

Interval timing as a computational pathway from early life adversity to affective disorders.

Nora C. Harhen (nharhen@uci.edu)^{1,*}

Aaron M. Bornstein (aaron.bornstein@uci.edu)^{1,2}

1. Department of Cognitive Sciences, University of California, Irvine, Irvine, CA 92697 USA

2. Center for the Neurobiology of Learning and Memory, University of California, Irvine, Irvine, CA 92697 USA

* To whom correspondence should be addressed.

I. ABSTRACT

Adverse early life experiences can have remarkably enduring negative consequences on mental health, with numerous, varied psychiatric conditions sharing this developmental origin. Yet, the mechanisms linking adverse experiences to these conditions remain poorly understood. Here, we draw on a principled model of interval timing to propose that statistically-optimal adaptation of temporal representations to an unpredictable early life environment can produce key characteristics of anhedonia, a transdiagnostic symptom associated with affective disorders like depression and anxiety, and is also often co-morbid with forms of schizophrenia. The core observation is that early temporal unpredictability produces broader, more imprecise temporal expectations. As a result, reward sensitivity is blunted, reward anticipation is diminished, and associative learning is slowed. When agents with such representations are later introduced to more stable environments, they demonstrate a negativity bias, responding more to the omission of reward than its receipt. Increased encoding of negative events has been proposed to contribute to disorders commonly associated with anhedonia. We then examined how unpredictability interacts with another form of adversity, low reward availability, and found that unpredictability's effect was most strongly felt in richer environments, potentially leading to categorically different phenotypic expressions. In sum, our formalization suggests a single mechanism can help to link early life adversity to a range of behaviors associated with anhedonia, and offers novel insights into the interactive impacts of multiple adversities which often co-occur.

II. INTRODUCTION

Across development, brain circuits adapt to reflect the environment's structure, preferentially encoding more frequent aspects of the world. The statistics of the early life environment tune sensory receptive fields, producing non-homogeneous sensitivity to perceptual stimuli and determining discrimination abilities in adulthood (Tanaka, Ribot, Imamura, & Tani, 2006; Efrati & Gutfreund, 2011). Early consistency in these sensory inputs are crucial for the future functionality of involved circuits (Li, Fitzpatrick, & White, 2006). Similar developmental processes may take place in reward and memory systems, those underlying associative learning, implying that the consistency or predictability of associations in early life may shape the acquisition of associations later on (Birnie et al., 2020).

Caregivers are primary contributors to the associative structure infants encounter. Associations may take the form of an infant performing an action to which a caregiver consistently responds, allowing the infant to accurately anticipate their response in the future. Caregiver interactions will vary in their valence and predictability. Prior work has largely focused on the effect of valence on later child mental health outcomes (Sroufe, 2005; NICHD Early Care Research Network, 2006; Belsky & Fearon, 2002; Hane, Henderson, Reeb-Sutherland, & Fox, 2010). However, recent work has highlighted how early life unpredictability, or ELU, may also contribute (Baram et al., 2012). Research done in animals has illustrated that offspring exposed to unpredictable caregiver signals show a reduction in both motivation and the experience of pleasure, characteristics of the trans-diagnostic symptom anhedonia (Bolton et al., 2018). Work in humans accords with these findings, showing a relationship between experiences of early life unpredictability and symptom severity in anhedonia, depression, and anxiety (Glynn et al., 2019; Spadoni et al., 2022).

Here, we propose that the study of early-life unpredictability can be understood in part via its influence on the development of temporal receptive fields (TRFs), which are thought to serve as basis sets for associative learning more generally (Jin, Fujii, & Graybiel, 2009; Howard et al., 2014). TRFs capture the intuition that the strength of learned associations is dependent on the time between events (Balsam, Drew, & Gallistel, 2010). These tuning curves are similar to those found in sensory areas, but rather than being tuned to visual angle or auditory pitch, are sensitive to the temporal duration between related events.

We specifically examine how early-life unpredictability can, via its influence on the adaptation of TRFs, result in an anhedonic phenotype. We extend a principled computational model of

interval timing (Ludvig, Sutton, & Kehoe, 2008) to simulate how enhanced volatility during an early period of heightened plasticity can, with minimal assumptions, affect later predictions of reward during maturity. With this model, we formally demonstrate that early unpredictability in timing, and adaptation of temporal receptive fields to this timing, can lead to several defining characteristics of anhedonia – including slowed reward learning, reduced motivation and sensitivity to increasing rewards, and a bias towards learning from negative events – in the absence of differences in the overall amount of reward. Our results reproduce empirical findings that poor mental health outcomes can emerge from unpredictability in early life experience even when controlling for overall reward availability (Glynn et al., 2019).

While we show that a singular type of adversity can alone produce an anhedonic phenotype, in the real world individuals are often subject to multiple adversities. Modeling the nature of these interactions and their combined effect on learning will be critical for characterizing the developmental trajectory of psychopathology. As a first step, we model how temporal unpredictability interacts with the availability of reward, or richness, of the environment to shape later learning and expectations of reward. Under the common cumulative risk approach to conceptualizing and measuring early life adversity (Felitti, 2002), these two adversities are assumed to have an additive effect on development: individuals facing both are predicted to have the most negative outcomes. Our model predicts that unpredictability always has a negative effect on reward learning, however, contrary to the cumulative risk prediction, this effect is most pronounced in richer environments. Both unpredictability and an abundance of rewards individually alter temporal representations to be more expansive or diffuse, producing the observed interaction. Our results highlight the potential value of computational psychiatric approaches to tackling the heterogeneity of early life adversity and making sense of its developmental consequences.

III. ISOLATING THE CONTRIBUTIONS OF ONE FORM OF ADVERSITY, UNPREDICTABILITY

A. Methods

During the initial phase (“critical period”), agents’ temporal receptive fields were allowed to adapt to the environment’s temporal statistics. Agents fell into one of two groups, early life unpredictability (ELU) and control, differentiated by the variance of the distribution their reward timings were sampled from. The ELU group’s distribution had the higher variance. In the second phase (“post-critical period”), rewards were delivered at the same time point on each reinforced

