Characterizing the Contributions of Reward and Emotion Information on Learning Through Adolescence

Camille V. Phaneuf (cphaneuf@g.harvard.edu)
Harvard University Department of Psychology, 52 Oxford Street
Cambridge, MA 02138 USA

Elizabeth A. Phelps (phelps@fas.harvard.edu)
Harvard University Department of Psychology & Center for Brain Science, 52 Oxford Street
Cambridge, MA 02138 USA

Leah H. Somerville (somerville@fas.harvard.edu)
Harvard University Department of Psychology & Center for Brain Science, 52 Oxford Street
Cambridge, MA 02138 USA

Abstract

To behave adaptively, individuals of all ages must heed value information in their environments. This study characterizes how incidental and integral value cues shape learning from childhood to adulthood (*N*=114, 8-22 years). Within a probabilistic reinforcement learning task, emotional expressions conveyed incidental information while monetary rewards conveyed integral information. In some conditions, emotion and reward contributed to value in a congruent manner: following either cue promoted learning. In other conditions, emotion and reward contributed to value in an incongruent manner: following the emotion cue impeded reward learning. Model-free and computational modeling analyses revealed that, indeed, although participants of all ages adopted condition-wise learning strategies, younger participants' learning was most disrupted by emotion-reward incongruency. Meanwhile, older participants leveraged emotion-reward congruency to guide their choices to the greatest degree. Together, this work sheds light on age-related changes in the use of incidental and integral information for goal-directed actions.

Keywords: adolescence; development; emotion; reinforcement learning

Background

We are faced with a barrage of information throughout the lifespan. How do multiple information streams guide our behavior when those streams are in harmony vs. in conflict? Value information frequently takes form as emotional expressions or monetary rewards, and the current research examines how these cues distinctly and synergistically modulate learning from childhood to adulthood. We focus on modulation through adolescence because this period is associated with increased salience of value signals (Hartley & Somerville, 2015; Somerville, Jones, & Casey, 2010). Modulation is tested with a reinforcement learning paradigm and accompanying model-free and computational modeling analyses.

Methods

Emotion-Reward Manipulation of Learning

Participants (*N*=114, 8-22 years) completed a probabilistic reinforcement learning task containing emotional face stimuli (Lundqvist, Flykt, & Öhman, 1998). On each of the 180 trials, participants saw a pair of card decks with happy or fear faces. They selected a deck, then received feedback (50 or 0 cents). Participants were paid a bonus based on their choices and $20 for their time.

Critically, the decks corresponded to a 2x2 factorial design (Figure 1) crossing emotion (happy vs. fear face) and reward (75% vs. 25% reinforced). Consequently, there was differential congruence of emotion and reward information across four conditions, which were pseudo-randomly interspersed throughout the task.

Model-Free Analyses

To tease apart the impact of emotions and rewards on learning, we implemented a logistic mixed-effects model from the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017) in R. The *accuracy* of each choice made in the task (i.e., select high reward face) was predicted by *age*, *trial number*, *condition*, and their interactions. *Participant ID* was associated with a random intercept to account for repeated measurements. The significance of the main effects and interactions were assessed with analyses of deviance (type II Wald χ^2 tests) using the Anova function from the *car* package (Fox & Weisberg, 2019) in R.

Computational Modeling Analyses

We also implemented a series of standard temporal difference models (Sutton & Barto, 1998) in MATLAB. Only the best-fitting model (determined by the lowest mean and median Akaike Information Criterion and highest number of participants for whom the model explained the most variance) is detailed here. Models were fit to participants' data by maximizing the log posterior of their choices with the fmincon function.

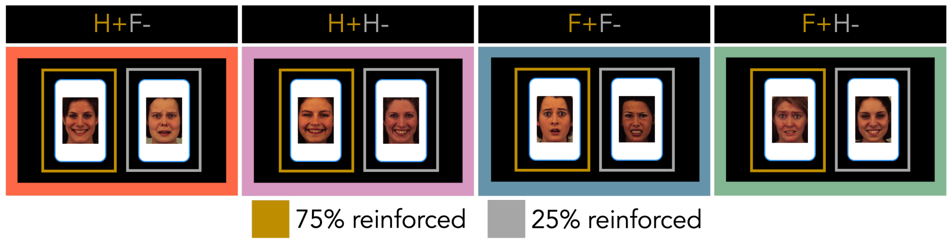

Figure 1: Orthogonal emotion-reward manipulation within a probabilistic reinforcement learning task.

