Early-life environments shape learning and decision-making strategies across development

Nora C. Harhen (nora.harhen@nyu.edu)[1], Noam Goldway[1,2], Jennifer C. Fielder[3], Maria Alejandra Martinez Ortiz[1,4], Yvette Q. Ma[1], Ethan McCormick[5] & Catherine A. Hartley[1]

[1]Department of Psychology, New York University
[2]Department of Psychology and Neuroscience, King's College London
[3]Division of Psychology and Language Sciences, University College London
[4]Department of Psychology, Stanford University
[5]Department of Education, University of Delaware

Abstract

During childhood and adolescence, sensitivity to the environment is often heightened, and individual differences in decision making emerge. However, it remains unclear which dimensions of experience produce these differences. Here, we focus on three candidate dimensions of early-life environments that theoretical work in reinforcement learning suggests should shape individuals' behavior. Environmental reward prevalence should govern learning from positive versus negative outcomes; controllability, reliance on reactive versus proactive control; and predictability, planning over a mental model. To test these predictions, we recruited a large sample of 10- to 25-year-olds to complete three reinforcement-learning tasks and a self-report measure of early-life experience. We found that behavior conformed to these predictions. Younger adolescents with less emotional support learned more slowly from positive outcomes, those with less control over their environment acted less proactively, and participants from less predictable environments planned more. Our results suggest that individual differences arise, in part, from rational adaptation to early-life environments.

Keywords: early-life experience; life-history effects; reinforcement learning; decision making

Introduction

Developmental researchers have long sought to understand how early-life environments shape individual differences. Research on early-life adversity has repeatedly shown that experiences such as abuse, neglect, and resource scarcity alter learning, decision making, and affect (Birnie et al., 2020; McLaughlin et al., 2021; Thomas et al., 2016). But, the mechanisms linking these experiences to later outcomes are incompletely understood. One proposed reason for this lack of mechanistic understanding is that adversity has historically been treated as a unitary construct, despite the heterogeneity of the experiences it encompasses. Dimensional approaches address this limitation by organizing adversity along dimensions such as threat, deprivation, and unpredictability, and hypothesizing that each dimension alters a distinct learning mechanism (McLaughlin & Sheridan, 2016). Recent studies adopting this approach have clarified *how* early-life experience shapes cognitive and affective processes. However, the question of *why* these experiences produce their effects remains open.

To understand *why* specific early-life experiences alter specific aspects of learning and decision making, we turn to normative modeling work from reinforcement learning (RL). These models specify how an agent ought to behave to maximize future expected reward in a given environment. Importantly, when human behavior aligns with these normative prescriptions, it *explains* individuals' behavior as a consequence of rational adaptation to the statistical structure of the environments they've encountered. Normative work in reinforcement learning identifies three key environmental dimensions for learning and decision making: reward prevalence, controllability, and predictability. These share similarities to dimensions used in prior studies of adversity. In addition, these normative models are particularly useful because they make precise predictions regarding which aspects of learning and decision-making strategies each dimension should influence, achieving a level of specificity not matched by prior applications of the dimensional approach. We focus on three predictions here. First, reward prevalence should shape learning from positive versus negative outcomes (Cazé & van der Meer, 2013; Jaskir & Frank, 2023). Second, controllability (i.e., the extent to which an agent's actions causally influence the environment) should determine reliance on innate reactive behaviors (i.e., approaching reward, avoiding threat) versus proactive, goal-directed control (Huys & Dayan, 2009; Moscarello & Hartley, 2017). Third, predictability should govern the use of a mental model for planning (Kool et al., 2017; Simon & Daw, 2011). When these environmental features change over the course of a single task, people systematically adjust how they learn and act in ways that align with these normative predictions (Jaskir & Frank, 2023; Kool et al., 2017; Raab et al., 2024). Here, we ask whether these three predictions also hold across developmental timescales. A core assumption of our approach is that development is an adaptive process in which individuals tune their learning and decision-making strategies to the structure of their environments (Nussenbaum & Hartley, 2024).

To test this account, we recruited a large developmental sample of 10- to 25-year-olds who completed self-report measures characterizing early-life experience along the dimensions of reward prevalence, controllability, and predictability, as well as three reinforcement-learning tasks indexing distinct strategies. Although our sample spanned adolescence and young adulthood, we did not have predictions *a priori* regarding how the relationship between the early-life environment and reinforcement-learning strategies might vary with age. If these normative predictions extend across development, they offer a principled explanation for why early-life environments give rise to specific patterns of individual differences in decision making and would provide a guide for generating novel, testable hypotheses.

