

Contextual Influences on Cognitive and Psychophysiological Mechanisms of
Learning in Early Adolescence

A DISSERTATION SUBMITTED TO THE FACULTY
OF UNIVERSITY OF MINNESOTA BY

Meriah Lee DeJoseph, M.A.

IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisors: Drs. Daniel Berry & Kathleen Thomas

June 2023

© Meriah Lee DeJoseph, 2023

Acknowledgements

It is with a deep sense of appreciation that I reflect on the vast collection of people who have helped me realize a dream once far beyond what I thought possible. My pursuit of higher education—a long journey filled with many challenges and uncertainties—was not only made manageable, but incredibly rewarding, because of them. While I know that a handful of pages can never fully capture my gratitude, I would like to take a moment to acknowledge the countless individuals who have inspired and championed me along the way.

I am indebted to the many mentors who have played a pivotal role in helping me navigate graduate school and grow both personally and professionally. My heartfelt thanks go to my advisors, Drs. Dan Berry and Katie Thomas, for their invaluable guidance and support throughout my doctoral training. Their steadfast belief in me afforded the space to take risks and challenge myself in ways essential for my development. I would also like to extend my sincere gratitude for my committee members, Drs. Willem Frankenhuis and Ann Masten, for generously offering their time and wisdom and whose pioneering work continues to inform my thinking. I also thank the many ICD faculty and staff who have enriched my time at ICD, especially Dr. Dante Cicchetti, who taught me to “turn fear into fire.” Finally, I remain humbled by the longstanding mentorship of Drs. Cybele Raver and Clancy Blair, whose thoughtful advocacy early in my training launched my academic career path.

My immense gratitude goes to my dedicated and hardworking research assistants Brittney Olivares, Amelia Spaeth, Sarah Kuplic, Cielo DeCastro, Meghan Harwell, Sofiya Briguene, Mari Belina, and Emilie Sparrow. They not only made the current dissertation study possible, but also made the entire process joyful and fulfilling in ways that I will forever cherish. I also extend my gratitude for the many other stellar mentees I have had the honor of working with and learning from throughout my doctoral training, especially Otiti Mayo, Salma Ibrahim, and Lilly Bendel-Stenzel.

Additional thanks to colleagues Kate Nussenbaum for consulting on the computational models presented in the current dissertation study, as well as my computational modeling peer reading group—Felix Pichardo, Seokyoung Kim, Brandon Almy, and Norwood Gaspie. I would also like to give a special shoutout to ChatGPT, my personal coding wizard, who never tired of my endless questions and magically solved many of my programming woes.

I am greatly appreciative of the many funding sources—especially the Ford Foundation Predoctoral Fellowship and UMN Doctoral Dissertation Fellowship—that provided me the privilege of focused time and resources needed for my doctoral training and dissertation work. I am also fortunate to be a recipient of ICD’s Dr. Carrie DePasquale award for first-gen students, an award created after Carrie’s sudden passing in 2020. Our community continues to remember the joy and brilliance she exuded as a young scholar.

I am also fortunate to have been surrounded by a supportive and inspiring community of friends and colleagues throughout my academic journey. I first thank my BAB cohort, who were the foundation and backbone of this entire experience. I would also like to thank my labmates—especially Isa Stallworthy, Robin Sifre, Keira Leneman, Alyssa Palmer, and Shreya Lakhan-Pal—who served as both dedicated peer mentors and close friends that generated some of my fondest memories in Minneapolis. I also thank the rich and inspiring community of colleagues I have gained from the NextGen Psych Scholars Program. Finally,

special thanks to the many other friends and dear colleagues I have gained along the way, including (but not limited to) Carolyn Lasch, Charisse Pickron, Annie Brandes-Aitken, Rose Perry, and Stephen Braren. You all uplifted and sustained me through this crazy marathon.

To my family, who shaped my early development. My whole-hearted gratitude goes to my mom, Maureen, whose love, integrity, and spirit have been a guiding light throughout my life. I owe a great deal of my success to the sacrifices she made and the obstacles we overcame and became stronger from, which instilled in me the determination needed to reach our shared vision of a brighter future. I am also fortunate to have a loving and vibrant extended family that helped raise me. I am especially grateful for my late grandparents, Jack and Sharon, who taught me to always shoot my shot; for my Aunts, Sharon and Theresa, who illuminated the magic of manifesting; and for my many cousins, as well as my best childhood friends, Madison and Ashley, who functioned like siblings and cheered me on every step of the way.

Finally, to my husband and biggest hype man, Neil, who has been my anchor since we met in college nearly 11 years ago. I am immeasurably grateful for his support and belief in me, as well as his courage and adaptability through the many twists and turns my academic path has taken us.

Dedication

To my mom, whose enduring strength and unwavering support propelled me to dream big and strive for more. This achievement is as much yours as it is mine.

Abstract

Learning is a central mechanism through which early experiences shape biological and behavioral development across the lifespan. One type of learning, called *reinforcement learning*, is posited to support youth's ability to engage and adapt to their unique worlds with links to long-term social and emotional outcomes. Yet, individual differences in reinforcement learning across diverse environmental and experimental contexts remains poorly characterized in developmental samples. The current dissertation study integrated reinforcement learning and dynamic systems frameworks and drew upon newly adapted methodologies to capture how cognitive and psychophysiological processes of learning are modulated by socioemotional context. In a sample of 56 youth aged 12-15-years-old, this study leveraged a within-person experimental design and quantified continuous behavior and heart rate (~700 observations per system, per person) during an adapted reinforcement learning task with stimuli that varied in socioemotional relevance. Based on a series of Bayesian multilevel models, findings revealed that compared to traditionally-used benign or non-emotional stimuli, learning from stimuli high in socioemotional arousal enhanced behavioral performance. The use of computational modeling afforded valuable insights into the differential cognitive processes and strategies youth recruited to achieve such a behavioral advantage, demonstrating that socioemotional salience may have elicited faster value-updating processes and qualitative shifts in more exploitative decision-making. Underlying psychophysiological engagement seemed to be particularly modulated not by socioemotional salience as hypothesized, but by heightened sensitivity to learning from rewards, such that faster value-updating in the context of rewards aligned with more optimal psychophysiological flexibility and organization. Taken together, this study provides an important step in clarifying the contexts and modulatory processes that serve to enhance and support the unique ways youth learn and make decisions. Open questions remain about the adaptive utility of these various patterns of behavior, cognition, and psychophysiology across a variety of learning contexts, how they are shaped by prior lived experiences across development, and how they predict later psychosocial adjustment outcomes. Such work will shed light on how youth learn from—and adapt to—different contextual demands, with the potential to inform programs and policies that support youth's ability to adjust to their dynamically changing ecologies.

Table of Contents

List of Tables.....	vi
List of Figures.....	vii
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: CURRENT STUDY.....	5
2.1 Developmental foundations of learning.....	5
2.2 Neurocognitive computational modeling perspectives.....	6
2.3 Learning in context: the role of socioemotional salience.....	9
2.4 Psychophysiological processes of engagement to contextual learning demands.....	12
2.5 The current study.....	15
CHAPTER 3: METHODS.....	19
3.1 Participants.....	19
3.2 Procedure.....	19
3.3 Open science statement.....	20
3.4 Measures.....	20
3.4.1 Learning task.....	20
3.4.2 Autonomic psychophysiology.....	23
3.5 Analytic strategy.....	24
3.5.1 Aim 1: Learning and cognitive processes across task versions.....	27
3.5.1.1 Aim 1a: Behavioral learning performance.....	27
3.5.1.2 Aim 1b: Cognitive learning processes.....	27
3.5.2 Aim 2: Psychophysiological engagement as a function of cognitive strategy and task context.....	33
3.5.3 Deviations from pre-registration.....	36
CHAPTER 4: RESULTS.....	37
4.1 Aim 1: Learning behavior and underlying cognitive processes across task versions.....	37
4.1.1 Behavioral learning performance as a function of task context.....	37
4.1.2 Cognitive learning processes as a function of task context.....	39
4.1.2.1 Computational model results.....	39
4.1.2.2 Bayesian multilevel model results.....	40
4.1.2.3 Aim 1 summary.....	43
4.2 Aim 2: Psychophysiological engagement as a function of cognitive process and task context.....	45
CHAPTER 5: DISCUSSION.....	50
5.1 Socioemotional salience enhances youths' behavioral performance and cognitive learning strategies (Aim 1).....	51
5.2 Youth psychophysiology is flexibly organized irrespective of task context and is linked to reward learning (Aim 2).....	55
5.3 Limitations and future directions.....	58
5.4 Conclusions.....	62
REFERENCES.....	64
SUPPLEMENTAL MATERIAL.....	72

List of Tables

<i>Table 1. Sociodemographics of the current sample.....</i>	<i>19</i>
<i>Table 2. Final Bayesian mixed effect models predicting behavior.....</i>	<i>39</i>
<i>Table 3. Reward-punishment computational model results.....</i>	<i>40</i>
<i>Table 4. Final Bayesian mixed effect models predicting computational parameters.....</i>	<i>41</i>
<i>Table 5. Model comparisons for Bayesian models in Aim 1.....</i>	<i>44</i>
<i>Table 6. Final Bayesian mixed effect models predicting psychophysiology.....</i>	<i>48</i>

List of Figures

<i>Figure 1. Computerized learning task schematic.....</i>	<i>23</i>
<i>Figure 2 Illustrative overview of the learning rate parameter.....</i>	<i>30</i>
<i>Figure 3. Illustrative overview of the inverse temperature parameter.....</i>	<i>32</i>
<i>Figure 4. Simulated illustration of systems with varying degrees of organization.....</i>	<i>35</i>
<i>Figure 5. Main effect of task version from Bayesian multilevel models on total points.....</i>	<i>39</i>
<i>Figure 6. Main effect of task version from Bayesian multilevel models on learning rates.....</i>	<i>42</i>
<i>Figure 7. Interactive effect of age and task version.....</i>	<i>43</i>
<i>Figure 8. Within-version correlations across computational parameters.....</i>	<i>45</i>
<i>Figure 9. Main effect of task version predicting psychophysiology.....</i>	<i>47</i>
<i>Figure 10. Main effect of computational parameters predicting psychophysiology.....</i>	<i>49</i>

Chapter 1: Introduction

“Growing up in the environment I came from, I am not proud of all I did. But it was a learning experience. My role models and mentors used to be drug dealers and gangsters. Now I am feeding off positive energy.” - Calvin Cordozar Broadus Jr. (Stern, 2012)

Growing up in poverty presents a complex constellation of daily stressors, opportunities, and constraints that shape how children learn from and engage with their environments. Although many poverty-related experiences can increase children’s risk of negative outcomes (Blair & Raver, 2016; Peverill et al., 2020), they may also contribute to the development of unique skills and strengths necessary for responding to both the harsh demands and positive opportunities that arise in the context of poverty (Frankenhuis & Nettle, 2020). For example, exposure to economic hardship and discrimination may foster motivation and strategic insights leveraged to resist such inequities (Roy et al., 2019). Coping with unstable housing and material scarcity may enhance children's ability to adapt flexibly to change and develop resourcefulness (Fields et al., 2021; Mittal et al., 2015). Navigating high-crime neighborhoods and under-resourced schools may enhance empathic development necessary for discerning threats from opportunities and for building strong social networks that maximize resource sharing and foster community (Kraus et al., 2023). At the more micro-level, while children in poverty are more likely to be exposed to family conflict, such as domestic violence and parental mental health challenges (Jenson et al., 2017), they are also exposed to rich cultural traditions and close family bonds (Coll et al., 1996). Learning to rapidly predict and adapt to such interpersonal volatility may alter children's physiology and behavior in ways that promote exploitation of moments of safety and love while avoiding or coping with moments of threat and instability (Del Giudice et al., 2011). Taken together, this constellation of experiences may initiate the development of a

unique set of learned strategies that facilitates adaptation and resilience despite potential trade-offs to positive adaptation in other contexts (Ellis et al., 2022; Masten, 2018).

My growing program of research seeks to illuminate and humanize the highly variable and nuanced experiences of growing up in poverty and examine how such experiences adaptively alter child development across levels of behavior (e.g, DeJoseph et al., 2021), physiology (e.g., DeJoseph et al., 2019; DeJoseph et al., under review), and the brain (e.g., DeJoseph et al. 2022). Fundamental to this work is the idea that the timing, type, and temporal variability of poverty-related experiences shapes contextually-adaptive developmental mechanisms that prepare children to meet the demands of their specific environments (Del Giudice et al., 2011; Ellwood-Lowe et al., 2021; Tooley et al., 2021; Ugarte & Hastings, 2023). If better understood and properly acknowledged, these developmental processes may be leveraged to serve as a key conduit for supporting youth thriving both *because of* and in *spite of* poverty-related adversity.

Automatized and evolutionarily-conserved learning processes are one candidate mechanism that may help explain how poverty-related experiences give rise to biological and behavioral changes across the lifespan. For instance, to successfully navigate the aforementioned experiences a child in poverty might encounter, they may develop heightened attunement to environmental or emotional volatility. In such contexts, more rapid updating of value-based learned associations would be adaptive, as it would afford flexibility when contingencies change while maximizing fleeting rewards or resources that may otherwise be scarce. However, such a strategy would be suboptimal in contexts characterized by high stability or high stochasticity, whereby slower value-updating that considers longer-term experiential history is best. Thus, a core question my growing research program aims to answer is: *how does such experiential learning*

unfold across development, and what learning strategies are adaptive for whom, when, and under what contexts?

Learning is a complex and experientially-dependent process that involves the acquisition of knowledge, skills, and behaviors through various mechanisms reflected in rich theoretical models such as associative learning (Pearce & Bouton, 2001), statistical learning (Krough et al., 2013), and reinforcement learning (Sutton & Barto, 2018). Broadly, these learning processes share the common theme of forming associations between environmental cues and responses, and adjusting behavior based on the outcomes of those associations. Importantly, decades of research has shown that these learning processes are intricately linked to mechanisms of neuroplasticity through which experiences become biologically embedded to shape developmental trajectories (Greenough et al., 1987; Lin et al., 2020; Nelson et al., 2012).