Condition Model Specification After choosing deck d on trial t and experiencing reward r_t, participants updated their current Q-value estimate $Q_t(d)$ according to a condition-wise learning rate $\alpha \in \{\alpha_{H+F-}, \alpha_{H+H-}, \alpha_{F+F-}, \alpha_{F+H-}\}$ and the prediction error $\delta_t = r_t - Q_t(d)$:

$$Q_{t+1}(d) = Q_t(d) + \alpha * \delta_t$$

Q-values were initialized to .50 and were converted to choice probabilities via a softmax function with an inverse temperature parameter β:

$$P(d_t) = e^{\beta * Q_t(d)} / \sum_{d \in D} e^{\beta * Q_t(d)}$$

This model providing the best fit to the data suggests that participants utilized different learning strategies, adjusting their behavior to emotion-reward congruency, in each condition.

Derived Q-Differences From the Condition Model, we extracted trial-wise Q-value estimates, then computed *Q-differences* within the decision pair by subtracting the Q-value of the low reward face from the high reward face[1]. We implemented another logistic mixed-effects model with the *accuracy* of each choice predicted by *age*, *condition*, *Q-differences*, and their interactions. *Participant ID* and *trial number* were associated with random intercepts to account for repeated measurements. Effects were assessed using Anova.

Results

Model-Free Results

We found main effects of *age* (p<.001) and *trial number* (p<.001), as well as a two-way interaction between them (p<.001); *accuracy* improved with age and time, but most sharply for older participants. We also found a main effect of *condition* (p<.001) and an interaction between *age* and *condition* (p<.01; Figure 2A); *accuracy* was highest in H+F- trials and lowest in F+H- trials, especially for younger participants.

Computational Modeling Results

We again found main effects of *age* (p<.001) and *condition* (p<.001), with the same pattern of results as the model-free analyses. Additionally, we found a main effect of *Q-differences* (p<.001); as expected, greater differences between Q-value estimates were associated with improved *accuracy*. These main effects were qualified by a two-way interaction between *condition* and *Q-differences* (p<.001), a two-way interaction between *age* and *Q-differences* (p<.001), and a three-way interaction between *age*, *condition*, and *Q-differences* (p<.05; Figure 2B). Ultimately, with increasing differences between Q-value estimates, *accuracy* improved most in older participants across conditions (steeper coral lines). Moreover, when the low reward face in a pair was erroneously valued above the high reward face (i.e., Q-differences were near -1), older participants achieved greater accuracy in the H+F- condition while younger participants did not achieve accuracy gains in this condition relative to the others (gold vs. silver circles).

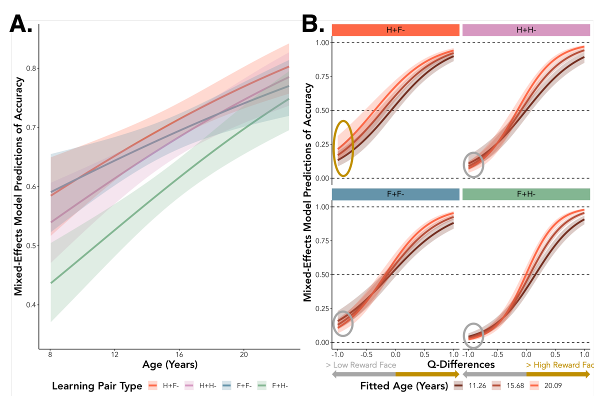

Figure 2: **A.** Learning improves with emotion-reward congruency most in younger participants. **B.** When Q-value estimates are mis-calibrated, emotion-reward congruency recovers accuracy most in older participants. Shading represents 95% confidence intervals around fitted lines.

Conclusions

In sum, this study characterizes age-related changes in the contributions of incidental emotion and integral reward information on learning with model-free and computational modeling analyses. Learning improves with age, across time, and with emotion-reward congruency; the latter is particularly observed in children. Differences in Q-value estimates extracted from the best-fitting Condition Model enrich the behavioral re-

[1]A Q-difference that reflects the reinforcement schedule of the task should be more likely to produce a correct response on that trial.

sults by uncovering a potential learning strategy. When Q-value estimates are mis-calibrated to the reinforcement schedule of the task, adults are more accurate than children and adolescents in the H+F- condition. Consistent with previous research indicating that diverse, flexible, and sophisticated choice strategies are increasingly adopted with age (Jacobs & Klaczynski, 2002), we show that while in a congruent emotion-reward environment, adults may adaptively rely on emotion information to make more optimal reward decisions.

Acknowledgments

This work was supported by Harvard University and a National Defense Science and Engineering Graduate Fellowship to Camille V. Phaneuf.

References

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression (third edition)*. Sage.

Hartley, C. A., & Somerville, L. H. (2015). The neuroscience of adolescent decision-making. *Current Opinion in Behavioral Sciences*, *5*, 108–115.

Jacobs, J. E., & Klaczynski, P. A. (2002). The development of judgment and decision making during childhood and adolescence. *Current Directions in Psychological Science*, *11*(24), 287–317.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). Karolinska directed emotional faces. *PsycTESTS Dataset*, *91*, 630.

Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: Behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition*, *72*(1), 124–133.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. *IEEE Transactions on Neural Networks*, *9*, 1054.