughout, as well as the NIMH (F31 MH129105 to K.N. and R01 MH126183 to C.A.H.) for funding.

References

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological Science*, *27*(6), 848-858. (PMID: 27084852) doi: 10.1177/0956797616639301

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*(6544), eabf1357. doi: 10.1126/science.abf1357

Lu, J., Xu, M., Yang, R., & Wang, Z. (2020). Understanding and predicting the memorability of outdoor natural scenes. *IEEE Transactions on Image Processing*, *29*, 4927-4941. doi: 10.1109/TIP.2020.2975957

Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018, dec). Offline replay supports planning in human reinforcement learning. *eLife*, *7*, e32548. doi: 10.7554/eLife.32548

Tambini, A., & Davachi, L. (2019). Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends in Cognitive Sciences*, *23*(10), 876-890. doi: https://doi.org/10.1016/j.tics.2019.07.008