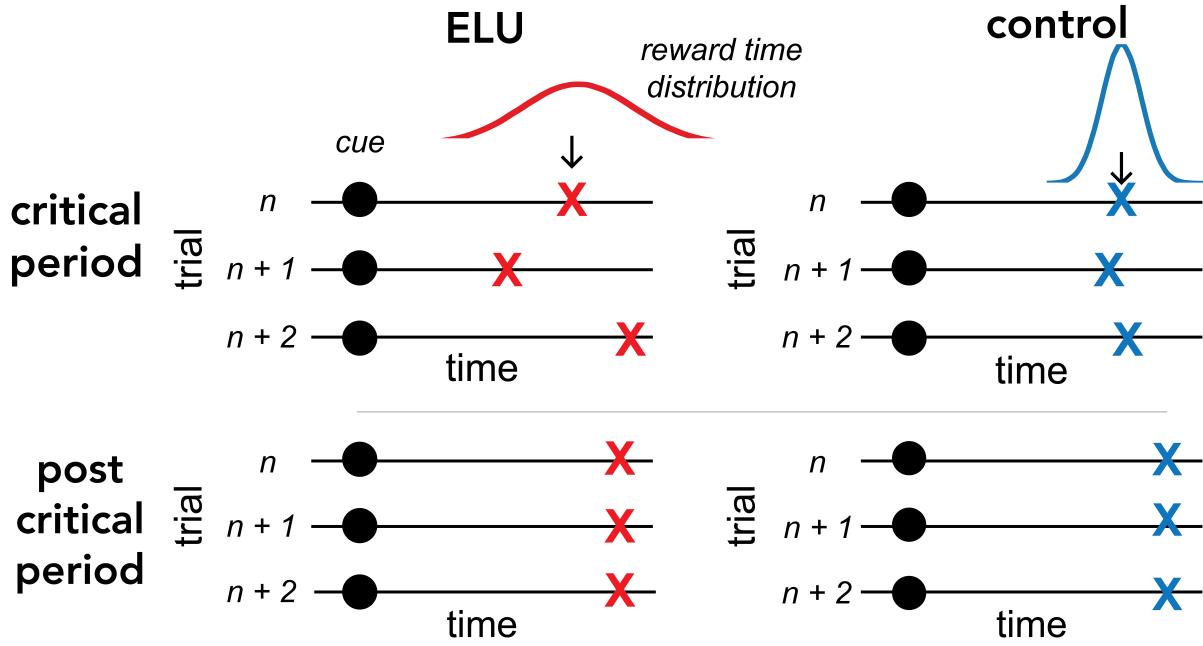


Fig. 1. Simulated agents learned to associate a cue and reward across two environments. The cue was partially reinforced in both environments — 75% of the time in the first and 50% in the second. On reinforced trials, the time point of reward delivery varied from trial to trial. During the initial phase (“critical period”), agents’ temporal receptive fields were incrementally adjusted by experience with the environment’s temporal statistics. Agents fell into one of two groups, early life unpredictability (ELU) and control, differentiated by the variance of the reward timing distribution, with the ELU group experiencing a more variable-timing environment. Critically, rewards were delivered according to the same probabilistic schedule for both groups. In the second phase (“post-critical period”), agents’ temporal receptive field parameters were frozen, and on each reinforced trial, rewards were delivered at the same time point for both groups.

trial for both groups, and critically, agents’ temporal receptive fields were no longer allowed to adapt to the test environment’s statistics.

1) *The Temporal-Difference Learning model:* Temporal-Difference (TD) models aim to accurately estimate the value of world states, V , in terms of the future rewards they predict. Time is explicitly represented in these models with each time step identifying a world state.

$$V^* = E \left[\sum_{k=1}^{\infty} \gamma^{k-1} r_{t+k} \right] \quad (1)$$

where r_t is the reward received at the current time step, and γ is a parameter controlling how heavily future rewards are discounted. Future rewards are less influential on the estimation of V when γ is low. A TD agent learns V via an error driven learning rule — the difference, δ_t ,

between the reward that was predicted (V_{t-1}) and what was actually received ($r_t + \gamma V_t$) is used to update the estimate of V at the next time step.

$$\delta_t = r_t + \gamma V_t - V_{t-1} \quad (2)$$

2) Microstimulus representation of time: All TD models explicitly represent time, but do so in various ways. Basic TD models use a complete-serial-compound (CSC) representation in which each time step is treated as independent from one another and agents are assumed to have perfect knowledge of when events occur. This representation prohibits temporal generalization, creating issues in environments where the time between cue and reward varies. The microstimulus representation addresses this problem by relaxing its temporal markers (Ludvig et al., 2008). CSC's discrete markers are replaced with continuous "microstimuli" which allow for temporal uncertainty to be represented. A stimulus, whether it be neutral, rewarding, or aversive is assumed to leave behind a memory trace that decays with time. The trace is represented by a basis set of overlapping temporal receptive fields — Gaussian distributions whose peak (μ) and standard deviations (σ) increase with the time after onset of the initial stimulus (y).

$$f(y, \mu, \sigma) = \frac{1}{\sqrt{2\pi}} e^{(-\frac{(y-\mu)^2}{2\sigma^2})} \quad (3)$$

A time step's value, V_t , is estimated as the weighted average of the microstimuli.

$$V_t = w_t^T x_t = \sum_{i=1}^n w_t(i)x_t(i) \quad (4)$$

This value is compared to the reward received. The error term, δ_t , adjusts the weights on the microstimuli, consequently updating the predicted value at the next time step.

$$w_{t+1} = w_t + \alpha \delta_t e_t \quad (5)$$

α is the learning rate controlling the time window over which trial to trial experiences are integrated. e_t is a vector containing each stimulus's eligibility traces.

$$e_t = \gamma \lambda e_t + x_t \quad (6)$$

Following the stimulus, its eligibility trace decays at a rate determined by γ and λ . γ is a temporal discounting factor as it was for the TD model using a CSC representation, while

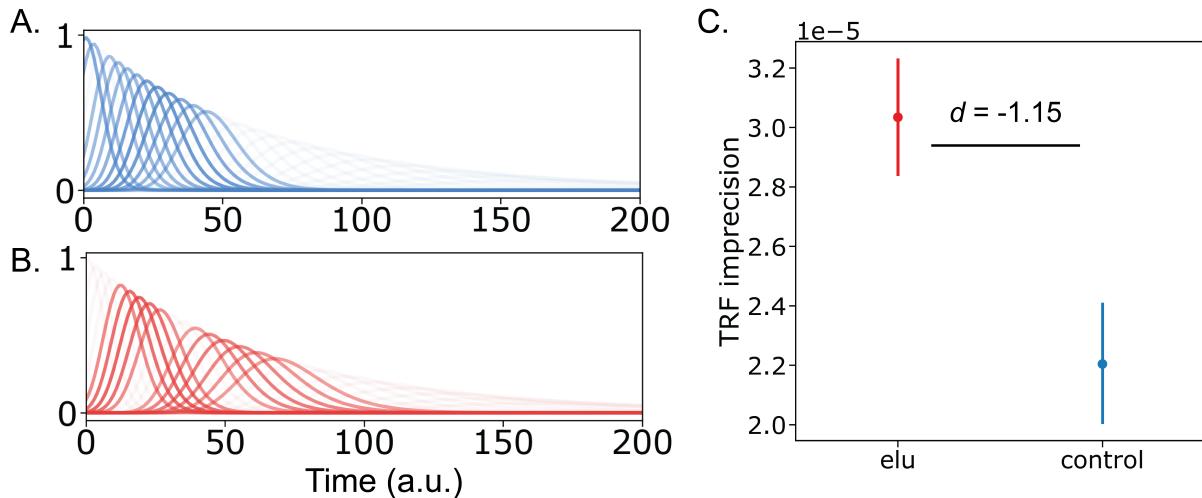


Fig. 2. Positively weighted temporal receptive fields. **A-B.** The ELU group developed a greater reliance on more imprecise temporal receptive fields, a consequence of their more volatile experience during the critical period. **C.** To summarize the difference between groups receptive fields, we computed the temporal receptive field (TRF) imprecision by taking a weighted average of the standard deviations of the positively weighted temporal receptive fields following the critical period. The ELU group relied on more broadly tuned, less precise temporal receptive fields relative to the control group.

λ controls the time window over which a stimulus can induce learning within a trial. For all simulations, we use the parameter settings from Ludvig et al, 2008 — $\alpha = 0.01$, $\gamma = 0.98$, $\lambda = 0.95$, $n = 50$, and $\sigma = 0.08$.

3) Simulating development: To model developmental changes in learning, we limit the period over which microstimuli weights can adapt to experience. We treat this as a critical period during which the temporal receptive fields are tuned to support accurate estimation of V . This adaptation process is designed to mimic the observed tuning of sensory receptive fields during analogous sensitive periods of development (Simoncelli & Olshausen, 2001).