While there is a large body of work documenting theoretically-informed model-based learning processes in adults (e.g., Frank, 2011; Eckstein et al., 2021), only recently have similar learning models been explored in developmental samples, leading to inconclusive findings (for a review, see Eckstein et al., 2022). One potential reason for these inconsistencies is a lack of studies that quantify interacting and dynamic modulatory processes (e.g., emotional stimuli that elicit arousal, moment-by-moment psychophysiological engagement) likely to influence learning. Further, even less research has sought to explore the role of early life adversity on learning (see McLaughlin et al., 2019 and Pollak, 2005 for a theoretical review), and most of this literature relies on deficit perspectives and study designs that limit the ability to capture potentially adaptive or ecologically-meaningful manifestations of learning. A greater understanding of how poverty-related experiences activate and shape learning across multiple systems in ways that both undermine and enhance development has the

potential to inform science, policy, and important strengths-based interventions and education.

The overarching goal of the present dissertation is to take a preliminary step toward my long-term goal of evaluating whether and how poverty-related experiences affect cognitive, behavioral, and physiological processes underlying learning, with the ultimate goal of identifying factors that optimize adaptation in specific contexts. In the dissertation described in the sections that follow, I take the necessary first step in assessing the role of *experimental context* in a nonpoverty sample of 12-15 year-olds. Conducting this empirical study afforded me the opportunity to gain expertise in theoretically-informed modeling approaches to learning and psychophysiology and test the hypothesis that socioemotionally-salient cues prioritizes arousal systems to enhance learning processes (Todd & Manaligod et al., 2018; Eckstein et al., 2022). This foundational work provides a proof of concept that warrants further exploration of how cues *salient to the context of poverty specifically*—from milliseconds (e.g., task stimuli) to years (e.g., type and temporal variation of lived experiences)—may similarly modulate how youth dynamically shift and optimize the way they learn from and navigate their uniquely complex worlds.

Chapter 2: Current Study

2.1 Developmental foundations of learning

Throughout development, we learn to navigate our unique worlds—or the many contexts, communities, and relationships we have—by attending to the predictability of objects and events, their meanings in relation to each other, and their associations with reward and punishment. Beginning as early as infancy, these foundational forms of value-based reinforcement learning are purported to facilitate evolutionarily-conserved adaptive behaviors that allow us to avoid threats, seek safety, and optimize rewards and opportunities in our environments (Frankenhuis et al., 2019). Such highly automatized learning processes guide attention and decision-making and are a central mechanism through which early experiences give rise to biological and behavioral changes across the lifespan (Greenough et al., 1987; Nelson et al., 2012; Thomas, & De Haan, 2012). Importantly, these learning-derived adaptations are posited to shape developing youth's ability to self-regulate (Keramati & Gutkin, 2014), develop social skills and relationships (Reeb-Sutherland et al., 2012), and make goal-directed decisions (Humphreys et al., 2016; Oudeyer et al., 2007), with far-reaching implications for long-term health and wellbeing (Brown et al., 2021; Vogel & Schwabe, 2016).

Early adolescence marks a particularly salient developmental stage that includes exposure to novel challenges and opportunities in the context of increasingly complex socioemotional demands co-occurring with rapid neurocognitive and biological changes (Fuhrmann et al., 2015). Revealing the complex processes by which youth engage and interact with a diverse array of environmental cues and transform that information to guide behavior is of critical importance for both science and policy. However, individual

differences in reinforcement learning across varied socioemotional cues and levels of analysis remain poorly characterized in developmental samples, as task paradigms are often limited to contextually arbitrary stimuli, and measures of learning are often limited to summary statistics of behavior. Grounded in a multilevel developmental approach, the current study draws upon several cross-disciplinary theoretical frameworks and leverages recent methodological advances to explore whether and how cognitive and psychophysiological mechanisms of learning among youth (ages 12-15 years) vary across socioemotional context.

2.2 Neurocognitive computational modeling perspectives of learning

Developmental researchers have used a wide range of tasks and performance indices to examine how value-based, or reinforcement learning (RL) unfolds. Many of these paradigms include reinforcement contingencies to choices involving abstract or neutral stimuli (e.g., shapes, colors), whereby participants make a series of choices between two to four options of which subsequently deliver positive (e.g., points gained) or negative outcomes (e.g., points lost). Reinforcement contingencies can be presented deterministically or probabilistically, and some can reverse, requiring participants to overcome prepotent (i.e., well-learned) associations to learn new ones (Shiu & Chan, 2006; Yaple & Yu, 2019). Broadly, this body of work has demonstrated that both aggregate performance accuracy (e.g., percentage correct responses, omission errors) and the asymptotic growth rates at which associations are accurately learned across task blocks increases over developmental time (Izquierdo et al., 2017; Bonawitz et al., 2014). Critically however, general indices of performance accuracy on these tasks obscure the underlying value-updating and cognitive processes theoretically posited to

drive learning. In other words, while performance can provide some assumed intel about how well someone is learning, it lacks insight into the many internal cognitive strategies one might employ to arrive at such performance.

In recent decades, pivotal advancements in the creation and application of mathematical RL models have afforded researchers with a powerful toolkit for describing fundamental cognitive learning processes, as well as the ways such processes map onto concurrent brain function (Rescorla et al., 1972; Sutton & Barto, 1998; Palminteri et al., 2017). More specifically, computational RL models leverage densely-sampled behavioral data from common learning tasks to estimate latent cognitive processes in the form of fitted model parameters (Ahn et al., 2017). To build an intuition for how these models work, consider an experimental task in which youth are instructed to maximize points by learning which stimuli to click in order to earn points. There may be several stimuli to learn from over the course of the task, and through trial and error they begin to assign higher value to the rewarded stimuli which subsequently guides their future decisions to click or not. Probing the child to explicitly report these value estimates for each stimulus would not only be disruptive, but would likely yield little information given that such cognitive processes are largely outside of conscious awareness. Instead, computational RL models link trial-by-trial behavior (e.g., decision to click versus not click) to falsifiable models that estimate individual-level values of theoretically-meaningful parameters posited to account for their observable behavior (Niv, 2021). These models operate in a feedback-dependent manner whereby they both produce predicted behavior and can also be inferred on the basis of those predictions (Ahn et al., 2017; Zhang et al., 2020).

In contrast to aggregate behavioral task performance indices, computational RL models can distinguish between cognitive processes that contribute to learning the value of different options and processes that translate those value estimates into choices (Ahn

et al., 2017). The initial value-updating process is captured by estimating a learning rate parameter, which reflects the extent to which individuals update their value estimates in response to a prediction error, or a discrepancy between the observed outcome and their expected outcome based on prior experience (Bolenz et al., 2017). Higher learning rates represent higher weight for prediction errors that are influenced by recent outcomes; the opposite is true for low learning rates (Nussenbaum & Hartley, 2019). For example, youth adopting a high learning rate in a task that has unpredictable outcomes will update their predictions more rapidly, causing them to adjust their behavior more quickly in response to shifts in the task. In contrast, youth adopting a low learning rate in a more predictable environment will update their predictions more slowly, leading to more consistent and stable behavior. Thus, learning rates reflect the extent to which past experiences guide subsequent actions. Although basic RL models estimate a single learning rate, prior work has shown that estimating separate learning rates scaled on positive/reward (i.e. better-than-expected outcomes) and negative/punishment (worse-than-expected outcomes) prediction errors better account for observed behavior than single learning rate models (Palminteri & Lebreton, 2022). This may be particularly true for developmental samples across a variety of learning tasks since adolescent youth tend to demonstrate amplified reward learning rates given enhanced sensitivity for rewards at that age (Master et al., 2020; Nussenbaum & Hartley, 2019). Empirical and simulated evidence further suggests that valence asymmetries in learning rates may reflect task-specific adaptation needed to reach optimal performance in accordance with task context, which may increase over age (Nussenbaum & Hartley, 2019; van den Bos et al., 2012; Christakou et al., 2013).

As value-based learned associations form via learning rates, they support the next step in the learning process: making choices and decisions based on the

information gathered. In RL models, this decision-making process is captured through the inverse temperature parameter (Cohen & Aston-Jones, 2005; Weiss et al., 2021). This parameter indicates the degree to which individuals explore lower-valued alternatives in their choices (low inverse temperature) or make high valued actions (high inverse temperature). Emerging research shows increases in inverse temperature over development (Decker et al., 2015; Nussenbaum & Hartley 2019; Palminteri et al., 2016), consistent with an explore-then-exploit developmental pattern (Gopnik et al., 2017). Multiple underlying cognitive and neurobiological processes are posited to govern these age-related changes. For instance, younger children with fewer constraints imposed by prior knowledge and high plasticity early in life tend to maximize exploration (Werchan & Amso, 2017). As youth age, it is speculated that increased executive function and increased knowledge about the world maximize reward as they switch to exploitation, becoming more rigid in their value-updating and decision-making processes (Weiss et al., 2021; Cohen et al., 2007). While individuals across ages can recruit exploratory decision-making strategies, children and youth tend to be more exploratory—and thus show lower inverse temperatures on average across tasks—than adults (Eckstein et al., 2022; Nussenbaum & Hartley 2019).

2.3 Learning in context: the role of socioemotional salience

Despite the value in capturing purported latent cognitive mechanisms that computational modeling affords, inconsistencies in study findings have led to calls for greater examination into additional contextual factors influencing within- and between-person differences in RL parameters (Nussenbaum & Hartley, 2019). A recent study using a large within-person design (N = 281; ages 8-30 years) showed that the inverse

temperature parameter followed similar developmental trajectories across a wide range of RL tasks (as shown in prior work; Somerville et al., 2017; Gopnik, 2020), whereas learning rates tended to differ across age and task context (Eckstein et al., 2022). The authors outlined several possibilities to account for observed discrepancies in learning rates from their study as well as in the developmental RL modeling literature more broadly. Here we focus on examining the possibility that RL parameters are influenced by modulatory processes that can be driven by ‘top-down’ (e.g., environmental uncertainty eliciting a stress response) or internally ‘bottom-up’ (e.g., anxiety increasing threat detection) and may serve to enhance, hinder, or shift RL parameter values. Relatedly, we consider the idea that value-updating (via learning rates) and decision-making processes (via inverse temperature) cannot be neatly composed into distinct component parts because such processes likely involve a complex combination of other processes that get integrated into individual parameter estimation.

The current study aimed to probe and expand upon the modulatory hypothesis by drawing from the priority state space framework (PSS; Todd & Manaligod, 2018). The PSS offers a particularly useful lens for understanding the role of modulatory processes underlying computational model-derived parameters of learning, as it posits that there are developmentally and contextually-prioritized sources of salience—both internal and external to individuals—that guide attention and thus subsequent learning. These sources of salience are hierarchically nested according to short and long-term goals and tuned by processes occurring at the moment-by-moment level and from experience over a larger developmental time scale. Integrating the concept of ‘state’ from reinforcement learning and ‘state space’ from dynamic systems frameworks, the PSS suggests that we navigate a diverse array of schemas representing different states of our unique worlds. In these schemas, features of the environment are prioritized based on what was most

important in past experiences in similar situations and constrained by the landscape of all possible states (i.e., state space) that may be activated in that context. Learning in the context of an experimental task, then, is inherently situated in a broader landscape of all the potential priorities that may guide behavior in a task, including the affective salience of task stimuli, short-term task related goals and motivations, and longer-term semantic associations shaped by prior learning across development.

One way to begin distinguishing how behavioral and cognitive indices of learning vary by modulatory processes is to examine learning from social and emotional stimuli, which are considered particularly important sources of external salience that guide attention in early adolescence (Bolenz et al., 2017; Rosenblum & Lewis, 2003; Rosen et al., 2018). Socioemotional stimuli may influence reinforcement learning in ways that either impair or enhance behavioral task performance and/or learning rates and inverse temperature. Impairment may occur if socioemotional stimuli interfere with or compete with limited cognitive processing resources, and thus impede learning performance or lead to less optimal value-updating and decision making processes. However, support for this notion is largely limited to studies of working memory (e.g., Garrison & Schmeichel, 2018; Mather et al., 2016), which becomes de-prioritized over the course of a typical reinforcement learning task when the number of stimuli to be learned far exceeds working memory capacity (Yoo & Collins, 2021). In other words, memory constraints will likely not diminish the ability to learn from socioemotionally engaging stimuli.

In contrast to impairment, enhancement may occur as a result of mobilizing cognitive and attentional resources when learning from socioemotionally engaging stimuli, thereby fostering greater alignment with associations shaped by prior real-world learning experiences that were emotional in nature. Indeed, prior research in adults has

shown that both positive and negative emotional information is prioritized (i.e., faster detection and improve recall) over neutral information across a variety of domains including attention (e.g., faster detection; Pool et al., 2016; Todd et al., 2012) and memory (e.g, improved recall; Talmi, 2013; Vuilleumier & Huang, 2009). Far less evidence exists in the learning domain, but recent work, albeit limited to adult samples, shows initial support for the enhancement hypothesis. Using a statistical learning task whereby participants had to learn probabilistic patterns within continuous streams of stimuli, Everaert and colleagues (2020) demonstrated higher recognition accuracies for emotionally negative and positive stimuli compared to neutral stimuli. Building on this research, Plate and colleagues (2022) found an advantage for statistical learning of emotional faces over neutral faces or shapes, evidenced by earlier learning from the emotional stimuli. Given robust developmental literature highlighting the prioritized salience of, and sensitivity to, emotional signals in early adolescence (e.g., Rosenblum & Lewis, 2003), socioemotionally arousing cues may facilitate the value-updating and decision-making processes underlying reinforcement learning.