Methods

Participants

We plan to recruit 1,000 individuals aged 10–25 to complete three reinforcement-learning tasks and an early-life experience questionnaire across four online sessions administered in a fixed order. Participants are assigned to a discovery sample for exploratory analyses or a replication sample for validation. We report results from 284 individuals in the discovery sample. Participants were recruited from our lab database, which includes individuals from diverse socioeconomic, racial, and ethnic backgrounds across the United States. Sampling was uniform across age, with equal numbers of females and males. For each of the four sessions, participants were compensated $15 per hour and had the opportunity to earn a bonus payment based on task performance.

Measuring early-life experience

We developed a self-report questionnaire to characterize early-life experiences along four dimensions: reward prevalence, threat prevalence, predictability, and controllability. Our primary dimensions of interest were reward prevalence, predictability, and controllability, with threat included given its prominence in prior developmental research (Felitti et al., 1998; Green et al., 2010; Kessler et al., 2010; McLaughlin et al., 2012). Following Berman et al., 2021, we assembled items from established measures, including the Deprivation–Threat Self-Report (Tsai et al., 2025), the Child Report of Parent Behavior (Schaefer, 1965), the Nowicki–Strickland Locus of Control Scale (Nowicki & Strickland, 1973), and the Neighborhood Context Questionnaire (York Cornwell & Cagney, 2014). The questionnaire assessed participants' subjective experiences in their households and communities, including relationship stability, emotional support from caregivers, peers, and teachers, and perceived neighborhood safety. Attention check questions were interspersed throughout; participants who failed two or more were excluded.

Although the questionnaire was designed around hypothesized dimensions, we evaluated its structure using exploratory factor analysis. The number of factors was determined via parallel analysis, with solutions evaluated for interpretability. Because parallel analysis can overestimate factor number (Michelini et al., 2019), we iteratively reduced the solution until each retained factor included at least three items with loadings ≥ 0.32 (Yong & Pearce, 2013). Factor scores from the final solution were extracted for each participant and used as measures of early-life experience along the identified dimensions.

Measuring valence asymmetries in learning

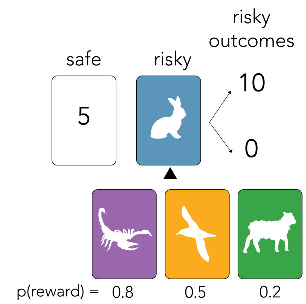

Figure 1: Participants repeatedly choose between a safe option (a certain +5 points) and a risky option (some probability of +10 or 0 points) and observe the outcomes of both the chosen and unchosen options. This design measures how individuals learn from positive vs. negative prediction errors.

Task We adapted a risk-sensitive reinforcement learning task (Fig. 1; Niv et al., 2012). On each of 240 trials, participants chose between a safe card that guaranteed 5 points and a risky card that yielded either 0 or 10 points, with reward probabilities of 20%, 50%, or 80%. When the risky card had a 50% probability, the two options were matched in expected value; therefore, a decision maker who weights positive and negative outcomes equally should be indifferent. However, asymmetrically weighing valenced outcomes will distort the values of the risky options such that someone who overweights positive outcomes will prefer the risky card, while someone who overweights negative outcomes will prefer the safe option. To ensure learning of the risky cards' expected values regardless of risk preferences, the outcome of the risky card was shown on every trial. Throughout the task, new cards were introduced to encourage continual learning.

Model We quantified participants' weighting of positive versus negative outcomes using a model developed by Niv et al., 2012. The model learns the expected value of each risky card, Q(c), through prediction-error-driven updating, where δ reflects the difference between the actual and expected outcome. Updates are scaled by a valence-dependent learning rate, α:

$$\begin{cases} Q(c) \leftarrow Q(c) + \alpha_{+}\delta & \text{if } \delta > 0 \\ Q(c) \leftarrow Q(c) + \alpha_{-}\delta & \text{if } \delta < 0 \end{cases} \tag{1}$$

A valence asymmetry index (AI) captures the relative weighting of positive versus negative outcomes:

$$AI = \frac{\alpha_{+} - \alpha_{-}}{\alpha_{+} + \alpha_{-}} \tag{2}$$

Positive values indicate greater weighting of positive outcomes, whereas negative values indicate stronger weighting of negative outcomes.

Exclusion Criteria Of the 284 discovery-sample participants, we excluded those who failed to respond or responded too quickly (<150 ms) on 10 or more trials, leaving 253 participants.