We simulated two groups of agents learning cue-reward pairings across two phases (Figure 1). One group of agents, the early life unpredictability or ELU group, experienced a volatile environment in the first phase, meaning that the delay between cue and reward could vary considerably across the entirety of the phase's trials. The other group of agents, the control group, experienced relatively much less variation during this phase.

On each of the 1000 simulated trials, a cue was always presented at 10 timesteps and there was a 75% probability of a reward following it. If a cue was reinforced on a trial, the timing of reward delivery was sampled from a normal distribution with μ set to 30 timesteps and truncated

at 10 and 70 timesteps. σ varied between agents. For agents in the ELU group, σ was sampled from a zero-truncated normal distribution with $\mu_{hyper,elu} = 10$ and $\sigma_{hyper,elu} = 3$. The control group experienced much less variability, with σ being sampled from a zero-truncated normal distribution with $\mu_{hyper,control} = 1$, $\sigma_{hyper,control} = 2$.

In the second phase, the TRF weights were frozen, allowing us to directly examine the influence of highly variable early-life experiences. The temporal statistics of these environments differed from the prior phase's in two ways: 1. The reward was delivered at the same time step every trial for both groups of agents. 2. This time step was later (50 timesteps) than the mean time of reward in the initial phase (30 timesteps). By testing ELU agents' learning in novel environments that are more stable than the environment they "developed" in, we formalize the Mismatch Hypothesis of Early Life Adversity and Depression (Schmidt, 2011). Under this hypothesis, depression and other mental illnesses are proposed to be the byproduct of a mismatch between the developmental environment to which neural systems are optimized for and the later adulthood environment. We were particularly interested in characterizing how an agent's early adaptation to unpredictability would later affect uncertainty's influence on their expectations of reward. Within a probabilistic Pavlovian conditioning task, as agents were simulated, uncertainty should rise once the mean time of reward delivery has passed and no reward has been delivered. It is unclear whether the reward is simply late or is being omitted altogether on the trial. Thus, we moved back the time point of reward delivery in the novel environments to examine how the ELU and control groups differ in their response to reward and its omission following this uncertainty. These environments varied in the probability of reward on each trial (25, 55, 75, 95%). This allowed us to test the model's ability capture another characteristic of anhedonia — desensitization to increasing rewards (Sherdell, Waugh, & Gotlib, 2012; Yang et al., 2014). All agents completed 1000 trials. On each reinforced trial, the cue arrived at 10 timesteps and the reward at 50 timesteps.

4) Statistical Analyses: Each simulated agent encountered a different sequence of reward timings during the initial "critical period phase". Thus, a potential concern is that our results are largely driven by a subset of simulated agents. To assess the reliability a the relationship between prediction error magnitude and unpredictable experience, we performed a bootstrap analysis across agents within a group (Kim, Lewis-Peacock, Norman, & Turk-Browne, 2014; Bornstein et al., 2023). For each group, we sampled agents with replacement from that group until we reached the total number of agents in that group (100). We then computed the test

statistic for a two sample t-test with the selected groups. We repeated this procedure 1,000 times to obtain a distribution of test statistics across shuffled permutations of the simulated groups. This re-sampling procedure provides a p-value that is the fraction of test statistic values with a different sign from the base effect size (the test statistic for the original two groups). We also computed the Cohen's d in order to evaluate the size of the difference between simulated populations. By convention, effect sizes greater than 0.80 are considered "Large", and thus reliable (Cohen, 1992).

B. Results

1) Critical Period: First, we validated that the initial environment shaped the tuning of temporal receptive fields by comparing the groups' microstimuli weights following the critical period. For each agent, we computed a temporal precision measure by taking a weighted average of the microstimuli's standard deviations. Consistent with our predictions about the influence of environment on temporal representations in these simulated agents, we found that the ELU group relied on more broadly-tuned receptive fields relative to the controls (Figure 2; $t(198) = -7.83$, $p < .001$).

Early life unpredictability has been shown to slow learning from reinforcement (Birn, Roeber, & Pollak, 2017; Dillon et al., 2009). Next, we examined the model's ability to capture this finding. As a proxy for learning, we used prediction error magnitude. The more strongly a cue has been learned to predict reward, the smaller the prediction error should be when reward is actually delivered and the larger the prediction error should be when reward is omitted. To compare prediction errors between groups, we computed, across time within each trial, the prediction error extremum for each agent (Figure 3). On reinforced trials, the maximum prediction error magnitude following the cue was taken and on omission trials, the minimum was taken. We then averaged this value, across trials, within each participant, and across participants within each group. We found that the ELU group's prediction errors were more positive relative to controls on reinforced trials (Figure 4, $t(198) = -12.95$, $p < .001$, Cohen's $d = -1.83$) but more negative on omission trials ($t(198) = -8.47$, $p < .001$, Cohen's $d = -1.20$). Despite experiencing the same amount of reward on average as the control group, the ELU group showed slower learning under reinforcement. This demonstrates that impaired reward learning, as observed in anhedonia, can emerge from experienced temporal volatility alone during a period of plasticity.

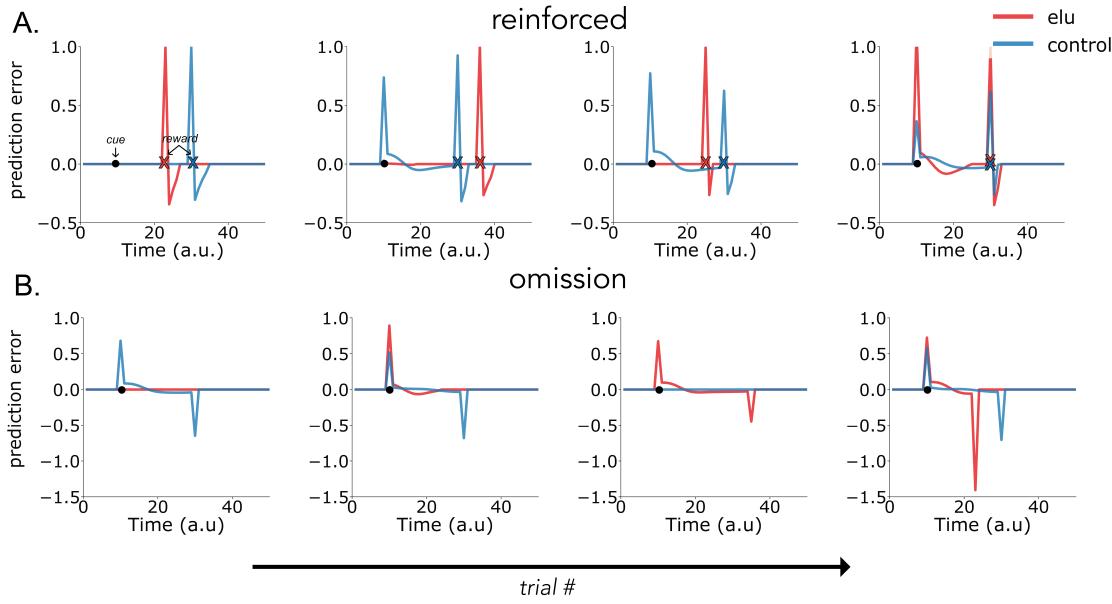


Fig. 3. Critical period results - prediction error δ , across time on reinforced and non-reinforced/omitted trials. **A. Reinforced trials.** The time point at which the ELU agent's prediction error signal peaks varies as a consequence of the reward delivery time itself varying from trial to trial. Additionally, the magnitude of their prediction error does not diminish in these early trials. In contrast, the control group consistently experienced a large prediction error near 300 ms. This consistency expedites the prediction error's transition from occurring at the time of the reward to the time of the cue. **B. Omission trials.** The control group, even very early on, experiences a negative prediction when reward is omitted. The ELU group, however, does not, as the unpredictable reward timing frustrates their ability to form strong expectations.