2.4 Psychophysiological processes of engagement to contextual learning demands

As top-down external modulation from socioemotional stimuli is likely to affect behavior and cognition during learning, processes at the biological level are similarly likely to be affected by such stimuli—in addition to exerting bottom-up modulation of behavior and cognition. For example, extant studies in adults have demonstrated differential brain activation and physiological patterns depending upon whether learning stimuli were emotional (e.g., Nashiro et al., 2012) or whether a pre-learning stressor

occurred before the learning task (e.g., Schwabe et al., 2008), that bidirectionally served to enhance behavioral learning performance in some cases. Although midbrain striatal dopaminergic circuitry has been the typical focus of reinforcement learning research (e.g., Wirz et al., 2018; Lin et al., 2020), the PSS framework emphasizes the regulatory role of the locus coeruleus norepinephrine (LC-NE) neuromodulator system in establishing attentional biases for particular categories of stimuli through learning. This is due to the input-output organization of the LC-NE system that directs arousal and attention, whereby inputs from bottom-up sensory stimuli via the brainstem and forebrain converge on inputs from top-down goal-directed regulation of behavior via the amygdala and prefrontal cortex (Breton-Provencher et al., 2021). In the context of reinforcement learning, the LC-NE system integrates and reconciles the influence of competing or mutually reinforcing sources of salience—including rewards, punishments, and type of stimuli in a given state (i.e., RL task), constrained by the larger attentional landscape and physiology shaped by prior experience. Because youth’s real-world attentional landscape is tuned more sensitively to socially and emotionally arousing cues, as well as to rewards more generally, the LC-NE system modulation is likely to stimulate greater physiological engagement in those contexts compared to non-emotional and/or punishing contexts.

According to the PSS and developmental systems frameworks, such physiological engagement during learning can be viewed as an emergent phenomenon arising from bottom-up and top-down LC-NE-driven neuromodulatory processes interacting across space and time. Neuroanatomically and functionally interwoven with the LC-NE system is the central autonomic nervous system (ANS), whereby the LC controls autonomic function from both direct projections to the spinal cord and projections to autonomic nuclei including the amygdala and vagus nerve implicated in

sympathetic and parasympathetic regulation, respectively (Thayer et al., 2009). Within-person autonomic coordination has been shown to differ as a function of affective context, driven by differences in the assembly of the coordination across the large network of bidirectional interactions in the LC-NE and autonomic systems (Gatzke-Kopp, Benson et al., 2020; Ulrich & Herman, 2009). Such dynamic coordination of the ANS is often characterized as a feature of a “complex system” reflective of a self-organized structure that arises from the assembly of multiple interconnected components (Smith & Thelen, 2003; Van Orden et al., 2003). Notably, this perspective contrasts with computational RL models as described above, which essentially “carves cognition at its joints” to derive specific components of cognitive processes underlying learning (Eckstein et al., 2022; though see Discussion section for a more thorough discussion on this point). Nonetheless, RL models are still considered to capture something about ‘process’, which is an improvement from aggregate accuracy scores. Mapping the concurrent real-time complex dynamics of psychophysiological modulation onto value-based and decision-making RL parameters may elucidate more nuanced multilevel processes supporting learning and reconcile inconsistent findings in the developmental RL literature.

One way to begin capturing these real-time psychophysiological dynamics during learning is through cardiac time series—specifically heart rate variability (HRV). HRV encompasses nonlinear organizational structures emerging from the interactive dynamics of the LC-NE and autonomic systems. These structures can be formalized mathematically by characterizing the system’s fractal patterns, or the repeated, self-similar organizational structure across nested levels over time. This ‘fractal complexity’ reflects organization in time that varies on a spectrum from very random to rigid, often indexed via the long-term interrelatedness or ‘color’ of the time series (Ihlen, 2012).

White noise is thought to represent random organization whereas brown noise represents highly self-similar or rigid temporal variability. Optimal temporal variability is indexed as pink noise, representative of a flexible self-organized system best suited to respond to environmental change (i.e., fractal). Fractal complexity is argued to represent the 'long-term memory' of a system such that preceding events within the system exert cascading effects on adaptive functioning, and has thus been associated with more positive mental and physical health outcomes (Beckers et al., 2006; de la Torre-Luque et al., 2016).

It remains an open question whether cardiac fractality is linked to learning, especially in developmental samples. However, some research has suggested cardiac fractality is a useful marker of cognitive processes more broadly (e.g., Favela, 2019; Van Orden et al., 2011; Wijnants, 2014). Relevant work in children from Berry and Stallworthy (2018) has demonstrated that cardiac complexity in the moderate (i.e., flexible pink noise) range was associated with better performance in an executive function task among youth who were actively challenged (i.e., not performing at ceiling). Similarly, in preschoolers, loosely organized patterns of fractality were associated with behavioral patterns of self-regulation, becoming more tightly organized during challenge (Berry et al., 2019). There is also evidence suggesting that fractality may serve as a proxy for social engagement (Stallworthy et al., 2020) and more general attentional engagement when viewing more complex images (Marlow et al., 2015). More work is needed before drawing strong substantive conclusions, but this emerging evidence suggests that cardiac fractality may provide novel insights into psychophysiological modulatory engagement underlying dynamic learning contexts.

2.5 The current study

The current pre-registered study integrates RL and developmentally-informed dynamic systems frameworks alongside recent methodological innovations to empirically examine the behavioral, cognitive, and psychophysiological processes of learning in early adolescence (ages 12-15 years old). Using a within-person experimental design adopting an adapted RL task (Finger et al., 2008) including a wide variety of learning stimuli ranging in socioemotional salience, this study sought to explore the role of key modulatory processes that serve to shift, enhance, and/or hinder learning. Two key aims were addressed.

Aim one examined whether youths' learning performance varies as a function of stimulus content (non-emotional versus socioemotionally-salient), and tested whether these various stimuli elicit different levels of traditional aggregate scores of behavioral performance (i.e., points earned) (1A) and computationally-derived cognitive mechanisms (i.e., learning rates and inverse temperature) (1B). Given that salient socioemotional stimuli likely enhance learning for youth, I hypothesized that performance would be more optimal (i.e., higher overall points) when learning from socioemotional stimuli compared to neutral, non-emotional stimuli. With respect to the RL model-derived parameters, I hypothesized that learning rates would reflect greater sensitivity to, and thus result in faster value-updating processes in the socioemotional context. For decision-making strategies via inverse temperature, it is likely that there will be no substantive version-based differences, given recent evidence for similar patterns across task demands (Eckstein et al., 2022). However, based on the PSS and some emerging research in the emotional learning literature (see above), I hypothesized that youths' decision-making strategies via inverse temperature may evidence greater exploitation when learning from socioemotional stimuli, given the role of affective salience in

recruiting prior experiential knowledge. In other words, youth may show patterns of decision-making that are more developmentally mature (i.e., exploitative rather than exploratory) when learning from affective socioemotional cues previously prioritized in day-to-day life.

Aim two explored the main and interactive effects of stimulus content (non-emotional versus socioemotionally-salient) and RL parameters on psychophysiological engagement, indexed via HRV fractality. I hypothesized that youth would exhibit temporal organizations suggestive of organized, yet flexible autonomic dynamics—as suggested by pink-noise dynamics—in the socioemotional condition than the non-emotional condition. Such a pattern would be reflective of greater arousal and engagement to task demands including affective salience. Hypotheses regarding the link between RL computational parameters and HRV fractality were more exploratory; however, a few possibilities were considered based on theory and some empirical work in other cognitive domains mentioned above. First, it is conceivable that relations could be linear, whereby faster value-updating processes (i.e., higher learning rates) and more exploitative decision-making (i.e., higher inverse temperature) requires greater attention and thus relies on the recruitment of higher psychophysiological modulatory engagement and potentially more rigid HRV coordination. In contrast, relations may show an inverted U-pattern, whereby extremely low or high learning rates and/or inverse temperature align with more random psychophysiological coordination reflective of greater disengagement and moderate parameter values align with more optimal psychophysiological coordination reflective of greater engagement. For both linear and nonlinear associations, task condition may moderate this association in different ways. Main effect relations may be functionally similar across conditions, but potentially magnified in the socioemotional condition. Alternatively, directional associations between cognitive

parameters and HRV fractality may only be apparent for the socioemotional condition, with a null relation in the non-emotional condition. In other words, alignment between cognitive and psychophysiological processes may only matter when youth are required to attend to and learn from emotionally affective cues that demands greater attention and regulation.

Chapter 3: Methods

3.1 Participants

A sample of 56 typically developing 12-15-year-old adolescents (M age = 13.3 years; 51.8% female) and their primary caregivers were recruited through the University of Minnesota's Institute of Child Development's Participant Pool. Exclusion criteria included (a) parent-reported developmental or learning disability and/or severe language, hearing, vision deficit; (b) diagnosed ASD, ADD/ADHD, ODD; (c) known deficits in fine motor and/or eye movement control; or (d) known cardio-respiratory abnormalities. Overall, participants were predominately White from middle to upper-middle class backgrounds. Sociodemographics are shown in Table 1 (also see supplemental Figure S1 for additional histograms).

Table 1. *Sociodemographics of the current sample.*

	N = 56	%
Gender identity		
Female	29	51.8%
Male	26	46.4%
Prefer not to say	1	1.8%
Race/Ethnicity*		
White	53	94.6%
Asian	11	19.6%
Hispanic	2	3.6%
Other (European)	1	1.8%
	Mean	SD
Age	13.3	0.9
Income-to-needs	6.1	3.5

*Note: Participants could report multiple race/ethnicities.

3.2. Procedure

All procedures were approved by the Institutional Review Board at the University of Minnesota (Study #8326). Data was collected during a one-and-a-half-hour laboratory

visit. After consent and assent, primary caregivers filled out the questionnaires in a separate room while youth completed a computerized learning task on a laptop as continuous heart rate was recorded. After the task, youth completed additional questionnaires. At the end of the visit, caregivers and youth were given \$10 and \$15 gift cards, respectively.

3.3 Open science statement

All procedures and analysis plans were pre-registered on OSF prior to data collection. Minor deviations and justifications for deviations are described in the Analytic Strategy section below. IRB-approved study materials, pre-registration, raw de-identified data, and code presented in the current manuscript can be found on the OSF project page here:

https://osf.io/xqj8w/?view_only=839bd73241844d3fae0c161a26fd38e9

3.4 Measures

3.4.1 Learning task

Youth completed two versions of an adapted computerized learning paradigm (Figure 1) in counterbalanced order while recording continuous cardiac activity from a wireless electrocardiogram. The 'nonemotional' version included 24 stimuli of line drawings (e.g., key, chair, shoe) from the original version of the task (Finger et al., 2008). The adapted 'socioemotionally-salient' version included 24 IAPS (Lang et al., 2005) stimuli of emotionally negative (e.g., neighborhood violence, couple fighting) and

positive (e.g., happy family, money) images rated high in arousal (Mean arousal = 10.4) and positively and negatively valenced (Mean negative = 14.8; Mean positive = 6.1). Stimuli were presented electronically using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

Youth were given the following verbal and written instructions, followed by a brief practice session (with images of shapes) before beginning the task:

“In this game, you are going to be presented with a series of pictures. Some of the pictures are winners and will gain you points if you click the spacebar when they are showing. Some are losers and will lose you points if you click the spacebar when they are showing. If you do not click, you will not gain or lose any points. Your goal is to win as many points as you can.

At some point, you will see pictures that show positive or negative emotional things in them. The positive ones do not necessarily mean they give points, and the negative ones do not necessarily mean they lose points. Remember that to learn which pictures are winners or losers, you should click a lot in the beginning and pay close attention as you continue through the task and new pictures are presented.”

Across both versions, the task started with an *acquisition phase* where youth learned stimulus-outcome associations via reward (where points are gained) or punishment (where points are lost). Subsequently, the outcome associated with half of

the stimuli switched to the opposite reinforcement value, forming the *reversal learning phase*. Each acquisition and reversal phase consisted of 8 blocks (96 trials) and together constituted one run (192 trials total per run). Each version of the task contained two runs.

For each run in each version, six stimuli (3 negative and 3 positive for the socioemotional version) were initially associated with reward: if a button press was made while the stimulus was displayed, the participant received positive feedback (“you WIN 100 points” paired with a pleasant “ding” sound) and a running point total was displayed. Another six images were associated with punishment and negative feedback (“you LOSE 100 points” paired with an aversive alarm noise). Each version of the task contained two runs, each with a new set of 12 stimuli presented in random order. Youth completed a total of 768 trials (384 per version) to learn a total of 48 stimulus-outcome pairs (24 each version). Each trial lasted 2300 ms, beginning with an image for 1100 ms followed by the feedback screen for 1000 ms and finally a fixation cross for 200 ms.

Performance on this task relied on the formation of stimulus–outcome associations in the acquisition phase and then the subsequent replacement of the original stimulus–outcome association with an updated association in a reversal phase. Correct hits (i.e., press on a reward stimulus) resulted in a point added, false alarms (i.e., press on a punishment stimulus) resulted in a point taken away, and a correct avoidance (i.e., correct miss on punishment stimulus) or miss (i.e., incorrect/accidental miss of reward stimulus) resulted in no points gained or lost. A summation of points across the three response types (see supplemental Figure S2 for descriptives) was used to index overall performance, where higher scores indicate more ‘optimal’ performance. In other words, higher scores reflected more correct hits on reward trials, fewer false alarms on punishment trials, and fewer incorrect misses.

Computational cognitive modeling was applied to generate more process-oriented aspects of learning across the task versions (see Analytic Plan).