Measuring reactive versus proactive control

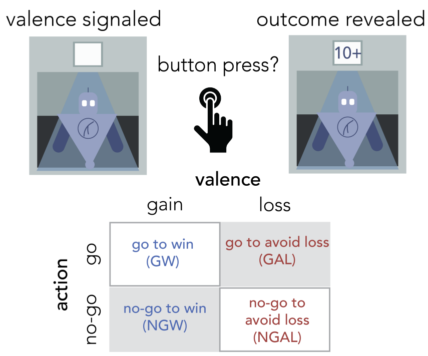

Figure 2: The color of the light (blue or red) cued whether a robot could deliver reward (gain) or punishment (loss). On each trial, participants must decide to "repair" (Go) or do nothing (No-Go) to the robot. If the participant performs the correct action, then there is an 80% chance that they will experience the better outcome (winning points or avoiding loss). Certain trials align with innate "Pavlovian" tendencies to approach reward or withdraw from potential punishment (i.e., Go to win, No-Go to avoid losing), whereas others are incongruent (e.g., No-Go to win). This task assesses the extent Pavlovian approach/avoid tendencies interfere with an individual's instrumental (action-outcome) learning.

Task Participants completed a task measuring the extent to which Pavlovian biases facilitate or interfere with instrumental learning (Fig. 2; Guitart-Masip et al., 2011). Pavlovian biases reflect tendencies to approach reward and avoid punishment, which can either support or conflict with goal-directed action. We used a child-friendly probabilistic go/no-go task (Raab & Hartley, 2020; Zorowitz et al., 2023) in which action (go vs. no-go) and valence (reward vs. punishment) were orthogonalized. On each of 240 trials, participants saw a robot marked with a unique symbol and chose whether to repair it (go) or withhold a response (no-go). A colored light (blue or red) indicated whether the correct response would yield points or prevent point loss, with correct responses producing the better outcome 80% of the time. Pavlovian bias was indexed as the difference in performance between trials in which action and valence were aligned (go to win; no-go to avoid loss) versus misaligned (no-go to win; go to avoid loss).

Model We quantified reactive responding as the extent to which Pavlovian biases influenced instrumental learning using a model developed by Guitart-Masip et al., 2011. The model learns action values, $Q_k(go)$ and $Q_k(no-go)$, for each robot, k via prediction-error-driven updating:

$$Q_k(action) \leftarrow Q_k(action) + \alpha\delta \tag{3}$$

where α is the learning rate and δ is the difference between observed and expected outcomes. Outcomes are coded as 1 for better and 0 for worse outcome.

Choices are made according to a softmax function:

$$p(action = go) = (1-\epsilon) * \frac{e^{\beta_v Q_k(go)+\tau_v}}{e^{\beta_v Q_k(go)+\tau_c} + e^{\beta_v Q_k(no-go)}} + \frac{\epsilon}{2} \tag{4}$$

where ϵ is the lapse rate, β_v is the valence-dependent (gain or loss) outcome sensitivity, and τ_v is the valence-dependent approach-avoidance bias. Pavlovian bias was defined as the difference between τ_{gain} and τ_{loss} with more positive values indicating a stronger bias and more reactive responding.

Exclusion Criteria Of the 284 discovery-sample participants, we excluded those whose accuracy across all trials was < 50%, leaving 279 participants.

Measuring forward planning

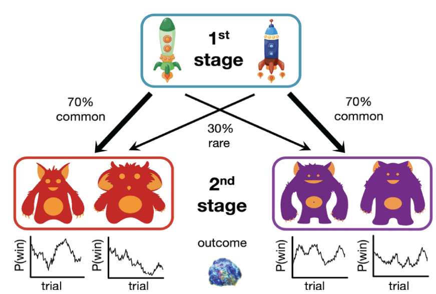

Figure 3: Participants make two-stage choices to collect treasure on different "planets," with a 70/30 transition from each spaceship to a planet. Rewards fluctuate slowly, requiring continual updating. Model-free learners tend to repeat rewarded first-stage choices regardless of whether a transition was common or rare, whereas model-based learners take into account the task transition structure (staying or switching based on transition type and outcome).

Task We used a child-friendly version (Fig. 3; Decker et al., 2016) of a two-stage decision task designed to dissociate model-free and model-based reinforcement learning (Daw et al., 2011). Model-free values are learned incrementally through trial and error, whereas model-based values are computed by planning over an internal model of the environment. On each of 200 trials, participants chose between two spaceships (first-stage choice), each of which transitioned probabilistically to one of two planets. One spaceship led to the red planet on 70% of trials (common transition) and to the purple planet on the remaining 30% (rare transition), with the opposite transition structure for the other spaceship. Upon arrival, participants chose between two aliens with slowly and independently drifting reward probabilities(second-stage choice). Responses following rare transitions are particularly diagnostic of which of the two strategies a participant is using. After a rare transition resulting in reward, a model-free learner will select the previously chosen spaceship again. In contrast, a model-based learner will switch to the alternative spaceship, which is more likely to lead to the planet on which they just collected a reward.