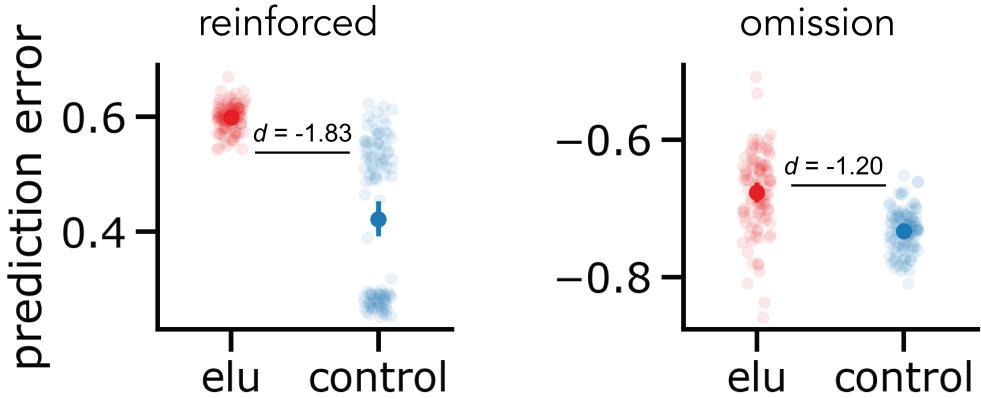


Fig. 4. Critical period results - mean prediction error. For each trial, following the cue, the time point at which prediction error was at its peak magnitude was taken. For each agent, an aggregate measure was computed by taking the median over the trials' extrema. The ELU group showed larger prediction errors on trials in which reward followed the cue but weaker prediction errors when reward was omitted following the cue. Both suggest slowed reward learning in the ELU group.

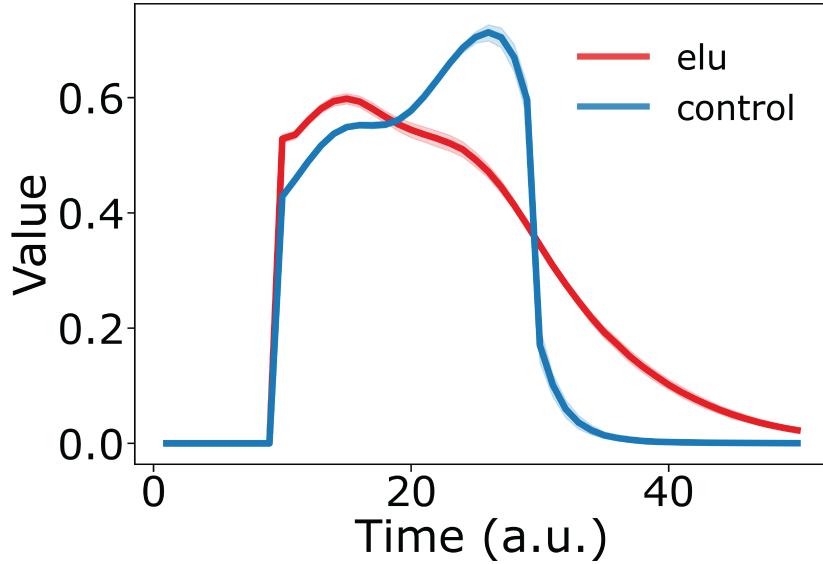


Fig. 5. Critical period results - value. V , at each time step averaged across trials. The ELU group's value signal decreased following the cue while the control group's increased. Once the mean time of reward was reached, the ELUs' value signal steadily and slowly dropped off while the controls' did so much more quickly.

Early life unpredictability has also been shown to impair motivation (Hanson, Williams, Bangasser, & Peña, 2021), potentially stemming from a reduced expectation of reward. Thus, we next compared the groups' expectation of value across time following the cue. The ELU group's value signal peaked early following the cue (Figure 5, mean = 15.43, $sd = 3.15$) and slowly decayed (mean = 49.87, $sd = 1.29$). We interpret this as ELU agents gradually growing less confident reward will come at all if not received immediately following the cue. In contrast, control agents' appear to increasingly anticipate the reward as the expected arrival time approaches. Control agents' value signal, in comparison, peaked much later (mean = 26.09, $sd = 2.20$, $t(198) = 27.53$, $p < .001$, Cohen's $d = 3.89$) and quickly dropped off, reaching its minimum soon after the average time of reward (mean = 37.80, $sd = 8.43$, $t(198) = -14.08$, $p < .001$, Cohen's $d = -1.99$).

2) *Post Critical Period*: During the second phase, all agents received reward at the same time on every reinforced trial. Critically, we closed the “critical period”, no longer allowing the weights on the temporal receptive fields to adapt to the novel environment. In the previous simulations, we demonstrated how early life unpredictability could slow agents' learning from rewards. We next examined whether early life unpredictability would desensitize agents to increasing rewards,

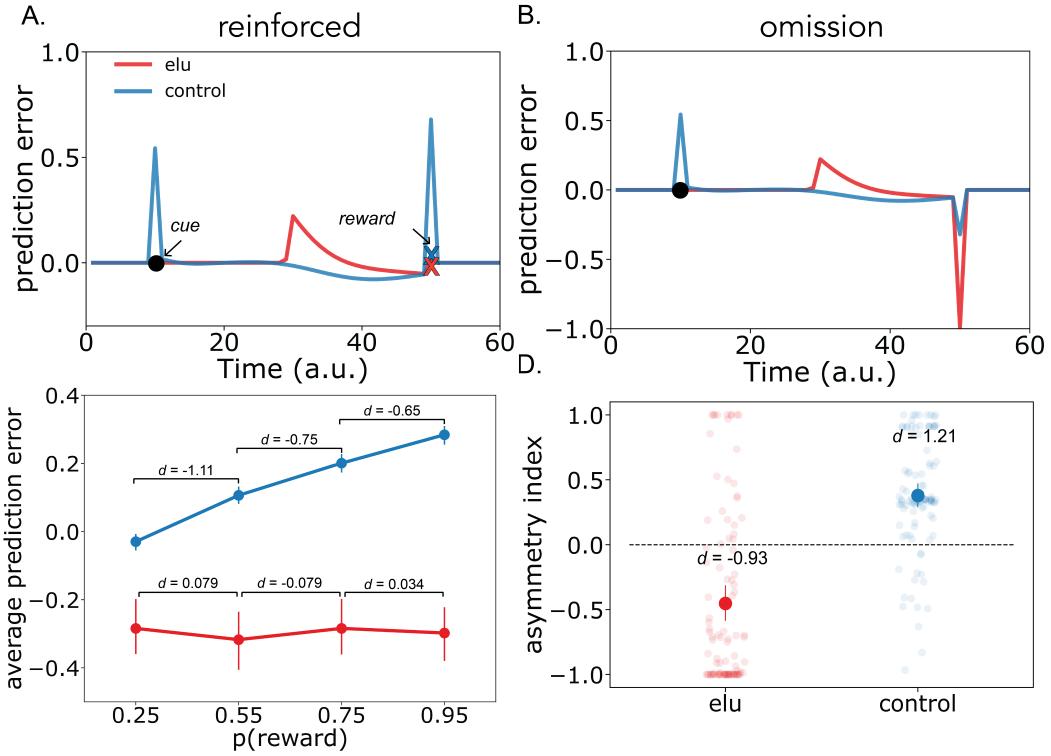


Fig. 6. A-B. Prediction errors at each time step for reinforced and omitted trials. A. On reinforced trials, control agents' relative to ELU agents', prediction errors were larger and earlier at the times of cue and reward. **B.** On omission trials, control agents' prediction errors were larger at the time of the cue. However, at the time of reward, ELU agents' prediction errors were larger. **C.** Prediction error extrema averaged across all trials. Control agents' average prediction errors grew more positive as the richness (probability of reward) of the environment increased. In contrast, ELU agents' average prediction errors were insensitive to increasing richness. **D.** Asymmetry indices from the 55% probability of reward environment. The ELU group showed a negative event bias, experiencing more extreme prediction errors on omission trials relative to reinforced trials. In contrast, the control group showed a positive event bias, experiencing larger prediction errors on reinforced trials than omission. Error bars are 95% bootstrapped confidence intervals.