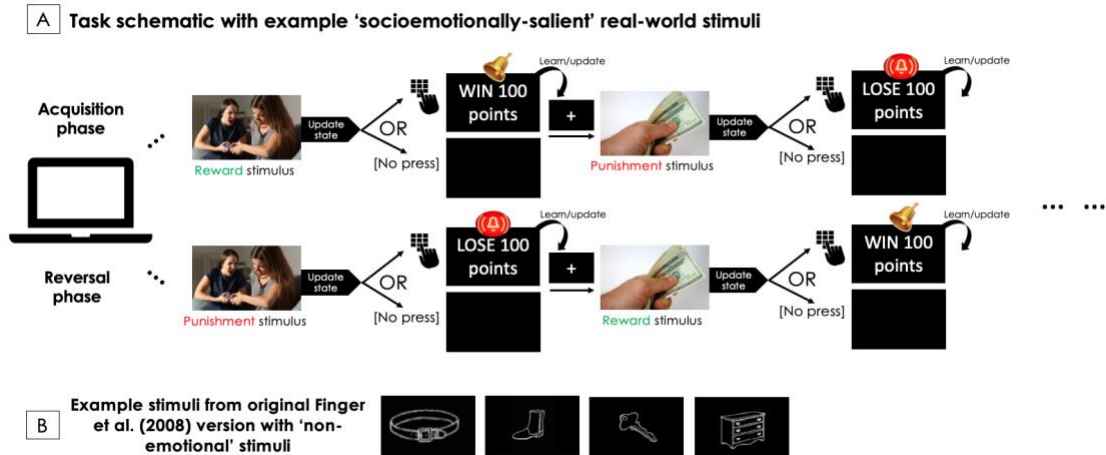


Figure 1. (A) Computerized learning task schematic with example 'socioemotionally-salient' real-world stimuli. Note: actual IAPS images included in the task are not shown due to copyright rules. During the acquisition phase (top row), a button press to a given stimulus was followed by either a reward paired with a pleasant ding sound (top left) or punishment paired with an unpleasant buzz sound (top right). During the reversal phase (bottom row), half of the stimuli switched their association. (B) Example 'nonemotional' stimuli from original Finger et al., (2008) task. Participants completed both the socioemotionally-salient and non-emotional version presented in randomized order.

3.4.2 Autonomic psychophysiology

Electrocardiography (ECG) was sampled at 1000 Hz from youth during a five-minute resting baseline (watching a nature video) and throughout the learning task using a BIOPAC MP150 system and a PC running Acqknowledge 5.0 software. Wireless ECG transmitters were applied to pre-gelled spot electrodes located on the adolescents' torso (Figure S3), using lead II (ECG) configurations. A USB-TTL was used to send event markers from E-prime to BIOPAC. These markers were used to

index the start/stop times of the phases and runs of the learning task. Inter-beat interval (IBI) measures, or the milliseconds between each successive heartbeat, were derived from the raw data using AcqKnowledge software and further segmented according to the extracted start/stop times of the runs for each version of the learning task.

Research assistants trained and achieved research reliability on gold-standard ECG artifact removal procedures (CardioEdit; Brain-Body Center for Psychophysiology and Bioengineering at the University of North Carolina) before visually inspecting and removing all artifacts that were likely due to recording error.

Temporal HRV complexity, a measure thought to reflect the dynamic of real-time organization of the autonomic system, was calculated from the raw artifact-free IBI data and used as an index of psychophysiological engagement. Further details about how this index was calculated are provided in the analytic strategy below.

3.5 Analytic strategy

All analyses were conducted in R using RStudio version 2022.7.2.576. Reproducible code can be found in the OSF link provided above (see Open Science Statement).

Substantive models for both aims were fit using a fully Bayesian approach (see Supplement for conceptual description comparing Bayesian with frequentist approaches). Specifically, multilevel Bayesian analysis was used as it is particularly amenable to the small sample size of the current study (see Lee & Song 2004; Hox et al., 2012 for simulation studies) and overcomes common limitations with maximum likelihood approaches by estimating the distribution of each free parameter over the group of participants and for each participant individually (van de Schoot et al., 2014).

All models were fit using the *brms* package in R (Burkner, 2017). The syntax and steps to conduct Bayesian models using *brms* was based on the tutorials and resources by Fusaroli and Cox (2022), McElreath (2020), Kurz (2022, 2023), and Veenman et al. (2022). Priors in all models outlined in the sections below were defined via a principled approach that applies accessible knowledge about the realistic parameter space as a reasonable starting point and subsequently examines the implications of those choices using prior sensitivity checks, making adjustments as needed to improve model convergence (Fusaroli & Cox 2022). This approach is preferred over the use of noninformative priors, which have been shown to lead to biased estimates, especially in small samples (McNeish, 2016; van de Schoot et al., 2014). This is because non-informative priors, such as a uniform distribution without bounds, assign equal probability density to a very large range of values and thus any effect becomes unlikely (Gelman, 2006; Veenman et al., 2022). In turn, this negatively influences model comparisons such that the support for the null hypothesis becomes increasingly large (Veenman et al., 2022).

For all outcome models (described in more detail in subsequent sections), a Student's t-distribution was specified as the model family to capture heterogeneity and/or account for potential outliers in the data. Priors were specified based on the realm of realistic parameter space using knowledge about the practical constraints of the current data (i.e., min and max possible values) as well as relevant prior empirical and theoretical literature. Given the relative lack of prior empirical work examining the associations of interest, priors were specified with large variances around the means to represent greater uncertainty in the true population values. Certain parameters were also given an upper and lower bound based on the minimum and maximum number of possible values on a given outcome or prior literature. Priors for standard deviation

parameters were drawn from an inverse gamma distribution (which only allows for non-negative values), as recommended by Veenman et al (2022). The family-level Student's t-distribution also requires an LKJ prior (R) for the correlation structure and a gamma prior (N) for the degrees of freedom, which were defined based on default recommendations (Fusaroli & Cox 2022). Prior and posterior predictive sensitivity checks were conducted on each model and priors were updated if necessary.

Posterior distributions were approximated using the Stan software package that implements a Markov Chain Monte Carlo (MCMC) algorithm called Hamiltonian Monte Carlo (HMC). All models were fit using typically recommended settings for Bayesian multilevel models (McElreath 2020): 4 chains, 4 cores, and 2 threads per chain, with a total of 4000 iterations and 1000 warmup samples. Statistically meaningful associations between predictors and dependent variables in a Bayesian framework are defined as a 95% posterior credible interval (CrI) of the regression coefficient excluding 0, analogous to a frequentist alpha of .05. Model fit was evaluated using convergence diagnostics including Rhat (1.00 - 1.04 indicates good parameter convergence) and graphical inspection of trace plots to examine convergence across chains. Pending successful convergence, competing models were compared using WAIC and LOOIC, which are measures of out-of-sample predictive accuracy computed from the posterior distribution (Vehtari et al., 2017). WAIC (widely applicable information criterion) is calculated as the sum of the log pointwise predictive density (LPPD) that functions as a penalty term for the effective number of parameters in the model. LOOIC (leave-one-out-cross validation information criterion) is calculated as the average log pointwise predictive density (LPPD) for each data point, with the model fit on all but one data point and the prediction made for the left-out data point. Lower WAIC and lower LOOIC values within a series of nested models indicate better fit.

Two methods common in Bayesian modeling were used to assess the level of support, uncertainty, and magnitude of effects of interest. First, formal hypothesis testing was conducted on the final best fitting model for each parameter of interest. Evidence ratios (i.e., Bayes factors) and credible intervals were used to determine the support and levels of uncertainty for each hypothesis within a Bayesian framework. Given that evidence ratios can take on a wide range of numeric values and depend upon model constraints, ratios were twice log transformed as recommended by Kass and Raftery (1995) to draw more interpretable conclusions. On this scale, evidence for a hypothesis ranges from weak (0-2), modest (2-6), strong (6-10), and very strong (>10) and is used solely as a descriptive index. Second, reported standardized effect sizes were calculated and scaled on the square root of the unconditional variance, based on the most relevant level of analysis (i.e. between-person versus within-person). Reported Cohen's d for task version were calculated as the difference in means between the two task conditions divided by the pooled standard deviation of the differences. Using this method, the effect size estimate reflects the magnitude of the difference between the two conditions relative to the variability within the subjects, which is appropriate for the within-person design of the current study.

All models including continuous variables were grand mean-centered.

Additional specifics for the application of aim-specific models is described below.

3.5.1 Aim 1: Learning and cognitive processes across task versions

3.5.1.1 Aim 1a: Behavioral learning performance as a function of task context

To examine relations between learning context and performance, I regressed the overall learning performance (total points sum score; see Measures above) on version (non-emotional vs. socioemotionally-salient) as fixed-effects, and a random intercept and slope for each participant:

$$\text{total points} \sim 1 + \text{version} + (1 + \text{version} | \text{subject})$$

Priors were specified as follows:

$$\begin{aligned} \text{total points} &\sim \text{student's } t(\mu_i, \sigma_e) \\ \mu_i &= \alpha_{\text{subject}[i]} + \beta * \text{version}_{\text{subject}[i]} \\ \alpha_{\text{subject}} &\sim \text{normal}(100, 25), \text{ lb} = -384, \text{ ub} = 384 \\ \sigma_{\text{subject}} &\sim \text{inverse-gamma}(5, 75) \\ \beta &\sim \text{normal}(0, 40) \\ \sigma_{\text{version}} &\sim \text{inverse-gamma}(5, 75) \\ \sigma_e &\sim \text{inverse-gamma}(5, 85) \\ \mathcal{R} &\sim \text{LJKcorr}(2) \\ N &\sim \gamma(2, .1) \end{aligned}$$

A second model was fit to total points with age (mean-centered and using a normal(0,40) prior) added as a fixed effect to examine whether age explained substantial variance in the outcome. Model comparisons (described above) determined the final model.

3.5.1.2 Aim 1b: Cognitive learning processes as a function of task context

Reinforcement computational learning models were also fitted to Bayesian multilevel models using the *hBayesDMR* package (Ahn et al., 2017) to examine latent cognitive learning processes that are posited to underlie performance on each version of the task. Under this reinforcement learning modeling framework, it is assumed that youth learn about the values of task stimuli and actions from their outcomes throughout the learning process (i.e., during the task). Decision-making relies on incremental reward-

prediction errors that apply a policy to maximize rewards via choosing the more valuable action.

To capture this reinforcement learning process in the current task, I applied the Reward-Punishment Model for Multiple-Block Probabilistic Reversal Learning (*pri_rp_multipleB* function using default parameters) where block = stimulus image, choice = press (1), no press (2), outcome = reward (1), neutral (0), punishment (-1). This model estimates three parameters: (1) punishment learning rate, (2) reward learning rate, and (3) inverse temperature.

Learning rate ($0 < \alpha < 1$) indexes how rapidly the expected value of an action $Q(a)$ is updated based on prediction error, or the difference between the received outcome value o and prior expected value Q at the previous time point t . The learning rate scales the prediction error, such that higher learning rates (i.e., closer to 1) more heavily weigh recent outcomes for a given stimulus, and lower learning rates (i.e., closer to 0) distribute the weight across prediction errors collected for that stimulus throughout the task (Figure 2). Thus, learning rates reflect sensitivity to past experiences and the extent to which past experiences guide subsequent actions. See Figure 2 for an illustrative overview of learning rate.

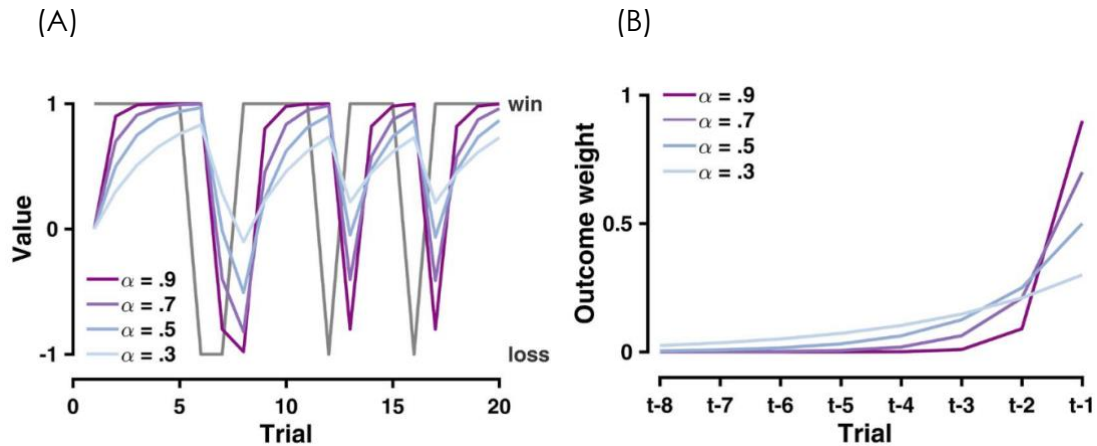


Figure 2. Illustrative overview of the learning rate parameter adapted from Zhang et al., 2020. (A) Simulated illustration of different learning rates on value update for a given stimulus-outcome association. In a learning environment where half of the stimuli switch reward contingencies every 96 trials, as in the current study, a moderately high learning rate ($\alpha = .7 - .9$) leads to faster value updating and thus more optimal performance. For example, $\alpha = .7 - .9$ will result in the updated value of a given stimulus to approximate its maximum after approximately two to four trials, if a reward was observed. Similarly, if a loss was observed on a later trial (i.e., a stimulus switches from reward to loss), then a fast value decrease is observed (see trials 7-8). (B) Simulated illustration of different learning rates on the weights of past outcomes, where trial denotes a trial on a given stimulus. As in (A), moderately high learning rates of $.7 - .9$, the more recent trials (i.e., trials $t-1 - t-4$) contribute to the outcome weight, and the weight on outcomes observed farther in the past are reduced. Taken together, in the context of the current learning task design (where reward schedule is volatile via reversal every 96 trials), more recent positive and negative feedback is informative in that it allows the participant to detect the reversal and recompute action values.