Model We quantified reliance on these strategies using a model developed by Daw et al., 2011. For each spaceship, s, the model maintains both model-free values, $Q_{MF}(s)$, and model-based values, $Q_{MB}(s)$. For each alien, a, on a planet, there is only a single value, $Q(planet, a)$. $Q_{MF}(s)$ and $Q(planet, a)$ are updated through error-driven learning. In contrast, model-based values are computed from the transition probabilities and alien values:

$$Q_{MB}(s) = \sum_{planet} p(planet|s) max_a Q(planet, a) \tag{5}$$

Choice probabilities combine model-free and model-based values:

$$p(action = s) \propto e^{\beta_{MF} Q_{MF}(s) + \beta_{MB} Q_{MB}(s)} \tag{6}$$

We use the model-based inverse softmax temperature, β_{MB}, as our measure of forward planning.

Exclusion criteria From the 284 discovery-sample participants, we excluded those who failed to respond or had response times <150 ms on 10 or more first- and second-stage choice trials, leaving 232 participants.

Model fitting

We fit all models using hierarchical Bayesian inference with Hamiltonian Monte Carlo. Each model was run with four chains (10,000 samples per chain; 5,000 warm-up), leaving 5,000 post-warm-up samples from the joint posterior. All parameters showed good convergence ($\hat{R}$ less than or equal to 1.01). Weakly informative priors were used to minimize parameter estimation bias.

Exploratory factor analysis of early-life experience

The results from the exploratory factor analysis largely recapitulated our hypothesized structure (Fig. 4), with parallel analysis and interpretability analyses supporting a four-factor solution. One factor, *predictability*, comprised items indexing the stability of access to caregivers, housing, food, and other resources (e.g., "How often did you not have a permanent place to live?"). 13 of 19 items assigned *a priori* to the predictability dimension had loadings ≥ 0.32 on this factor. Items assigned beforehand to reward prevalence split into two factors based on the source of social reward/support: *adult support* (16/21 items; e.g., "How often did a caregiver show an interest in your thoughts and feelings?") and *peer support*

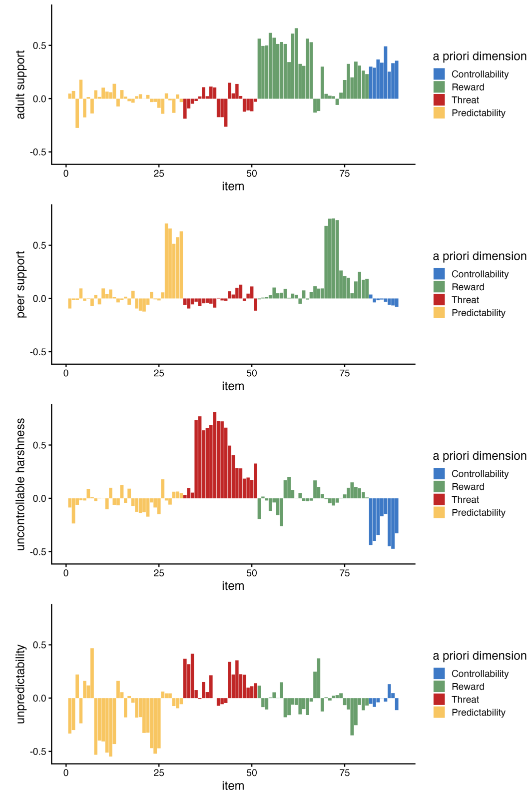

Figure 4: Bars indicate the item loadings for each factor. They are color-coded to indicate the the dimension the item was a priori assigned to. Largely, the factor structure reflects the dimensional structure we constructed the questionnaire around. Items capturing environmental predictability loaded strongly and coherently onto the same factor which we term the *predictability* factor. Items assessing environmental reward prevalence split into two factors, *adult support* and *peer support*, based on the provider of social rewards. Items measuring the dimensions of controllability and threat prevalence loaded onto the same factor which we term *uncontrollable harshness*.

(4/9 items; e.g., "I felt safe sharing my feelings and problems with my close friends"). A fourth factor, *uncontrollable harshness*, comprised items *a priori* assigned to the controllability (6/18 items; e.g., "When you got punished, how often did it seem like it was for no good reason at all?") and threat dimensions (12/18 items; e.g., "How often did an adult seriously act like or told you they were going to hurt you?"). This factor structure broadly aligned with our hypothesized dimensions, sufficiently so to enable us to test our three hypotheses. In subsequent analyses, we use factor scores as our measures of early-life experience, treating adult and peer support as reflecting reward prevalence, uncontrollable harshness as reflecting low controllability, and predictability as reflecting, naturally, predictability.