another characteristic of anhedonia. Averaging prediction errors across all trials, control agents had more positive prediction errors than ELU agents, and they showed a sensitivity to increasing rewards as their prediction error average increased with the richness of the environment ($\beta_{elu} = -0.15, p = .0074, \beta_{rich} = 0.45, p < .001, \beta_{elu*rich} = -0.46, p < .001$). Now, focusing on the magnitude of the prediction error trial to trial, ELU agents showed marginally lower positive prediction errors on reinforced trials relative to controls but more strongly negative prediction errors on omission trials (reinforced - $t(198) = 2.98, p < .001$, Cohen's $d = 0.42$; omission - $t(198) = 13.29, p < .001$, Cohen's $d = 1.88$). The greater the magnitude of the prediction

error, the more the agent learns. Because valence asymmetries in learning have been proposed to be clinically relevant (Rouhani, Norman, Niv, & Bornstein, 2020; Pike & Robinson, 2022), we next wanted to compare the magnitude of prediction errors on reinforced and omission trials for each agent to check for such asymmetries. We computed an asymmetry index for each agent as follows:

$$\text{index} = \frac{|PE_+| - |PE_-|}{|PE_+| + |PE_-|} \quad (7)$$

Because the asymmetry index is computed by separately averaging over positive and negative prediction errors, the index is not sensitive to the richness of the second phase environment. Thus, we only report indices from the 55% probability of reward environment. ELU agents' asymmetry indices were overall negative (Figure 6D; $t(199) = -6.60$, $p < .001$, Cohen's $d = -0.93$) while the control agents' were positive ($t(199) = 8.56$, $p < .001$, Cohen's $d = 1.21$). Because prediction error magnitude enhances learning and memory, this suggests that negative events would have an outsized influence on ELU agents, making their value estimates overly pessimistic while control agents' overly optimistic (Sharot, 2011). Our model provides a mechanism through which both of these biases could emerge under minimal assumptions.

IV. INTERACTIONS BETWEEN MULTIPLE FORMS OF ADVERSITY

A. Methods

1) *Critical Period*: To examine the interaction between multiple forms of early life adversity — temporal unpredictability and reward availability, we additionally manipulated the richness of the critical period environment and observed its effect on both groups' reward learning. This allowed us to test the assumptions of the cumulative risk conceptualization of early life adversity which assumes an additive effect of adversities on developmental outcomes. We simulated groups of the ELU and control agents in environments with 25, 55, 75, and 95% probability of reward. As in previous simulations, the time of reward delivery was sampled from a normal distribution with $\mu = 30$ timesteps and truncated at 10 and 70 timesteps, and the distribution's σ differed between groups — ELU agents' σ were sampled from a zero-truncated normal distribution with $\mu_{\text{hyper},\text{elu}} = 10$ and $\sigma_{\text{hyper},\text{elu}} = 3$ and controls' were sampled from a zero-truncated normal distribution with $\mu_{\text{hyper},\text{control}} = 1$, $\sigma_{\text{hyper},\text{control}} = 2$.

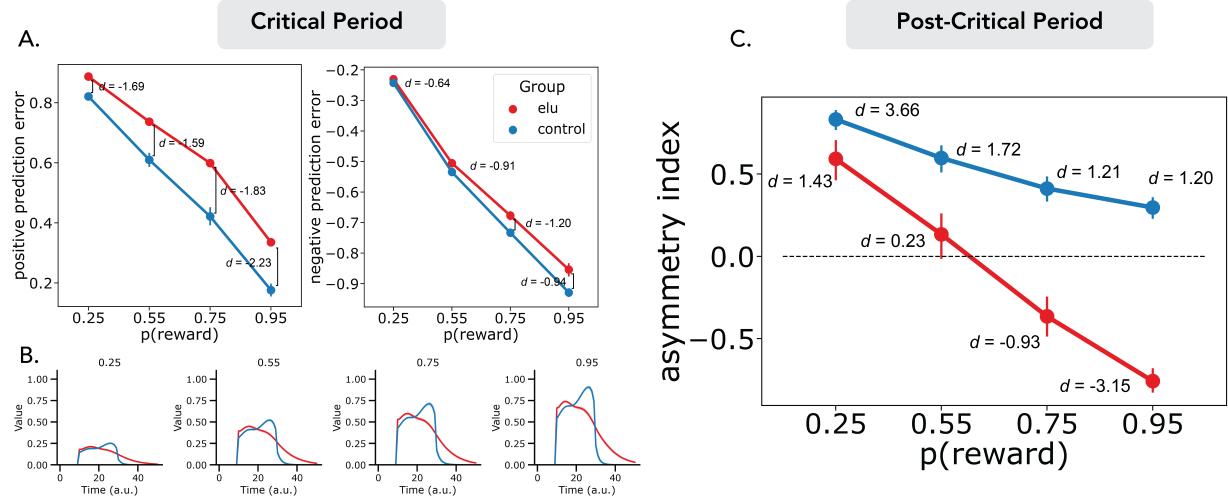


Fig. 7. Varying critical period environment richness to examine the impact of multiple adversities. **A.** Critical Period Positive Prediction Errors. The magnitude of positive prediction errors were altered both by reward availability (probability of reward) and temporal unpredictability (ELU vs. control). Greater unpredictability produced larger prediction errors (implying slowed reward learning) while greater reward availability produced smaller prediction errors (suggesting faster learning). These two dimensions of adversity mitigated one another — unpredictability and availability interacted such that group differences were reduced in the poorer environments. This is contrary to what would be predicted under the cumulative risk conceptualization of adversity. **B.** Critical Period Negative Prediction Errors. Low reward availability and temporal unpredictability's effects on negative prediction error magnitude similarly suggested slowed learning. Negative prediction errors were more negative in richer environments and for control agents. Again, poorer environments mitigated the differences between groups. **C.** Critical Period Value Signal. Environment richness increased agents' expectation of reward as reflected in their heightened value signal. In the richest environment, group differences at various time points were the greatest. These differences diminished as richness diminished. **D.** Post-Critical Period Asymmetry Indices. Control agents demonstrated a positivity bias across all environments while ELU agents only showed a positivity bias in the poorest environments but a negativity bias in the richest. Error bars are 95% confidence intervals, computed via permutation test.

2) *Post Critical Period:* In the novel environment agents encountered following the critical period, there was a 55% chance of reward on each trial. Within a trial, the cue was presented at 10 timesteps and if it was a reinforced trial, reward was delivered at the same time step (50).