In the current model, learning rates (α) are parameterized in a feedback-dependent manner, separated based on whether feedback was positive or negative. In other words, when feedback for a given stimulus is better than expected, a positive prediction error ensues and increases the decision weight to press. When feedback is worse than expected, a negative prediction error ensues and decreases the decision to press. The impact of these separate types of prediction errors is scaled by estimated reward (α_{rew}) and punishment (α_{pun}) learning rates:

$$Q_{t+1}(a) = \begin{cases} Q_t(a) + \alpha_{\text{rew}} [o_t - Q_t(a)], & \text{if } o_t = 1 \\ Q_t(a) + \alpha_{\text{pun}} [o_t - Q_t(a)], & \text{if } o_t = -1 \end{cases}$$

In the RL modeling framework, after values for a given stimulus are updated, the next step is to use those values the next time that stimulus is observed to make a decision. These action values inform choices probabilistically based on a softmax choice function, where a is the selected action and a_{ns} is the non-selected action scaled by the inverse temperature parameter ($\beta > 0$):

$$P_t(a) = \frac{\exp(\beta Q_t(a))}{\exp(\beta Q_t(a)) + \exp(\beta Q_t(a_{ns}))}$$

Conceptually, this function transforms learned value-estimates of different options (i.e., press versus not press) into choice probabilities, where the magnitude of the differences in the options is scaled by the inverse temperature (Figure 3). Inverse temperature β is posited to reflect the extent of choice stochasticity, or the degree of exploration-exploitation in the decision making process. Higher inverse temperatures represent more exaggerated differences in the choice options, resulting in greater rigidity in choices, greater risk aversion, and exploitation of higher valued choices. Lower inverse temperatures represent a smaller magnitude of difference between choices, resulting in greater randomness in choices, greater risk taking, and exploration of lower valued choices. See Figure 3 for an illustrative overview of inverse temperature.

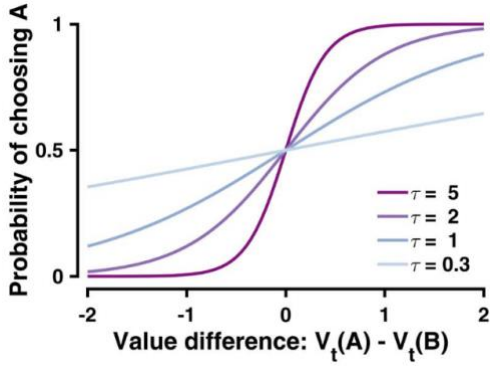


Figure 3. Illustrative overview of the inverse temperature parameter adapted from Zhang et al., 2020. In the current study, where the choice is to click on the stimulus or not, the softmax choice function is simplified as a logistic curve whereby the input is a value difference (x-axis) between the choices, and the output is the probability of choosing the selected action. As denoted by the legend, different values of inverse temperature represent different slopes of the sigmoid curve, which captures choice consistency. When the inverse temperature is higher, as is optimal in the current study, the curve is steeper, which represents more consistent choices that favor or exploit the rewarded option.

Learning rates and inverse temperature parameters were fit separately for the non-emotional and socioemotionally-salient versions of the task and individual-level estimates of each parameter were extracted, resulting in two sets of parameters for each person. Punishment learning rate, reward learning rate, and inverse temperature were then regressed separately on task version using the same model fitting procedures as Aim 1a, where:

$$\text{computational parameter} \sim 1 + \text{version} + (1 + \text{version}|\text{subject})$$

Baseline model priors for the learning rate outcomes (punishment and reward) and inverse temperature outcome were defined as follows:

learning rate \sim student's $t(\mu_i, \sigma_e)$
 $\mu_i = \alpha_{\text{subject}[i]} + \beta * \text{version}_{\text{subject}[i]}$
 $\alpha_{\text{subject}} \sim \text{normal}(.5, .5), \text{lb} = 0, \text{ub} = 1$
 $\sigma_{\text{subject}} \sim \text{inverse-gamma}(3, .6)$
 $\beta \sim \text{normal}(0, 1)$
 $\sigma_{\text{version}} \sim \text{inverse-gamma}(3, .6)$
 $\sigma_e \sim \text{inverse-gamma}(3, .7)$
 $\mathcal{R} \sim \text{LJKcorr}(2)$
 $N \sim \gamma(2, .1)$

inverse temperature \sim student's $t(\mu_i, \sigma_e)$
 $\mu_i = \alpha_{\text{subject}[i]} + \beta * \text{version}_{\text{subject}[i]}$
 $\alpha_{\text{subject}} \sim \text{normal}(2, 2), \text{lb} = 0, \text{ub} = 6$
 $\sigma_{\text{subject}} \sim \text{inverse-gamma}(3, .6)$
 $\beta \sim \text{normal}(0, 1)$
 $\sigma_{\text{version}} \sim \text{inverse-gamma}(3, .6)$
 $\sigma_e \sim \text{inverse-gamma}(3, .7)$
 $\mathcal{R} \sim \text{LJKcorr}(2)$
 $N \sim \gamma(2, .1)$

Similar to Aim 1a, the above models were additionally fit with age as a covariate and model comparisons determined the final model for each computational model-derived learning outcome.

3.5.2 Aim 2: Psychophysiological engagement as a function of cognitive strategy and task context

Psychophysiological engagement in response to contextual learning demands was quantified within individuals using techniques from complexity science. Specifically, an index of cardiac complexity was estimated using detrended fluctuation analysis (DFA; Peng et al., 1995) fit to each subject's extracted IBI time-series measured during each version of the task. Broadly, DFA is a method used to analyze long-term correlations and scaling properties in time-series data. DFA is posited to capture the extent to which a system (e.g., physiology) exhibits temporal noise consistent with different levels of system organization (ranging from highly predictive or 'rigid' organizational properties (brown noise), to random organization (white noise), where the middle of the distribution (pink noise) represents a system organization thought to reflect flexible blend of to the two. Pink noise temporal structure is consistent with fractal organization. The extent of fractal structure along this continuum of temporal noise is represented by the relationship between power (P; size of change) and frequency (f; how frequently changes of P occurs), which can be mathematically inferred by calculating a scaling

exponent, α (Figure 4). Fractal structure serves as a putative index of emergent self-organization in a given system, reflected in long-range dependencies and nested scale-invariant patterns of variability (Ihlen, 2012), thought to underlie cognitive processes such as learning (Kello et al., 2008; Van Orden et al., 2011).

Autonomic fractal organization was assessed using the *RHRV* package in R (Martinez et al., 2017). An alpha (α) score for each participant was generated for each version of the task, providing a summary index of levels of psychophysiological engagement via how organized or presumably 'fractal' their system was. To compute this score, the cleaned IBI time-series (see Measures) is converted to a random walk by subtracting the mean value and integrating a cumulative sum. Then, the time-series was divided into segments of different sizes, and each segment was detrended using a linear function. The root-mean-square (RMS) fluctuation was calculated for each segment, and the log-log relation between the segment size and RMS was used to estimate the scaling exponent (i.e., Hurst exponent). The alpha score was then calculated as the average of the scaling exponents across all segment sizes.

Continuous values of α fall between 0 to 2, which are purported to align with specific patterns of fluctuations described as colored noise (Coey et al., 2012). White noise values tend to fall near an α of 0.5 and are posited to reflect more random fluctuations, pink noise values fall near an α of .7 to 1 and are posited to reflect greater self-similarity and flexible organization, and brown noise values fall near an α of 1.5 to 2 and are posited to reflect a more tightly organized or rigid structure (Figure 4).

In the context of the current study, alpha was calculated for each phase (i.e. acquisition and reversal phases) of the learning task separately, with a mean IBI length of 347 and standard deviation of 67. Based on previous recommendations (Ihlen, 2012; Berry et al., 2019) and constraints of the current data, the following DFA parameters

were defined: linear detrending, minimum segment size = 10, maximum segment size = 50, and total segment sizes = 10. The regression range (i.e. slope estimation indexing how fast the overall RMS of local fluctuations grows with increasing segment sample size) was set to capture both short and long range fluctuations in the time series. The mean and standard deviations of the DFA-derived alpha scores for each phase of the task were used to descriptively examine the extent of stability and change across the task and inform subsequent modeling decisions.

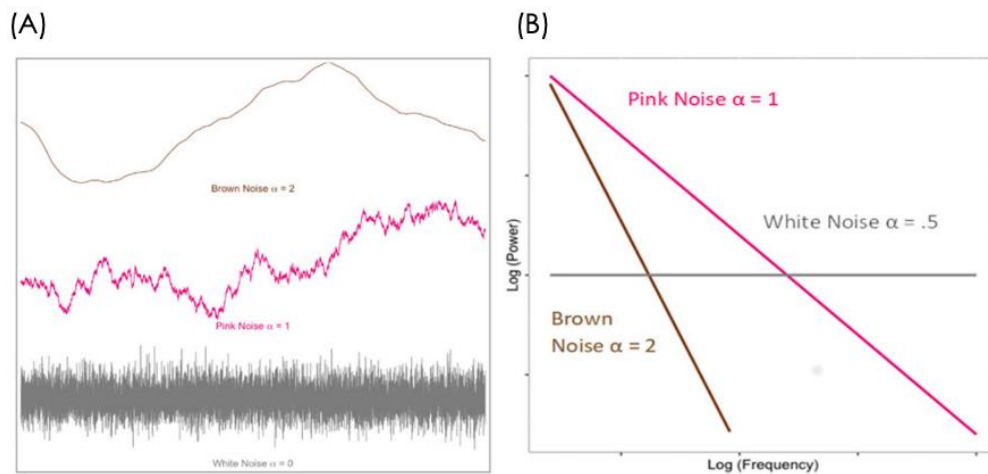


Figure 4. (A) Simulated illustration of systems with varying degrees of organization and flexibility. Highly rigid organization is represented in brown, flexible organization is represented in pink, and random organization is represented in gray. (B) Relationship between power and frequency for a highly flexible (pink) system.

Applying the same Bayesian multilevel modeling steps used in Aim 1, I examined whether and how cognitive learning processes are associated with individual differences in cardiac complexity across the non-emotional and socioemotionally-salient task conditions. Specifically, participants' alpha scores were regressed on task version and each computational-model parameter (reward and punishment learning rates, inverse temperature) as main effects. Interactions between computational parameters and

version were examined in subsequent models and model comparisons determined the final model.

Priors in the full interaction model (most complex fitted model) were defined as follows:

cardiac fractality (α) \sim 1 + punishment learning rate * version + reward learning rate *
version + inverse temperature * version + (1+ version |subject)

```
alpha ~ student's t( $\mu_i$ ,  $\sigma_e$ )
   $\mu_i = \alpha_{\text{subject}[i]} + \beta * \text{version}_{\text{subject}[i]} + \beta * \text{learning parameter}$ 
   $+ \beta * \text{learning parameter: version}$ 
 $\alpha_{\text{subject}} \sim \text{normal}(.85, .5)$ , lb = 0, ub =2
 $\sigma_{\text{subject}} \sim \text{inverse-gamma}(3, .6)$ 
 $\beta \sim \text{normal}(0, 1)$ 
 $\sigma_{\text{version}} \sim \text{inverse-gamma}(3, .6)$ 
 $\sigma_e \sim \text{inverse-gamma}(3, .7)$ 
 $\mathcal{R} \sim \text{LJKcorr}(2)$ 
 $N \sim \gamma(2, .1)$ 
```

3.5.3 Deviations from pre-registration

I deviated from the analytic plan proposed in the pre-registration in two notable ways due to unanticipated constraints in the data as well as finding improved Bayesian modeling strategies. First, for Aim 1a, I originally proposed analyzing the percentage of correct hits in the task, separated by acquisition and reversal phases, to index task performance. Upon further reflection, this index arguably did not seem to capture “optimal” performance, and thus a more holistic scoring approach that leveraged all types of responses throughout the task was used. Second, the pre-registration described using uninformative priors for the Bayesian multilevel models. While learning more about various approaches to defining priors, I found that a principled approach to establishing informative priors that account for uncertainty was more appropriate.

Chapter 4: Results

The final sample for Aim 1 included behavioral data from 56 youth. Due to a BIOPAC device malfunction in data collection, 55 youth provided psychophysiological data for Aim 2. A correlation matrix containing variables used in analysis can be found in supplemental Figure S4. Additional descriptives (Figures S2) and figures and tables referenced below can be found in supplementary material.

Bayesian prior and posterior sensitivity checks (Figures S5 - S7 for behavioral outcomes; Figures S8 - S16 for computational outcomes; Figures S17- S19 for psychophysiological outcomes) can also be found in supplementary material.

4.1 Aim 1: Learning behavior and underlying cognitive processes across task versions

4.1.1 Behavioral learning performance as a function of task context

Bayesian multilevel regression analyses found that on average, youth performed (via summation of correct hits minus false alarms) slightly more optimally in the socioemotionally-salient version compared to the non-emotional version of the learning task ($B = 6.42$, 95% CrI = [-1.42, 14.22], $d = 0.22$) (Figure 5). Random effects of intercept and slope demonstrated individual heterogeneity around this average effect (sd intercept = 20.24, 95% CrI = [13.11, 27.67]; sd slope = 13.47, 95% CrI = [6.72, 23.29]). In other words, accounting for individual differences, learning from socioemotional stimuli is likely to have an overall small positive effect on task performance for youth, with a mean increase of 6.42 points and a range of possible values from -1.42 to 14.22

reflective of substantial uncertainty. Hypothesis testing demonstrated moderately strong evidence in support of the hypothesis that the population average effect of task version is greater than 0, with a posterior probability of 95% (i.e., percentage of posterior distribution that crosses 0). Log-transformed evidence ratios (see Analytic Plan above) indicated that there was approximately four times more evidence (i.e., falling within the “modest” range of support) that the socioemotionally-salient version results in higher performance than the null hypothesis (i.e., no difference between task version).