Relating early-life experience to reinforcement learning strategies

To examine how early-life experience relates to reinforcement-learning strategies, we fit five regression models, each predicting a different model-derived parameter. Parameters included were the valence asymmetry index (and its component individual learning rates), Pavlovian bias, and model-based inverse softmax temperature. Predictors included age, age squared (to capture nonlinear effects), all factor scores derived from the exploratory factor analysis, and their interactions with age and age squared. All predictors were z-scored prior to analysis. We report effects that were significant ($p < .05$) or trending ($p < .1$) given ongoing data collection.

Results

Early-life reward prevalence predicts sensitivity to positive outcomes

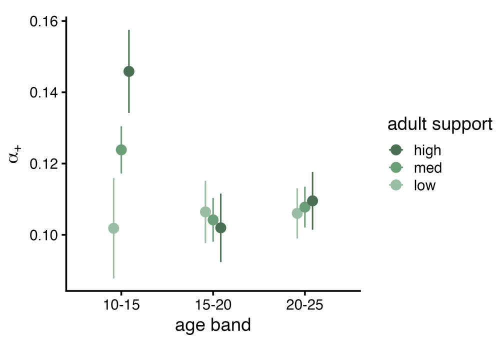

Figure 5: Estimated marginal means plots showing the interaction between age and adult support in predicting the learning rate for positive prediction errors. Among younger participants, lower social support from adults was associated with reduced positive learning rates; this effect attenuated with age (adult support x age: $b = -0.010$, SE=0.005, $p = .05$; adult support x age squared: $b = 0.011$, SE=0.010, $p = .04$). Age and adult support are plotted as discrete bins for visualization, but all analyses treat these predictors as continuous. Here, age has been split into three 5-year age groups, and adult support into terciles. Error bars denote S.E.M.

First, we examined whether early-life experience predicted variation in the valence asymmetry index (AI).There were no main effects of early-life experience. However, there was a main effect of age. AI decreased with age ($b = -0.033$, SE=0.009, $p < .001$), indicating that older participants had a stronger bias toward learning from negative relative to positive outcomes. There was no quadratic effect of age. We also observed no interactions between early-life experience and age.

We next examined positive and negative learning rates separately. There were no main effects of early-life experience on either, but age was associated with both learning rates (positive: $b = -0.009$, SE=0.004, $p = .02$; negative: $b = 0.006$, SE=0.002, $p < .001$). We then examined interactions. Positive learning rates showed a significant interaction between age and adult support. Among younger participants, lower adult support was associated with reduced positive learning rates (Fig. 5), consistent with normative accounts suggesting that sensitivity to positive outcomes should be heightened in richer environments (Jaskir & Frank, 2023). This association attenuated with age (adult support x age: $b = -0.010$, SE=0.005, $p = .05$; adult support x age squared: $b = 0.011$, SE=0.010, $p = .04$). No interactions were observed for negative learning rates or other early-life factors.

Early-life controllability predicts Pavlovian bias

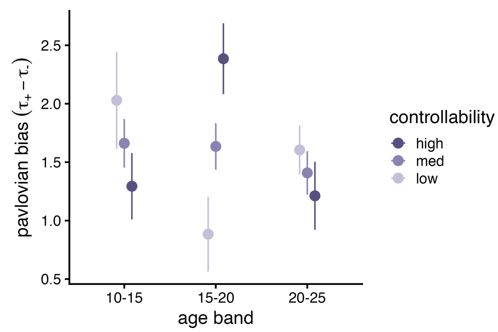

Figure 6: Estimated marginal means plots showing the interaction between age and uncontrollable harshness in predicting Pavlovian bias. There was a trending quadratic effect of age on Pavlovian bias and no main effect of uncontrollable harshness, but a significant interaction between the two ($b = -0.31$, SE=0.15, $p = .04$). Younger participants raised in less controllable environments show a stronger Pavlovian bias, while participants in the middle of our age range showed a *weaker* bias. The effect of uncontrollable harshness was weakest in participants at the end of the age range.

Next, we tested whether early-life experience predicted variation in Pavlovian bias.There were no main effects of early-life experience, but there was a trending quadratic effect of age ($b = 0.26$, SE=0.14, p = .06), with participants in the middle of the age range showing the weakest Pavlovian biases. There was no linear effect of age.