B. Results

1) *Critical Period:* Again, we interpret less positive prediction errors on reinforced trials and more negative prediction errors on omission trials as indicators of better learning. Under this interpretation, both unpredictability and low reward availability were found to slow reward learning. On reinforced trials, prediction errors were overall more positive for the ELU group and

both groups' prediction errors became less positive with environment richness (Figure 7A, $\beta_{elu} = 0.040, p = .011, \beta_{rich} = -0.91, p < .001$). These two dimensions interacted, with the difference between groups increasing as environment richness increased ($\beta_{elu*rich} = 0.15, p < .001$). We observed a similar pattern on omission trials. Prediction errors were more negative with increasing environment richness, again suggesting slower learning (Figure 7B, $\beta_{rich} = -0.98, p < .001$). The effects of unpredictability on learning were only observed in richer environments, with no main effect of group but an interaction effect between group and richness ($\beta_{elu} = -0.015, p = .11, \beta_{elu*rich} = 0.092, p < .001$). Taken together, our results reveal that the effect of reward unpredictability is most fully felt when reward is abundant, a consequence of both dimensions increasing the imprecision of temporal representations ($\beta_{elu} = 9.56e - 06, p < .0001, \beta_{rich} = 1.26e - 05, p < .001, \beta_{elu*rich} = -1.76e - 07, p = .95$). When rewards are both unpredictably timed and sparse, it reduces the range of timings an agent's representation must accommodate and particularly, early or late rewards can be attributed as noise rather than a defining feature of a volatile environment.

The value signal reveals a similar impact of temporal unpredictability and low reward availability on learning. The value signal overall strengthens with increasing reward availability (Figure 7C, $\beta_{rich} = 0.31, p < .001$). At the same time, only when the environment is sufficiently rich can unpredictability exert its blunting effect on the signal ($\beta_{elu} = 0.0027, p = .66, \beta_{elu*rich} = 0.027, p = .0033$).

2) *Post Critical Period:* In the post critical period phase, we observed the same complex relationship between reward unpredictability and availability, in which greater reward availability allows for unpredictability to have a greater influence. Across all environments, control agents maintained a bias towards learning from positive events as suggested by positive asymmetry indices (Figure 7D, $25\% - t(99) = 29.91, p < .001$, Cohen's $d = 3.66$; $55\% - t(99) = 12.14, p < .001$, Cohen's $d = 1.72$; $75\% - t(99) = 8.56, p < .001$, Cohen's $d = 1.21$; $95\% - t(99) = 8.46, p < .001$, Cohen's $d = 1.20$). The valence of ELU agents' biases, in contrast, was dependent on the richness of the developmental environment. ELU agents who experienced the sparsest rewards during the critical period exhibited a positivity bias, similar to control agents albeit weaker ($25\% - t(99) = 10.09, p < .001$ Cohen's $d = 1.43$), those who experienced a less sparse environment showed no bias ($55\% - t(99) = 1.62, p = 0.11$, Cohen's $d = 0.23$), and those who experienced an environment abundant with rewards exhibited a negativity bias ($75\% - t(99) = -6.60, p < .001$, Cohen's $d = -0.93$; $95\% - t(99) = -22.24, p < .001$,

Cohen's $d = -3.15$). This pattern of results are a byproduct of the reward expectations built up during the critical period. ELU agents whose representations are adapted for richer environments have a stronger prior expectation that reward will have a delayed arrival rather than being omitted altogether. Thus, when reward is omitted on a trial, they experience a particularly large negative prediction error. Our simulations contradict the predictions that would be made under the cumulative risk approach which assumes an additive effect of adversities.

V. DISCUSSION

Here, we propose a novel computational link between early life unpredictability and the emergence of anhedonia — the optimization of temporal representations to the early life environment. By simply assuming that temporal receptive fields are adapted to the statistics of the early life environment, several behaviors associated with anhedonia emerge — impaired learning from reinforcement, reduced anticipation of reward, insensitivity to increasing rewards, and a greater response to the omission of events.

These findings are consistent with behavioral outcomes observed in the laboratory and clinical settings. One representative set of such findings is of an asymmetric attentional bias in anhedonia. If we treat the omission of reward as a negatively valenced event and the presence of reward as a positive event, this suggests a negative attentional bias in the ELU group and positive bias in the controls, reproducing empirical findings (Dillon & Pizzagalli, 2018; Frank, 2004). Larger negative prediction errors may not only affect attention in the moment but also have longer lasting consequences via memory. Surprising events, like prediction errors, are known to be more easily retrieved from memory (Rouhani et al., 2020; Sinclair & Barense, 2018). This provides a mechanism by which singular negative events can have an outsized influence on expectations and consequently, shape mood over the longer term (Eldar, Rutledge, Dolan, & Niv, 2016). Frequent large negative prediction errors could produce the persistent negative mood that characterizes anhedonia (Dillon et al., 2009). We found that the development of this negativity bias was critically dependent on the overall richness of the environment. To experience a pronounced negative prediction error when reward was omitted, agents needed to have a strong expectation that reward would come but a weak expectation of when that would be. Only in environments rich with variously timed rewards did such expectations emerge.

Our results contradict the assumptions and predictions of the cumulative risk conceptualization of early life adversity (Felitti, 2002). The cumulative risk approach has been crucial in

establishing the robust association between negative events early in life and a wide array of negative outcomes later in development. However, aggregating over heterogeneous experiences may obscure the mechanisms linking such experiences to later psychopathology (Smith & Pollak, 2021; McLaughlin, Sheridan, Humphreys, Belsky, & Ellis, 2021). One proposed alternative are dimensional models which identify influential features of the early life environment on development and seek to characterize how these features exert their influence. Supporting the dimensional approach, recent work has found divergent associations between measures of threat and deprivation in the early life environment with later developmental outcomes including amygdala reactivity to threat, aversive learning, cognitive control, and pubertal timing (Lambert, King, Monahan, & McLaughlin, 2017; Machlin, Miller, Snyder, McLaughlin, & Sheridan, 2019; Miller, Machlin, McLaughlin, & Sheridan, 2021; Rosen et al., 2020; Sheridan, Peverill, Finn, & McLaughlin, 2017; Sumner, Colich, Uddin, Armstrong, & McLaughlin, 2019; Sun, Fang, Wan, Su, & Tao, 2020). However, adopters of these approaches have been criticized for an unprincipled choice of dimensions, particularly lacking neurobiological grounding (Smith & Pollak, 2021). Given the potential relevance of reward systems to psychopathology, it may be valuable to look at the statistical properties of the environment known to influence reward learning as potential candidate dimensions.

Thus far in our interpretation of the results, we've treated the cue-paired outcome as reward. However, the model is agnostic to the valence of the outcome — allowing for different interpretations where the outcome is treated as neutral or aversive. Different valences will suggest different behavioral phenotypes. Treating the outcome as aversive, like a shock, the ELU group's prolonged expectation of a negative outcome's appearance could be interpreted as sustained hypervigilance (perhaps akin to a form of “paranoia”), a symptom of anxiety. Treating the outcome as neutral, impairments in reward learning become more general impairments in relational learning. This may explain memory deficits and alterations in hippocampal structure in ELU individuals (Granger et al., 2021; Molet et al., 2016) and anhedonia's associated memory deficits. Prior work has suggested that anhedonia is characterized not only by the inability to experience pleasure in the moment but also the inability to recall past and anticipate future pleasurable experiences (Dillon & Pizzagalli, 2018).

Here, we've only considered the mechanism under Pavlovian learning conditions. However, it also suggests differences in ELU individuals' instrumental learning and action selection. The inability to accurately predict the timing of future outcomes diminishes an individual's perceived

controllability of the environment, which has also been implicated in psychiatric disorders such as anxiety (Bishop & Gagne, 2018).