The best fitting model included an interaction between version and age ($B = 5.57$, 95% CrI = [-2.93, 14.01]) (Figure 5), indicating that youth tended perform slightly more optimally in the socioemotionally-salient condition than the non-emotional condition with increasing age. Examination of simple slopes via Bayesian hypothesis tests demonstrated greater certainty for the positive relation in the socioemotional condition (posterior slope = 9.38, [0.69, 17.93], 57% two-sided posterior probability) compared to the non-emotional condition (posterior slope = 3.81, [-3.89, 11.75], 90% one-sided posterior probability).

See Table 2 for final model results.

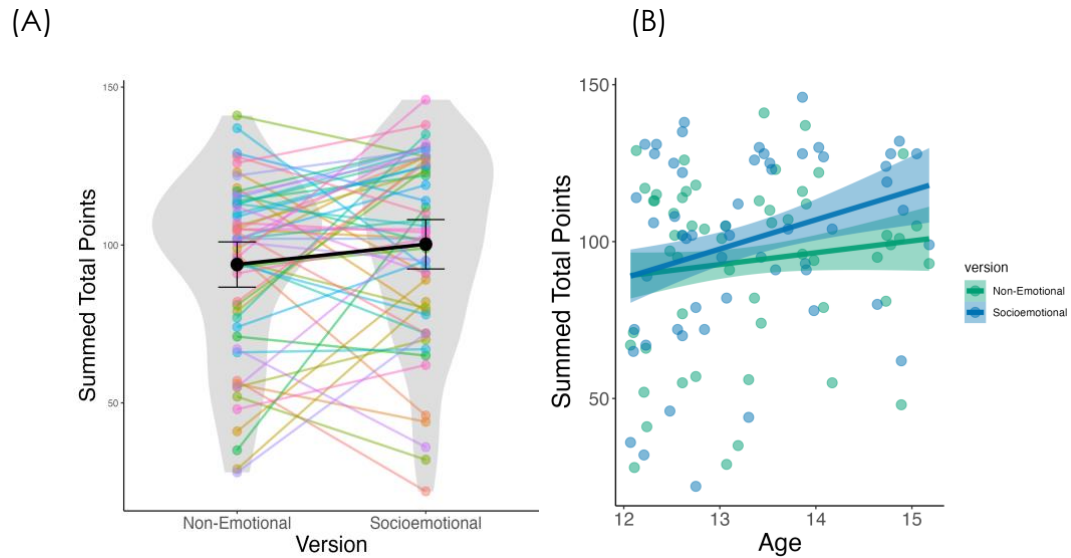


Figure 5. (A) Main effect of learning task version from Bayesian multilevel models on total points, indexing the degree of optimal performance (left). Black line indicates the predicted average fixed effects of task version with 95% credible intervals, superimposed on the raw participant-level data for the total points outcome. (B) Interactive effect of age and learning task version superimposed on raw data points.

Table 2. Final Bayesian mixed effect models predicting behavioral performance (total summed points) on the learning task.

	Posterior Mean	SE	95% Credible Interval	Rhat
<i>Fixed effects</i>				
(Intercept)	93.84	3.66	[86.65, 100.99]	1.00
Version (socioemotional)	6.42	3.95	[-1.42, 14.22]	1.00
Age	3.81	4.00	[-3.89, 11.75]	1.00
Age x Version	5.57	4.32	[-2.93, 14.01]	1.00
<i>Random effects</i>				
(Intercept)	20.24	3.71	[13.87, 30.71]	1.00
Version (socioemotional)	13.47	4.34	[0.51, 28.83]	1.00
Cor (Intercept, Version)	-0.04	0.33	[-0.70, 0.71]	1.00

4.1.2 Cognitive learning processes as a function of task context

4.1.2.1 Computational model results

The results of the reward-punishment reinforcement learning model are presented in Table 3 (see Figure S20 for convergence checks and S21 for descriptive histograms). Descriptively, average punishment and reward learning rates for this age group were in the moderate range (Eckstein et al., 2022), with slightly elevated rates in the socioemotional version indicating potential task-specific modulation. Within task version (non-emotional versus socioemotional), learning rates were slightly higher for reward than punishment learning rates, reflective of greater sensitivity and thus faster value-updating to learning from rewards. Inverse temperature was on the lower end and nearly identical across versions, indicating the difference in the value of the options (press versus not press) for a given stimulus was minimal and thus resulted in values reflective of greater choice stochasticity irrespective of task version, on average. Models demonstrated good mixing and convergence, as indicated by trace plots (Figure S20) and Rhat values being at or close to 1.

Table 3. *Reward-punishment computational model results across learning task version.*

	Non-Emotional Version					Socioemotional Version				
	Posterior Mean	SD	2.5% CI	97.5% CI	Rhat	Posterior Mean	SD	2.5% CI	97.5% CI	Rhat
Punishment learning rate	0.39	0.04	0.31	0.48	1.00	0.54	0.05	0.44	0.63	1.00
Reward learning rate	0.65	0.05	0.56	0.74	1.00	0.76	0.04	0.68	0.84	1.00
Inverse temperature	2.09	0.09	1.91	2.27	1.00	2.09	0.09	1.91	2.27	1.00

4.1.2.2 Bayesian multilevel models predicting computational parameters from task version

Table 4 presents final multilevel model results for each computational model parameter outcome, Table 5 presents model comparisons across fitted models, and supplementary Figures S8-S16 provide prior-posterior checks.

Table 4. Final Bayesian mixed effect models predicting computational model parameters: punishment learning rate, reward learning rate, and inverse temperature .

	Punishment learning rate				Reward learning rate				Inverse temperature			
	Posterior Mean	SE	95% Credible Interval	Rhat	Posterior Mean	SE	95% Credible Interval	Rhat	Posterior Mean	SE	95% Credible Interval	Rhat
<i>Fixed effects</i>												
(Intercept)	0.42	0.03	[0.36, 0.49]	1.00	0.63	0.03	[0.57, 0.69]	1.00	2.13	0.07	[1.99, 2.27]	1.00
Version (socioemotional)	0.11	0.04	[0.04, 0.18]	1.00	0.10	0.03	[0.03, 0.16]	1.00	0.01	0.09	[-0.16, 0.18]	1.00
Age	0.03	0.04	[-0.04, 0.10]	1.00					-0.16	0.08	[-0.31, -0.01]	1.00
Age x Version	-0.07	0.04	[-0.15, 0.01]	1.00					0.25	0.09	[0.07, 0.43]	1.00
<i>Random effects</i>												
(Intercept)	0.19	0.03	[0.12, 0.25]	1.00	0.15	0.03	[0.09, 0.21]	1.00	0.33	0.09	[0.16, 0.50]	1.00
Version (socioemotional)	0.14	0.04	[0.07, 0.23]	1.00	0.13	0.04	[0.07, 0.20]	1.00	0.23	0.10	[0.08, 0.48]	1.00
Cor (Intercept, Version)	-0.26	0.30	[-0.72, 0.47]	1.00	-0.66	0.21	[-0.94, -0.14]	1.00	0.12	0.37	[-0.57, 0.82]	1.00

Models predicting punishment and reward learning rates demonstrated similar patterns supporting initial hypotheses. On average, youth exhibited moderately higher punishment ($B = 0.11$, 95% CI = [0.04, 0.18], $d = 0.46$) and reward ($B = 0.10$, 95% CI = [0.03, 0.16], $d = 0.52$) learning rates in the socioemotionally-salient version compared to the non-emotional version of the learning task (Figures 6 and 7). Random effects of intercept and slope demonstrated individual heterogeneity around this average effect (Table 4). Hypothesis testing with each learning rate outcome demonstrated strong evidence (100% posterior probability) in support of the hypothesis that youth’s value-updating tends to be more rapid in the context of learning from socioemotionally-salient stimuli. Specifically, log-transformed evidence ratios indicated approximately 14 times more evidence and 12 times more evidence (i.e., falling within the “very strong” range of support) than the null hypothesis that the socioemotionally-salient version resulted in higher punishment and reward learning rates, respectively.

In addition to main effects of task version, the best fitting model for punishment learning rate included an interaction between version and age ($B = -0.07$, 95% CrI = [-0.15, 0.01]) (Figure 7), indicating that younger youth tended to show higher punishment learning rates in the socioemotional condition than the non-emotional condition.

Examination of simple slopes via Bayesian hypothesis tests demonstrated a positive relation between age and punishment learning rates in the nonemotional condition (posterior slope = 0.03, [-0.04, 0.10], 95% one-sided posterior probability) whereas a negative relation was demonstrated in the socioemotional condition (posterior slope = -0.04, [-0.11, 0.04], 95% two-sided posterior probability). For reward learning rate, none of the models including age improved model fit.

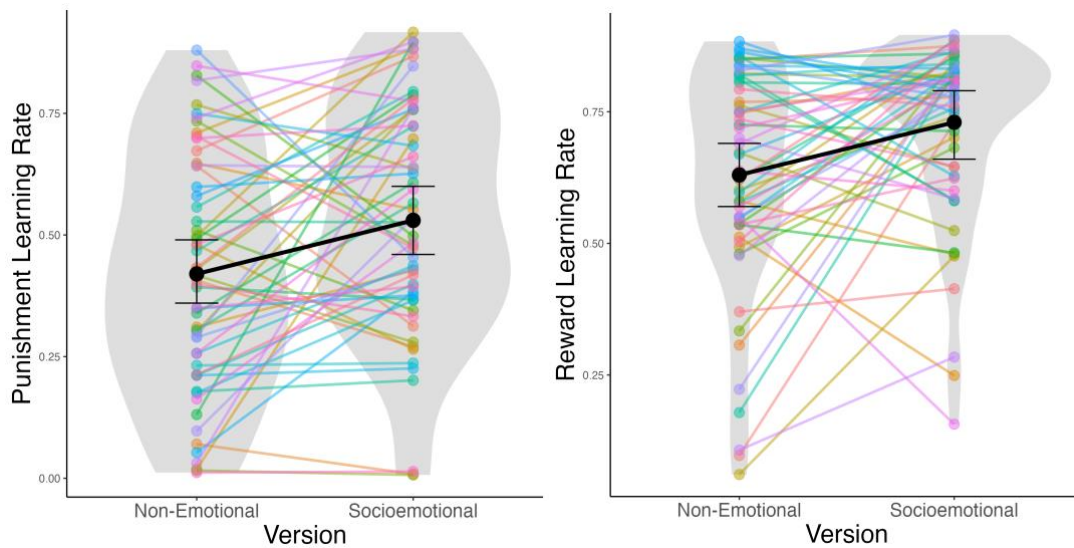


Figure 6. Main effect of learning task version from Bayesian multilevel models on punishment learning rate (left) and reward learning rate (right). Black line indicates the predicted average fixed effects of task version with 95% credible intervals, superimposed on the raw participant-level data for each outcome.

Contrary to hypotheses, models predicting inverse temperature did not demonstrate a statistically meaningful main effect of task version ($B = 0.01$, 95% CI = [-0.16, 0.18], $d = 0.02$). However, the best fitting model included an interaction between version and age ($B = 0.25$, 95% CI = [0.07, 0.43]) (Figure 7), indicating that, with increasing age, youth tended toward relatively higher inverse temperatures, or more exploitative decision-making strategies, in the socioemotionally-salient condition. In contrast, youth demonstrated lower inverse temperatures, or relatively more exploratory strategies, in the non-emotional condition with increasing age. Examination of simple

slopes via Bayesian hypothesis tests demonstrated greater certainty for the negative relation in the non-emotional condition (posterior slope = -0.16, [-0.31, 0.01], 61% two-sided posterior probability) compared to the socioemotionally-salient condition (posterior slope = 0.09, [-0.08, 0.26], 90% one-sided posterior probability).

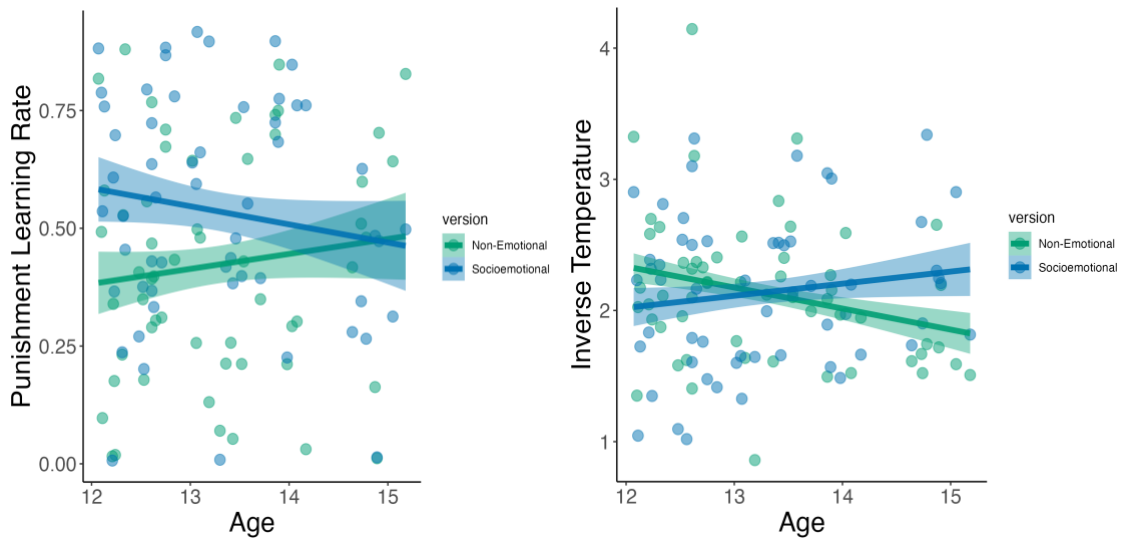


Figure 7. Interactive effect of age and learning task version from Bayesian multilevel models predicting punishment learning rate (left) and inverse temperature (right), superimposed on raw data points.