Next, examining the interactions between early-life experience and age, we found that uncontrollable harshness was the only factor to significantly interact with age ($b = -0.31$, SE=0.15, $p = .04$). Among younger participants, lower controllability was associated with stronger Pavlovian bias (Fig. 6), consistent with the normative predictions that environmental controllability should shape reliance on innate Pavlovian reactive responding versus more proactive, instrumental responding (Huys & Dayan, 2009; Moscarello & Hartley, 2017). Strikingly, this relationship reversed for individuals in the middle of our age range such that lower controllability was associated with *weaker* Pavlovian biases.

Early-life predictability predicts planning

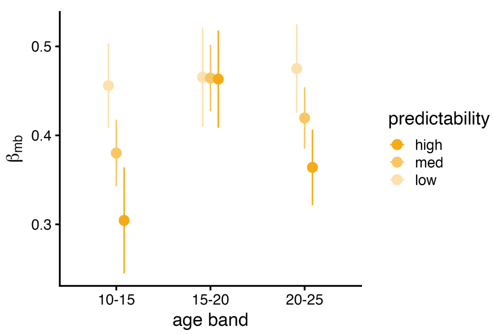

Figure 7: Estimated marginal means plots. Predictability was associated with greater engagement in planning across our age range, but particularly so for younger adolescents (predictability: b=0.061, SE=0.031, p = .06; age squared x predictability: b=0.045, SE=0.025, p = .08).

Finally, we tested whether early-life experience predicted planning over a mental model of the environment, indexed by the model-based inverse softmax temperature. We observed a trending main effect of predictability (Fig. 7; $b = -0.061$, SE=0.031, p = .06), such that individuals from less predictable environments planned more, consistent with normative accounts suggesting that model-based control is advantageous in unpredictable environments because it supports flexibility and rapid adaptation(Simon & Daw, 2011). There were no main effects of other early-life factors. We also observed a trending quadratic effect of age ($b = -0.045$, SE=0.025, p = .08), with participants in the middle of the age range engaging in planning the most.

We then examined the interactions between early-life experience and age. No interactions were observed between predictability and age, but there were for the three remaining factors. Among younger participants, lower controllability was marginally associated with less planning (age x controllability: $b = -0.041$, SE=0.025, p = .09), while both lower adult and peer support was associated with more planning (age x adult support: b = 0.047, SE=0.028, p = .09; age x peer support: b = 0.038, SE=0.021, p = .07). These associations were attenuated in older participants.

Discussion

Here, we examined whether individual differences in learning and decision making are associated with features of the early-life environment. We tested three normative predictions about how humans might adapt their strategies to features of their environments over developmental timescales.Younger participants with less adult support learned more slowly from positive outcomes, those from less controllable environments relied more on Pavlovian responding, and those from less predictable environments engaged in greater planning. Our findings suggest a compelling possibility for why we see the pattern of individual differences we do: they reflect an individual's rational adaptation to the structure and dynamics of their environment.

A striking feature of our results is that associations between early-life experience and learning varied with age, with effects strongest in younger participants. One possibility is that these strategies are continuously calibrated by the current environment even after adolescence. Although all participants reported retrospectively, adults have greater distance from the experiences they're asked to recall. Future work could test this account by incorporating measures of adults' current environmental features, which may interact with their early-life experiences in shaping behavior.

Another possibility is that effects of early-life experience are difficult to detect in simplified tasks, but are more apparent in real-world behavior. Real-world decisions are made under considerable uncertainty and their consequences can unfold over many years. Early-life effects may only reveal themselves under greater stress and pressure. The robust association between early-life experience and mental illness is consistent with this possibility Felitti et al., 1998; Green et al., 2010. Longitudinal studies tracking environmental features, reinforcement-learning strategies, and real-world behavior will be important for testing this account.

For planning and valence asymmetries in learning rates, early-life effects attenuated with age. In contrast, for Pavlovian bias, the effect reversed: younger participants from less controllable environments showed stronger Pavlovian biases, whereas participants from the middle of our age range showed weaker biases. It could be that older adolescents from less controllable homes increasingly seek opportunities for instrumental control outside of the home. Consistent with this, items loading onto our uncontrollable harshness factor primarily index relationships with caregivers and the home environment. Future work could extend this measure to capture more experiences outside the household.