Hidden-state inference models capture a similar idea as the microstimulus model at a different level of analysis (Starkweather, Babayan, Uchida, & Gershman, 2017). Often, the true state of the world is unknown or hidden and must be inferred from observations. This inference process is in part driven by prediction errors (Rouhani et al., 2020), and by extension is more difficult in volatile environments. As a result, ELU individuals may infer fewer states in the world (or, analogously, more states in an environment where negative prediction errors predominate) and group their experiences accordingly as a result of this early volatility. We have previously shown that this assumption of reduced sensitivity with a hidden-state inference model can produce reduced exploration in a foraging task (Harhen & Bornstein, 2021), a behavior found in ELU populations (Lloyd, McKay, & Furl, 2022), and may also explain why individuals who experience early life unpredictability are at higher risk of developing substance use disorders and relapsing following treatment (Harhen, Baram, Yassa, & Bornstein, 2021).

Our model is predicated on the assumption that prediction error learning can serve as a mechanism of environmental adaptation across multiple timescales — within a task and across development. Embodying an extreme form of sensitive period, adulthood is conceptualized as a period in which learning has altogether ceased. Future work could examine how more realistic, relaxed constraints on learning — in which developmental experience lays the groundwork for the architecture of neural systems which later adulthood experience can modify and reorganize (Galván, 2010; Karmiloff-Smith, 1994) — affects predicted learning outcomes. This particular assumption implies that the prior biases instilled by the developmental environment should have the greatest influence in few shot or one shot learning experiences. When current experience underdetermines what an agent should expect or do, past experience should largely influence the conclusion an agent reaches, with early life experience having a particularly privileged role (Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010). Such inductive biases will facilitate learning in environments aligned with these biases and frustrate it in misaligned environments. If the influence of the developmental environment on expectations and choice is greatest in environments in which the agent has limited experience, this has implications for when symptoms for disorders like anxiety and substance use disorder should worsen (Sharp, Miller, Dolan, & Eldar, 2020; Bornstein & Pickard, 2020).

Our results highlight the key role time plays in shaping reinforcement learning and conse-

quently its impact on behaviors associated with mental illness. The model's ability to produce varied phenotypes from the same computations suggests that the model's implication extend beyond anhedonia. Potentially it provides a common origin for a number of psychiatric disorders, which could explain their high co-morbidity rates (Jacobi et al., 2004; Kessler, Chiu, Demler, Merikangas, & Walters, 2005; Krueger, Chentsova-Dutton, Markon, Goldberg, & Ormel, 2003). Further research is needed to empirically test the model's behavioral predictions, namely, for early life unpredictability's impact on interval timing, and interval timing's relationship with psychiatric disorders. Finally, our results offer a demonstration of the value of computational modeling to understanding the development of psychopathology. By drawing on a reinforcement learning framework, we can formalize the changing relationship between the agent and their environment across development, produce testable predictions of how the environment shapes the latent computations underlying clinically relevant behaviors, like learning, and propose mechanistic links between altered computations and the later emergence of psychiatric symptoms.

VI. ACKNOWLEDGEMENTS

This work was supported by a seed grant under NIH P50MH096889 (PI: TZ Baram) and a NARSAD Young Investigator Award from the Brain and Behavior Research Foundation, both to AMB. NCH was supported by a National Defense Science and Engineering Graduate fellowship. The authors thank Elliot Ludvig for providing the microstimulus model code, and Tallie Z Baram, Meredith Doherty, Steven J Granger, Steven V Mahler, Sophia C Levis, and Michael A Yassa for helpful conversations.

REFERENCES

- Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and associative learning. *Comp. Cogn. Behav. Rev.*, 5, 1–22.
- Baram, T. Z., Davis, E. P., Obenau, A., Sandman, C. A., Small, S. L., Solodkin, A., & Stern, H. (2012, September). Fragmentation and unpredictability of early-life experience in mental disorders. *Am. J. Psychiatry*, 169(9), 907–915.
- Belsky, J., & Fearon, R. M. P. (2002, December). Early attachment security, subsequent maternal sensitivity, and later child development: does continuity in development depend upon continuity of caregiving? *Attach. Hum. Dev.*, 4(3), 361–387.

- Birn, R. M., Roeber, B. J., & Pollak, S. D. (2017, December). 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, May). Plasticity of the reward circuitry after Early-Life adversity: Mechanisms and significance. *Biol. Psychiatry*, *87*(10), 875–884.
- Bishop, S. J., & Gagne, C. (2018, July). Anxiety, depression, and decision making: A computational perspective. *Annu. Rev. Neurosci.*, *41*, 371–388.
- Bolton, J. L., Molet, J., Regev, L., Chen, Y., Rismanchi, N., Haddad, E., ... Baram, T. Z. (2018, January). Anhedonia following Early-Life adversity involves aberrant interaction of reward and anxiety circuits and is reversed by partial silencing of amygdala Corticotropin-Releasing hormone gene. *Biol. Psychiatry*, *83*(2), 137–147.
- Bornstein, A. M., Aly, M., Feng, S. F., Turk-Browne, N. B., Norman, K. A., & Cohen, J. D. (2023, September). *Associative memory retrieval modulates upcoming perceptual decisions*.
- Bornstein, A. M., & Pickard, H. (2020, May). “chasing the first high”: memory sampling in drug choice. *Neuropsychopharmacology*, *45*(6), 907–915.
- Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*(3), 98–101.
- Dillon, D. G., Holmes, A. J., Birk, J. L., Brooks, N., Lyons-Ruth, K., & Pizzagalli, D. A. (2009, August). Childhood adversity is associated with left basal ganglia dysfunction during reward anticipation in adulthood. *Biol. Psychiatry*, *66*(3), 206–213.
- Dillon, D. G., & Pizzagalli, D. A. (2018, March). Mechanisms of memory disruption in depression. *Trends Neurosci.*, *41*(3), 137–149.
- Efrati, A., & Gutfreund, Y. (2011, May). Early life exposure to noise alters the representation of auditory localization cues in the auditory space map of the barn owl. *J. Neurophysiol.*, *105*(5), 2522–2535.
- Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016, January). Mood as representation of momentum. *Trends Cogn. Sci.*, *20*(1), 15–24.
- Felitti, V. J. (2002). The relation between adverse childhood experiences and adult health: Turning gold into lead. *Perm. J.*, *6*(1), 44–47.
- Frank, M. J. (2004). *Dynamic dopamine modulation of striato-cortical circuits in cognition: Converging neuropsychological, psychopharmacological and computational studies* Unpublished