4.1.2.3 Aim 1 summary

Analyses of the behavioral and cognitive processes of youth’s learning revealed that youth exhibited more optimal performance and learned faster—especially from reward-based associations—in the context of more socioemotionally-salient learning demands than neutral or non-emotional contexts. Effects for punishment learning rate further varied as a function of age, such that the magnitude of difference across task versions decreased with increasing age. Decision-making strategies across contexts also varied as a function of age, such that the degree of exploitative tendencies showed

a slight increase across age in the socioemotional condition, whereas the opposite was true in the non-emotional condition.

An exploratory descriptive cross-walk examining the joint parameter space between the behavioral and cognitive parameters (Figure 8), highlights that moderate punishment learning rates combined with high inverse temperature tends to align with more optimal performance on the learning tasks overall. A slightly different pattern was observed for reward learning rates, whereby high reward learning rates combined with moderate inverse temperature tended to align with more optimal performance. These descriptive patterns seemingly generalize across task versions, but with greater variation in the socioemotional condition. See supplemental Figure S22 for similar descriptive plots as a function of age.

Table 5. Model comparisons for fitted Bayesian multilevel models in Aim 1 .

	Total points		Punishment learning rate		Reward learning rate		Inverse temperature	
	WAIC	LOOIC	WAIC	LOOIC	WAIC	LOOIC	WAIC	LOOIC
<i>Fitted model</i>								
1+ version + (1+ version subject)	1027.2	1041.6	-34.5	-17.3	-67.7	-58.5	179.3	186.0
1+ version + age + (1+ version subject)	1026.8	1040.6	-34.6	-17.2	-67.7	-58.7	180.5	187.1
1+ version + age + age ² + (1+ version subject)	-	-	-	-	-67.5	-57.4	-	-
1+ version + age + age: version + (1+ version subject)	1025.7	1039.9	-35.5	-18.1	-68.2	-57.9	167.8	176.8

Note. Quadratic age effects were only examined if descriptive plots indicated a potential quadratic relation with the outcome.

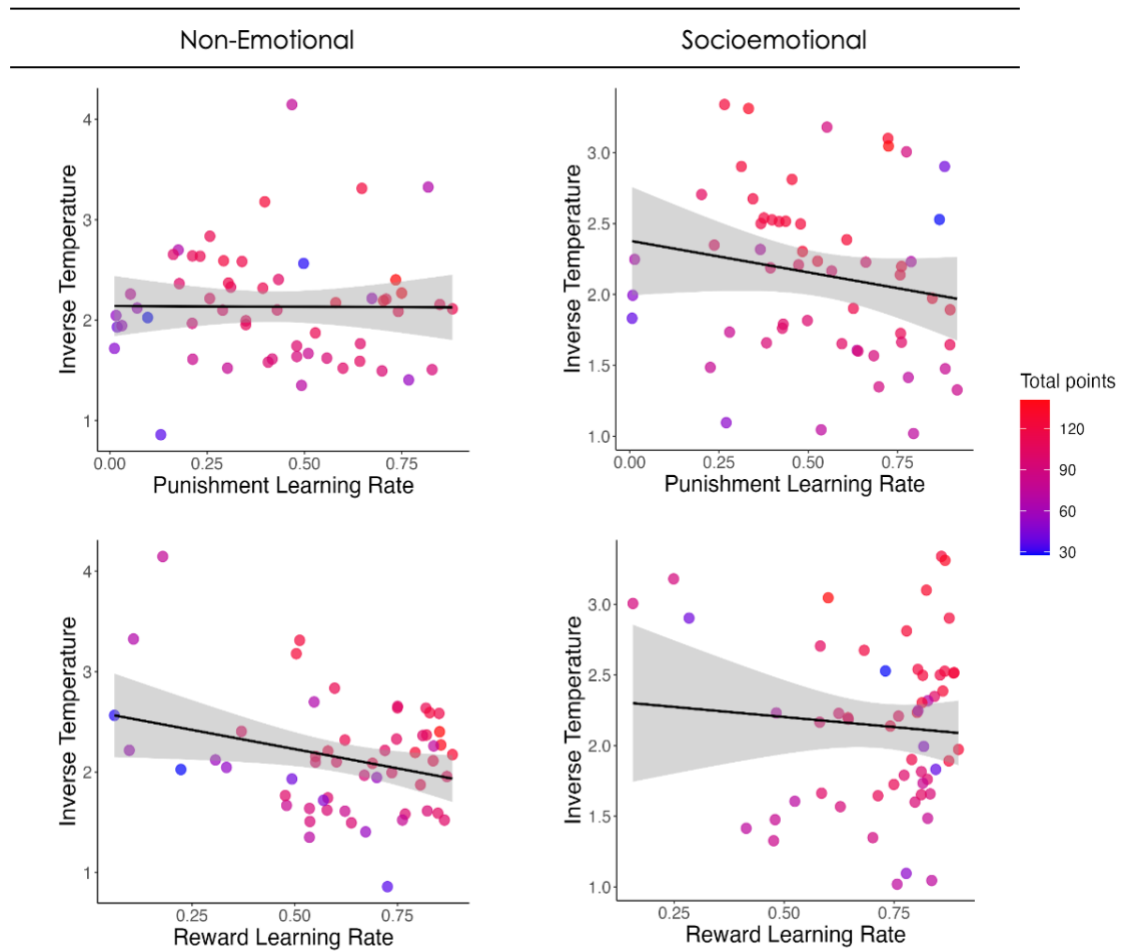


Figure 8. Within-version (left: non-emotional; right: socioemotional) correlations across learning rates and inverse temperature parameters colored by total points.

4.2 Aim 2: Psychophysiological engagement as a function of cognitive process and task context

Example IBI and DFA plots for randomly selected youth can be found in Supplement Figure S23. While there were individual differences in the scores (Figure S24 and S25), average scores for each task phase were relatively stable and thus supported the pre-registered plan of modeling scores collapsed across time. Average scores were in the ‘pink noise’ range, indicating moderately flexible and organized

physiology (mean range: .82 - .86; standard deviation range: .18 - .21). One participant demonstrated very low alpha scores across the task (range .3 - .45) indicative of highly random, or 'white' noise; thus the multilevel model presented below was generated without this outlier as a sensitivity check. Sensitivity analyses indicated that estimates were nearly identical (i.e., within 0.01 or less difference) (see supplementary Table S1) and thus findings from models conducted on the full sample are presented.

The best fitting Bayesian multilevel model predicting average cardiac complexity (alpha) scores included main effects of version, computational parameters, and age (Table 6). However, only the main effect of reward learning rate demonstrated a statistically meaningful effect with a relatively high degree of certainty ($B = 0.22$, 95% CrI = [0.09, 0.35], $\beta = .29$). Specifically, increases in reward learning rate are associated with a 0.22 increase, on average, in cardiac complexity, with a range of possible values from 0.09 to 0.22. Hypothesis testing demonstrated strong evidence in support of this moderately positive effect, with a posterior probability of 100% (i.e., percentage of posterior distribution that crosses 0). Log-transformed evidence ratios indicated that there was approximately 15 times more evidence (i.e., falling within the "very strong" range of support) that higher reward learning rates resulted in greater cardiac complexity than the null hypothesis (i.e., zero or negative association).

The degree of support for the average main effects of task version and the other computational model parameters were also examined. First, task version demonstrated a small positive effect ($B = 0.01$, 95% CrI = [-0.03, 0.05]), with a 69% posterior probability of being above zero (compared to a 31% posterior probability of being below zero). In other words, cardiac complexity during learning was only somewhat likely to be higher in socioemotional condition compared to the non-emotional condition (Figure 9). Scaling on their respective within-person variances, youths' cardiac complexity (as

indexed by alpha) was approximately 0.07 standard deviations higher during the socioemotional condition than the non-emotional condition. This finding is minimal in absolute terms, but provides partial support for initial hypotheses; nonetheless, no strong conclusions can be drawn based on this level of uncertainty in the estimated effect.

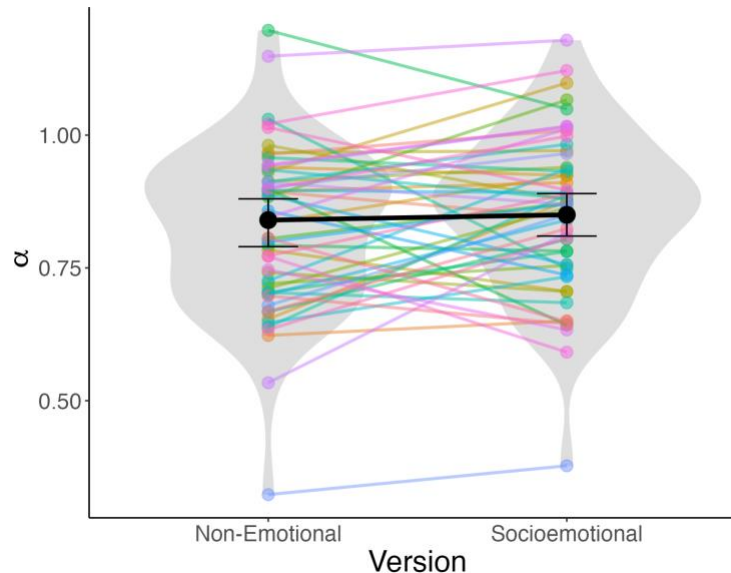


Figure 9. Main effect of learning task version from Bayesian multilevel models predicting psychophysiological (cardiac) complexity indexed via alpha (α). Black line indicates the predicted average fixed effects of task version with 95% credible intervals, superimposed on the raw participant-level data for each outcome. Note: this effect was not statistically meaningful but is presented for descriptive purposes (see text and Table 6 for more details on the level of support for this effect).

Second, with respect to the remaining computational parameters, a small negative effect was demonstrated for punishment learning rate ($B = -0.02$, 95% CrI = [-0.13, 0.10], $\beta = -0.03$), with a 63% posterior probability supporting this negative effect. In other words, contrary to reward learning rate, higher punishment learning rates (irrespective of socioemotional salience of the stimuli) was associated with small declines in cardiac complexity. Inverse temperature demonstrated a small positive effect (77% posterior probability of being above zero) suggesting slight increases in cardiac

complexity as decision-making becomes more exploitative ($B = 0.02$, 95% CrI = [-0.03, 0.07], $\beta = 0.08$). Of note, a model that included a quadratic effect of inverse temperature showed nearly identical model fit to the main effects model (Table 6), which provides preliminary support for the possibility of an inverted-U relation as initially hypothesized.

Lastly, youth age demonstrated a small positive association with cardiac complexity ($B = 0.04$, 95% CrI = [-0.01, 0.08], $\beta = 0.24$) with a 96% posterior probability supporting this effect. In other words, increases in age may be matched with slight increases in cardiac complexity reflective of psychophysiological systems maturing over developmental time.

See Figure 10 and Table 6 for the aforementioned main effect relations and supplementary Figure S26 for relations colored by total points to explore qualitative links with behavioral performance.

Table 6. Final Bayesian mixed effect model predicting psychophysiological engagement (α).

	Posterior Mean	SE	95% Credible Interval	Rhat
<i>Fixed effects</i>				
(Intercept)	0.84	0.02	[0.79, 0.88]	1.00
Version (socioemotional)	0.01	0.02	[-0.03, 0.05]	1.00
Punishment learning rate	-0.02	0.06	[-0.13, 0.10]	1.00
Reward learning rate	0.22	0.07	[0.09, 0.35]	1.00
Inverse temperature	0.02	0.02	[-0.03, 0.07]	1.00
Age	0.04	0.02	[-0.01, 0.08]	1.00
<i>Random effects</i>				
(Intercept)	0.13	0.02	[0.10, 0.17]	1.00
Version (socioemotional)	0.08	0.02	[0.05, 0.11]	1.00
Cor (Intercept, Version)	-0.26	0.26	[-0.70, 0.34]	1.00
<i>Model comparisons</i>		WAIC	LOOIC	
1+ version + (1+ version subject)		-182.8	-158.1	
1+ version + rew + pun + temp + (1+ version subject)		-202.8	-174.3	
1+ version + rew + pun + temp + temp ^2 + (1+ version subject)		-201.8	-173.7	
1+ version + rew + pun + temp + age + (1+ version subject)		-203.3	-175.8	
1+ version + rew + pun + temp + rew:version + pun:version + temp:version + age:version + (1+ version subject)		-195.9	-166.5	