Our work aligns with a large body of work showing that early-life adversity shapes learning and decision making (Amir et al., 2018; Birn et al., 2017; Guassi Moreira et al., 2022; Hanson et al., 2017; Lloyd et al., 2022). More recent work has moved beyond cumulative adversity to identify environmental dimensions that exert distinct effects on development (Davis et al., 2017; Doom et al., 2016; Ellis et al., 2009; Farkas & Jacquet, 2024; Lambert et al., 2017; Machlin et al., 2019; Sheridan et al., 2017; Simpson et al., 2012; Xu et al., 2023). Our results help bridge dimensional models of adversity with normative and mechanistic theories of learning and decision making, and lay the groundwork for using reinforcement learning to guide hypothesis generation in future work.

References

Amir, D., Jordan, M. R., & Rand, D. G. (2018). An uncertainty management perspective on long-run impacts of adversity: The influence of childhood socioeconomic status on risk, time, and social preferences. *J. Exp. Soc. Psychol.*, *79*, 217–226.

Berman, I., McLaughlin, P., Katie A, Tottenham, N., Godfrey, K., Seeman, T. E., Eric, L. B., Reiss, D., Suomi, S., Danese, A., & Sheridan, P. D., Margaret A. (2021). Measuring early life adversity across the lifespan.

Birn, R. M., Roeber, B. J., & Pollak, S. D. (2017). Early childhood stress exposure, reward pathways, and adult decision making. *Proc. Natl. Acad. Sci. U. S. A.*, *114*(51), 13549–13554.

Birnie, M. T., Kooiker, C. L., Short, A. K., Bolton, J. L., Chen, Y., & Baram, T. Z. (2020). Plasticity of the reward circuitry after early-life adversity: Mechanisms and significance. *Biol. Psychiatry*, *87*(10), 875–884.

Cazé, R. D., & van der Meer, M. A. A. (2013). Adaptive properties of differential learning rates for positive and negative outcomes. *Biol. Cybern.*, *107*(6), 711–719.

Davis, E. P., Stout, S. A., Molet, J., Vegetabile, B., Glynn, L. M., Sandman, C. A., Heins, K., Stern, H., & Baram, T. Z. (2017). Exposure to unpredictable maternal sensory signals influences cognitive development across species. *Proc. Natl. Acad. Sci. U. S. A.*, *114*(39), 10390–10395.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

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. *Psychol. Sci.*, *27*(6), 848–858.

Doom, J. R., Vanzomeren-Dohm, A. A., & Simpson, J. A. (2016). Early unpredictability predicts increased adolescent externalizing behaviors and substance use: A life history perspective. *Dev. Psychopathol.*, *28*(4pt2), 1505–1516.

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk. *Hum. Nat.*, *20*(2), 204–268.

Farkas, B. C., & Jacquet, P. O. (2024). Characterizing the role of unpredictability within different dimensions of early life adversity. *Dev. Psychopathol.*, 1–15.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. the adverse childhood experiences (ACE) study. *Am. J. Prev. Med.*, *14*(4), 245–258.

Green, J. G., McLaughlin, K. A., Berglund, P. A., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2010). Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: Associations with first onset of DSM-IV disorders. *Arch. Gen. Psychiatry*, *67*(2), 113–123.

Guassi Moreira, J. F., Méndez Leal, A. S., Waizman, Y. H., Saragosa-Harris, N., Ninova, E., & Silvers, J. A. (2022). Early caregiving adversity differentially shapes behavioral sensitivity to reward and risk during decision-making. *J. Exp. Psychol. Gen.*, *151*(12), 3249–3267.

Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J. M., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J. Neurosci.*, *31*(21), 7867–7875.

Hanson, J. L., van den Bos, W., Roeber, B. J., Rudolph, K. D., Davidson, R. J., & Pollak, S. D. (2017). Early adversity and learning: Implications for typical and atypical behavioral development. *J. Child Psychol. Psychiatry*, *58*(7), 770–778.

Huys, Q. J. M., & Dayan, P. (2009). A bayesian formulation of behavioral control. *Cognition*, *113*(3), 314–328.

Jaskir, A., & Frank, M. J. (2023). On the normative advantages of dopamine and striatal opponency for learning and choice. *Elife*, *12*(e85107).

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., Aguilar-Gaxiola, S., Alhamzawi, A. O., Alonso, J., Angermeyer, M., Benjet, C., Bromet, E., Chatterji, S., de Girolamo, G., Demyttenaere, K., Fayyad, J., Florescu, S., Gal, G., Gureje, O., . . . Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO world mental health surveys. *Br. J. Psychiatry*, *197*(5), 378–385.

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychol. Sci.*, *28*(9), 1321–1333.

Lambert, H. K., King, K. M., Monahan, K. C., & McLaughlin, K. A. (2017). Differential associations of threat and deprivation with emotion regulation and cognitive control in adolescence. *Dev. Psychopathol.*, *29*(3), 929–940.