- doctoral dissertation. Ann Arbor, United StatesUniversity of Colorado at Boulder.
- Galván, A. (2010, June). Neural plasticity of development and learning. *Hum. Brain Mapp.*, *31*(6), 879–890.
- Glynn, L. M., Stern, H. S., Howland, M. A., Risbrough, V. B., Baker, D. G., Nievergelt, C. M., ... Davis, E. P. (2019, April). Measuring novel antecedents of mental illness: the questionnaire of unpredictability in childhood. *Neuropsychopharmacology*, *44*(5), 876–882.
- Granger, S. J., Glynn, L. M., Sandman, C. A., Small, S. L., Obenhaus, A., Keator, D. B., ... Davis, E. P. (2021, February). Aberrant maturation of the uncinate fasciculus follows exposure to unpredictable patterns of maternal signals. *J. Neurosci.*, *41*(6), 1242–1250.
- Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010, August). Probabilistic models of cognition: exploring representations and inductive biases. *Trends Cogn. Sci.*, *14*(8), 357–364.
- Hane, A. A., Henderson, H. A., Reeb-Sutherland, B. C., & Fox, N. A. (2010, September). Ordinary variations in human maternal caregiving in infancy and biobehavioral development in early childhood: A follow-up study. *Dev. Psychobiol.*, *52*(6), 558–567.
- Hanson, J. L., Williams, A. V., Bangasser, D. A., & Peña, C. J. (2021). Impact of early life stress on reward circuit function and regulation. *Front. Psychiatry*, *12*, 1799.
- Harhen, N., Baram, T. Z., Yassa, M. A., & Bornstein, A. M. (2021, May). Formalizing the relationship between early life adversity and addiction vulnerability: The role of memory sampling. *Biol. Psychiatry*, *89*(9), S189.
- Harhen, N., & Bornstein, A. M. (2021). Structure learning as a mechanism of overharvesting. *Proceedings of the 19th International Conference on Cognitive Modeling*.
- Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014, March). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *J. Neurosci.*, *34*(13), 4692–4707.
- Jacobi, F., Wittchen, H.-U., Holting, C., Höfler, M., Pfister, H., Müller, N., & Lieb, R. (2004, May). Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the german health interview and examination survey (GHS). *Psychol. Med.*, *34*(4), 597–611.
- Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009, November). Neural representation of time in cortico-basal ganglia circuits. *Proc. Natl. Acad. Sci. U. S. A.*, *106*(45), 19156–19161.
- Karmiloff-Smith, B. A. (1994, January). Beyond modularity: A developmental perspective on

- cognitive science. *Eur. J. Disord. Commun.*, 29(1), 95–105.
- Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005, June). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Arch. Gen. Psychiatry*, 62(6), 617–627.
- Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014, June). Pruning of memories by context-based prediction error. *Proc. Natl. Acad. Sci. U. S. A.*, 111(24), 8997–9002.
- Krueger, R. F., Chentsova-Dutton, Y. E., Markon, K. E., Goldberg, D., & Ormel, J. (2003, August). A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *J. Abnorm. Psychol.*, 112(3), 437–447.
- Lambert, H. K., King, K. M., Monahan, K. C., & McLaughlin, K. A. (2017, August). Differential associations of threat and deprivation with emotion regulation and cognitive control in adolescence. *Dev. Psychopathol.*, 29(3), 929–940.
- Li, Y., Fitzpatrick, D., & White, L. E. (2006, May). The development of direction selectivity in ferret visual cortex requires early visual experience. *Nat. Neurosci.*, 9(5), 676–681.
- Lloyd, A., McKay, R. T., & Furl, N. (2022, January). Individuals with adverse childhood experiences explore less and underweight reward feedback. *Proc. Natl. Acad. Sci. U. S. A.*, 119(4).
- Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2008, December). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural Comput.*, 20(12), 3034–3054.
- Machlin, L., Miller, A. B., Snyder, J., McLaughlin, K. A., & Sheridan, M. A. (2019, May). Differential associations of deprivation and threat with cognitive control and fear conditioning in early childhood. *Front. Behav. Neurosci.*, 13, 80.
- McLaughlin, K. A., Sheridan, M. A., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2021, November). The value of dimensional models of early experience: Thinking clearly about concepts and categories. *Perspect. Psychol. Sci.*, 16(6), 1463–1472.
- Miller, A. B., Machlin, L., McLaughlin, K. A., & Sheridan, M. A. (2021, April). Deprivation and psychopathology in the fragile families study: A 15-year longitudinal investigation. *J. Child Psychol. Psychiatry*, 62(4), 382–391.
- Molet, J., Maras, P. M., Kinney-Lang, E., Harris, N. G., Rashid, F., Ivy, A. S., ... Baram, T. Z. (2016, December). MRI uncovers disrupted hippocampal microstructure that underlies memory

- impairments after early-life adversity. *Hippocampus*, 26(12), 1618–1632.
- NICHD Early Care Research Network. (2006, January). Infant-mother attachment classification: risk and protection in relation to changing maternal caregiving quality. *Dev. Psychol.*, 42(1), 38–58.
- Pike, A. C., & Robinson, O. J. (2022, April). Reinforcement learning in patients with mood and anxiety disorders vs control individuals: A systematic review and meta-analysis. *JAMA Psychiatry*, 79(4), 313–322.
- Rosen, M. L., Hagen, M. P., Lurie, L. A., Miles, Z. E., Sheridan, M. A., Meltzoff, A. N., & McLaughlin, K. A. (2020, July). Cognitive stimulation as a mechanism linking socioeconomic status with executive function: A longitudinal investigation. *Child Dev.*, 91(4), e762–e779.
- Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020, October). Reward prediction errors create event boundaries in memory. *Cognition*, 203, 104269.
- Schmidt, M. V. (2011, April). Animal models for depression and the mismatch hypothesis of disease. *Psychoneuroendocrinology*, 36(3), 330–338.
- Sharot, T. (2011, December). The optimism bias. *Curr. Biol.*, 21(23), R941–5.
- Sharp, P. B., Miller, G. A., Dolan, R. J., & Eldar, E. (2020, September). Towards formal models of psychopathological traits that explain symptom trajectories. *BMC Med.*, 18(1), 264.
- Sherdell, L., Waugh, C. E., & Gotlib, I. H. (2012, February). Anticipatory pleasure predicts motivation for reward in major depression. *J. Abnorm. Psychol.*, 121(1), 51–60.
- Sheridan, M. A., Peverill, M., Finn, A. S., & McLaughlin, K. A. (2017, December). Dimensions of childhood adversity have distinct associations with neural systems underlying executive functioning. *Dev. Psychopathol.*, 29(5), 1777–1794.
- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annu. Rev. Neurosci.*, 24, 1193–1216.
- Sinclair, A. H., & Barense, M. D. (2018, August). Surprise and destabilize: prediction error influences episodic memory reconsolidation. *Learn. Mem.*, 25(8), 369–381.
- Smith, K. E., & Pollak, S. D. (2021, January). Rethinking concepts and categories for understanding the neurodevelopmental effects of childhood adversity. *Perspect. Psychol. Sci.*, 16(1), 67–93.
- Spadoni, A. D., Vinograd, M., Cuccuruzzu, B., Torres, K., Glynn, L. M., Davis, E. P., ... Risbrough, V. B. (2022, October). Contribution of early-life unpredictability to neuropsychiatric symptom patterns in adulthood. *Depress. Anxiety*, 39(10-11), 706–717.

- Sroufe, L. A. (2005, December). Attachment and development: a prospective, longitudinal study from birth to adulthood. *Attach. Hum. Dev.*, 7(4), 349–367.
- Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017, April). Dopamine reward prediction errors reflect hidden-state inference across time. *Nat. Neurosci.*, 20(4), 581–589.
- Sumner, J. A., Colich, N. L., Uddin, M., Armstrong, D., & McLaughlin, K. A. (2019, February). Early experiences of threat, but not deprivation, are associated with accelerated biological aging in children and adolescents. *Biol. Psychiatry*, 85(3), 268–278.
- Sun, Y., Fang, J., Wan, Y., Su, P., & Tao, F. (2020, September). Association of Early-Life adversity with measures of accelerated biological aging among children in china. *JAMA Netw Open*, 3(9), e2013588.
- Tanaka, S., Ribot, J., Imamura, K., & Tani, T. (2006, April). Orientation-restricted continuous visual exposure induces marked reorganization of orientation maps in early life. *Neuroimage*, 30(2), 462–477.
- Yang, X.-H., Huang, J., Zhu, C.-Y., Wang, Y.-F., Cheung, E. F. C., Chan, R. C. K., & Xie, G.-R. (2014, December). Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry Res.*, 220(3), 874–882.