Lloyd, A., McKay, R. T., & Furl, N. (2022). Individuals with adverse childhood experiences explore less and underweight reward feedback. *Proc. Natl. Acad. Sci. U. S. A.*, *119*(4).

Machlin, L., Miller, A. B., Snyder, J., McLaughlin, K. A., & Sheridan, M. A. (2019). Differential associations of deprivation and threat with cognitive control and fear conditioning in early childhood. *Front. Behav. Neurosci.*, *13*, 80.

McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2012). Childhood adversities and first onset of psychiatric disorders in a national sample of US adolescents. *Arch. Gen. Psychiatry*, *69*(11), 1151–1160.

McLaughlin, K. A., & Sheridan, M. A. (2016). Beyond cumulative risk: A dimensional approach to childhood adversity: A dimensional approach to childhood adversity. *Curr. Dir. Psychol. Sci.*, *25*(4), 239–245.

McLaughlin, K. A., Sheridan, M. A., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2021). The value of dimensional models of early experience: Thinking clearly about concepts and categories. *Perspect. Psychol. Sci.*, *16*(6), 1463–1472.

Michelini, G., Barch, D. M., Tian, Y., Watson, D., Klein, D. N., & Kotov, R. (2019). Delineating and validating higher-order dimensions of psychopathology in the adolescent brain cognitive development (ABCD) study. *Transl. Psychiatry*, *9*(1), 261.

Moscarello, J. M., & Hartley, C. A. (2017). Agency and the calibration of motivated behavior. *Trends Cogn. Sci.*, *21*(10), 725–735.

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J. Neurosci.*, *32*(2), 551–562.

Nowicki, S., & Strickland, B. R. (1973). A locus of control scale for children. *J. Consult. Clin. Psychol.*, *40*(1), 148–154.

Nussenbaum, K., & Hartley, C. A. (2024). Understanding the development of reward learning through the lens of meta-learning. *Nature Reviews Psychology*, 1–15.

Raab, H. A., Goldway, N., Foord, C., & Hartley, C. A. (2024). Adolescents flexibly adapt action selection based on controllability inferences. *Learn. Mem.*, *31*(3), a053901.

Raab, H. A., & Hartley, C. A. (2020). Adolescents exhibit reduced pavlovian biases on instrumental learning. *Sci. Rep.*, *10*(1), 15770.

Schaefer, E. S. (1965). Children's reports of parental behavior: An inventory. *Child Dev.*, *36*(2), 413.

Sheridan, M. A., Peverill, M., Finn, A. S., & McLaughlin, K. A. (2017). Dimensions of childhood adversity have distinct associations with neural systems underlying executive functioning. *Dev. Psychopathol.*, *29*(5), 1777–1794.

Simon, D., & Daw, N. (2011). Environmental statistics and the trade-off between model-based and TD learning in humans. *Adv. Neural Inf. Process. Syst.*, *24*.

Simpson, J. A., Griskevicius, V., Kuo, S. I.-C., Sung, S., & Collins, W. A. (2012). Evolution, stress, and sensitive periods: The influence of unpredictability in early versus late childhood on sex and risky behavior. *Dev. Psychol.*, *48*(3), 674–686.

Thomas, A. W., Caporale, N., Wu, C., & Wilbrecht, L. (2016). Early maternal separation impacts cognitive flexibility at the age of first independence in mice. *Dev. Cogn. Neurosci.*, *18*, 49–56.

Tsai, A. P.-T., Halpin, P. F., Lurie, L., Gruhn, M., Rosen, M., Baucom, D. H., Sarabosing, M. B., Boda, S. S., McLaughlin, K. A., & Sheridan, M. A. (2025). The development and validation of a dimensional childhood adversity measure. *Assessment*, (10731911251367142), 10731911251367142.

Xu, Y., Harms, M. B., Green, C. S., Wilson, R. C., & Pollak, S. D. (2023). Childhood unpredictability and the development of exploration. *Proc. Natl. Acad. Sci. U. S. A.*, *120*(49), e2303869120.

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutor. Quant. Methods Psychol.*, *9*(2), 79–94.

York Cornwell, E., & Cagney, K. A. (2014). Assessment of neighborhood context in a nationally representative study. *J. Gerontol. B Psychol. Sci. Soc. Sci.*, *69 Suppl 2*(Suppl 2), S51–63.

Zorowitz, S., Karni, G., Paredes, N., Daw, N., & Niv, Y. (2023). Improving the reliability of the pavlovian go/no-go task for computational psychiatry research. *PsyArXiv*.