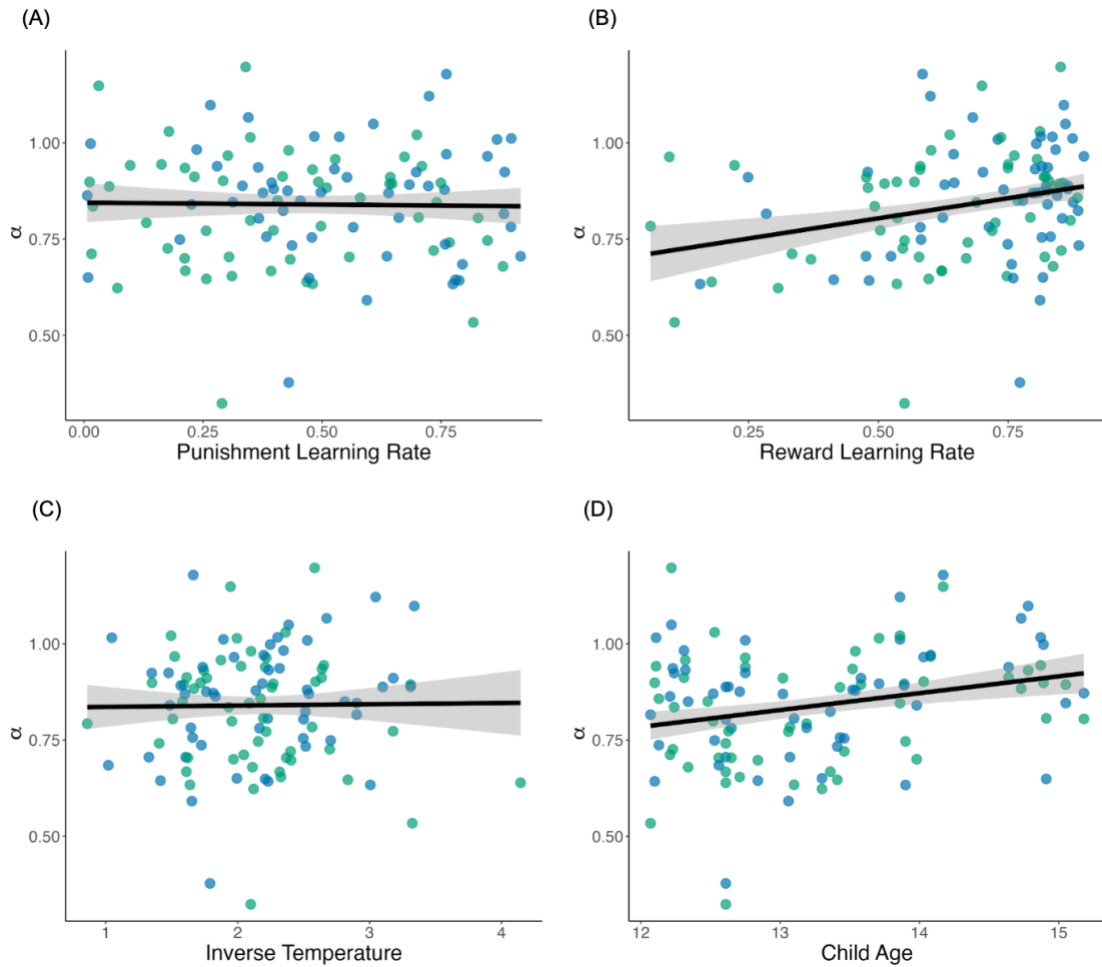


Figure 10. Main effects of (A) punishment learning rate, (B) reward learning rate, (C) inverse temperature, and (D) child age from Bayesian multilevel models predicting psychophysiological (cardiac) complexity indexed via alpha (α). Raw data points are presented in the background, where colors in plot A represent individual subjects and colors in plots indicate task version (green = non-emotional; blue = socioemotional). Note: only reward learning rate (B) and age (D) had a posterior probability of >95%. Other effects (A, C) are presented for descriptive purposes (see text and Table 6 for more details on the level of support for each effect).

Chapter 5: Discussion

Early adolescence marks a formative developmental transition that includes heightened engagement with novel socioemotional contexts that youth learn to navigate by drawing upon foundational value-based reinforcement learning (RL) processes. A rapidly growing body of research has begun elucidating the neurocognitive and biological mechanisms that help explain how youth learn from the predictability of environmental cues and their associations with rewards and punishments to guide decision-making. While studies have begun examining the role of developmentally-relevant modulatory processes in learning (e.g., role of volatility [Eckstein et al., 2022b]), there are several additional modulatory mechanisms that remain to be tested. The current pre-registered study sought to advance our understanding of these complex value-based reinforcement learning processes in youth aged 12-15 years old. Grounded in a multilevel developmental framework, we examined the role of socioemotional salience on youth's overall behavioral performance as well as underlying cognitive and psychophysiological mechanisms posited to underlie learning. To do this, we leveraged a within-person experimental design and drew upon recent methodological advances in neurocognitive and dynamic systems modeling that allowed us to capture more theoretically-informed processes of reinforcement learning.

Broadly, results demonstrated that learning from socioemotionally-salient stimuli, compared to non-emotional stimuli, enhanced behavioral performance and elicited greater sensitivity to learning from rewards and punishments. Older youth also evidenced relatively more exploitative decision-making strategies in the socioemotional learning context. Intriguingly, youth psychophysiology was flexibly organized irrespective of task context and was uniquely linked to enhanced value-updating in the context of

rewards specifically. Findings from this study build upon existing work to provide new insights into the context-specific modulatory mechanisms through which learning unfolds in early adolescence.

5.1 Socioemotional salience enhances youths' behavioral performance and underlying cognitive learning strategies (Aim 1)

Our first aim sought to determine whether stimulus content (non-emotional versus socioemotionally-salient) elicits different levels of behavioral performance and computational model-derived learning processes. Consistent with hypotheses, youths' performance on the task was slightly more optimal—indexed via a summation of correct hits minus false alarms—in the socioemotional version of the learning task. Aligned with the priority state space (PSS; Todd & Manaligod, 2017) theoretical framework, this finding suggests that youth attended to, and thus learned better from, stimulus-outcome associations that are more affectively-salient. Prior work adopting within-person designs in adult samples has shown similar behavioral findings from learning tasks, including greater accuracy for emotionally negative and positive stimuli compared to neutral stimuli (Everaert et al., 2020; Plate et al., 2022). Together, this line of evidence suggests that compared to benign or contextually-neutral information, socioemotionally-salient information may mobilize greater processing resources to facilitate learning (Todd et al., 2012; Schwabe et al., 2008).

We additionally observed a small task version by age interaction that raises the possibility that enhanced learning performance from socioemotionally-salient stimuli may increase in magnitude in later adolescence. More specifically, while total points tended to show the expected increase across age irrespective of task version, the slope for the

socioemotional version was steeper than the non-emotional version and showed a greater degree of certainty. Combined with work in adults (e.g., Everaert et al., 2020; Plate et al., 2022), this finding suggests the intriguing possibility that context-specific performance enhancements are present in mid adolescence and potentially continue increasing or plateauing into adulthood. Indeed, throughout adolescence, youth become increasingly less reliant on scaffolding from their caregivers and thus must learn to expertly navigate and make decisions across a wide range of novel social and emotional situations (Casey & Galvan, 2008).

Rapid neurodevelopmental changes in adolescence afford the recruitment of higher order cognitive resources (Crone et al., 2007) and increasingly complex learning strategies (DePasque & Galvan, 2017) that may explain the behavioral performance advantage observed in the socioemotional task condition. To reveal these learning mechanisms thought to underlie behavior, we applied computational RL modeling that estimated several parameters indexing value-updating and decision-making strategies (Ahn et al., 2017; Nussenbaum & Hartley, 2019).

Beginning with our findings for value-updating processes, youth demonstrated moderately higher punishment and reward learning rates in the socioemotionally-salient compared to the non-emotional version of the learning task. In other words, on average, youth were more sensitive to—and thus exhibited faster value-based updates—when learning from socioemotional stimuli. Effects for punishment learning rate further varied as a function of age, such that the magnitude of difference across task versions decreased with increasing age. These findings are consistent with recent evidence showing that learning rates (both within and between individuals) tend to be specific to task demands (Eckstein et al., 2022). Moreover, tasks applying emotional stimuli have been recently demonstrated to elicit faster learning (i.e., more rapid asymptotic

accuracy), perhaps illustrative of the ability of emotional arousal to assist in resolving ambiguity or uncertainty in the learning environment (Wu et al., 2021; Plate et al., 2022; Walle et al., 2017). In the current study, task demands (e.g., reward and reversal schedules) were the same across task versions with the main difference lying in the affective and socioemotional salience of the stimuli to be learned. Given that half of the reinforcement contingencies reversed every 96 trials, an optimal strategy would be to learn moderately quickly from rewards, as reflected by higher learning rates as opposed to low learning rates (Zhang et al., 2020). This allows learners to reach asymptotic accuracy faster while still allowing for flexible shifts in value updating as some stimuli shift their reward structure. Supporting initial hypotheses, socioemotional cues used in the current study may have required greater attentional resources and thus perhaps more readily elicited faster learning rates to need to achieve better performance (Todd & Manaligod, 2018).

With respect to cognitive strategies youth use to transform learned values to choices via the inverse temperature parameter, youth demonstrated an increase in relatively more exploitative (and less exploratory) decision-making in the socioemotional version across age. The opposite was true in the non-emotional version. In other words, youths' choices about whether to click (or not) on a given stimulus were more likely to be dictated by the potential to yield reward when the stimulus content was socioemotionally-salient. When exposed to benign or neutral learning stimuli, youth were less sensitive to value-based differences between choices and thus adjusted their choice behavior less. Of note, the slope for the non-emotional version was steeper than the socioemotional version and showed a greater degree of certainty. Surprisingly, this negative relation observed for the non-emotional version does not align with past work showing increases in inverse temperature with age (Eckstein et al., 2022; Christakou et al., 2013; Decker et

al., 2015; Giron et al., 2022; Javadi et al., 2014; Palminteri et al., 2016; Rodriguez Buritica et al., 2019). Rather, inverse temperature only increased with age in the socioemotional version. Given the uncertainty in these effects, future work is necessary to corroborate these findings. Nonetheless, it challenges the notion that inverse temperatures are generalizable across task demands (Eckstein et al., 2022; Nussenbaum & Hartley, 2019) by raising the possibility that socioemotional salience may elicit slightly more mature decision-making strategies. This may be driven by a prioritized attunement to maximizing rewards particularly when the decisions involve affective cues or cues eliciting prior experiential knowledge (i.e. the prioritized state space per the PSS; Todd & Manaligod, 2018).

Descriptive examinations of the joint parameter space and behavioral performance provided additional insights into the multilevel processes at play when youth learn in different experimental contexts. Specifically, moderate levels of punishment learning rates combined with high levels of inverse temperature tended to align with more optimal performance irrespective of task version. A slightly different pattern was observed for reward learning rate, whereby high levels of reward learning rates combined with moderate levels of inverse temperature tended to align with more optimal performance. These divergent joint parameter patterns suggest that youth deployed slightly different combinations of cognitive strategies depending on whether outcomes were negatively or positively reinforced. In the face of worse-than-expected outcomes, youth showed more optimal performance when adopting more moderate (i.e., slower) value updating while making more exploitative choices that avoided punishing outcomes. Speculating on what this may mean, youth may have employed greater caution, and accounted for past experience more, when inferring the value of negatively reinforced stimuli and induced less exploratory decision-making. In contrast, in the face

of better-than-expected outcomes, youth showed more optimal performance when adopting higher (i.e., faster) value updating while balancing exploration and exploitation. Perhaps youth were more confident in their value-estimates, and weighed recent experience more, when inferring the value of positively reinforced stimuli and thus induced a relatively more balanced explore-exploit decision-making strategy. Notably, these descriptive patterns seemingly generalized across task versions, but with greater variation in the socioemotional version—perhaps reflective of variation in prior lived experiences that may prioritize youth’s attention to various real-world stimuli differently (Ciranka & van den Bos, 2019; 2021).

5.2 Youth psychophysiology is flexibly organized irrespective of task context and is linked to reward learning (Aim 2)

Our second aim sought to explore whether psychophysiological engagement exhibited qualitative differences as a function of task version and cognitive RL processes derived from aim 1. This aim was informed by burgeoning theory that the locus coeruleus norepinephrine (LC-NE) and interwoven autonomic neuromodulator systems integrates and reconciles the influence of reinforcing sources of salience such as rewards, punishments, and type of stimuli elicited by a learning task and shaped by past experience (Todd & Manaligod et al., 2017; Thayer et al., 2009). Given that early adolescence is characterized by heightened sensitivity to socioemotionally-salient situations (e.g., Fuhrmann et al., 2015), we speculated that LC-NE and autonomic system modulation would stimulate greater psychophysiological engagement in similarly affective contexts compared to non-emotional or benign contexts. To capture the complex and emergent psychophysiological dynamics occurring across these

neuromodulating systems, we adopted time-series modeling from the dynamic systems literature (Thelen, 2012), allowing us to derive a person-specific summary index (alpha) of temporal cardiac organization/complexity across the learning task.

Contrary to our hypothesis that youth would demonstrate more moderate ranges of cardiac complexity in the socioemotional version indicative of greater arousal and engagement with task demands, cardiac complexity was only descriptively likely to be more temporally organized when learning from socioemotional compared to non-emotional stimuli. Rather, youth evidenced flexibly organized patterns of cardiac fractality irrespective of stimulus content. Given past work demonstrating that psychophysiology in this fractal 'pink' noise range is indicative of healthy system dynamics (e.g., Peng et al., 1995; Wijants, 2014) and greater behavioral self-regulation (Berry et al., 2019), we speculate that attentional arousal elicited from the socioemotional stimuli may not have been powerful enough to significantly disrupt homeostatic levels of psychophysiological engagement. Similarly, perhaps both types of stimuli were equally arousing. Thus, it is unlikely that psychophysiology served as a meaningful modulator in regulating behavior and cognition in the current study. However, given that most of the youth in our sample exhibited alpha scores in this 'pink' noise range, it is likely that the sociodemographic homogeneity of our convenience sample highly constrains the inferences we can draw here. Another possible explanation may be that learning in affective contexts relies on other neuromodulatory systems and/or processes that our summary physiological index was unable to fully capture. For example, perhaps heart rate was too peripheral, and more direct measures of brain activation via neuroimaging would have given a level of precision needed to measure meaningful differences across experimental contexts. Indeed, prior studies, albeit primarily in adults, have demonstrated differential patterns of brain network activation

depending upon whether learning stimuli were emotional (e.g., Nashiro et al., 2012) or whether a pre-learning stressor occurred before the task (e.g., Schwabe et al., 2008). More research is needed to better understand these neuromodulatory processes in youth as they relate to learning across a range of contexts.

We further explored the link between psychophysiological engagement and the RL computational parameters from Aim 1 to examine the possibility that cognitive processes more broadly, or in interaction with task stimulus content, may be bi-directionally associated with organizational dynamics of related neuromodulator systems. Several exploratory hypotheses were proposed based on theory, but the only statistically meaningful effect observed, with a strong degree of certainty, was a main effect of reward learning rate. Specifically, reward learning rate was associated with moderate increases in cardiac complexity, supporting the hypothesis that faster value-updating processes in the context of reward elicits greater attention and thus relies on the recruitment of higher psychophysiological modulatory engagement and coordination. This finding partially aligns with prior research in infants that showed greater attentional system complexity (i.e., alpha scores derived from eye-tracking data) in the context of social stimuli (Stallworthy et al., 2020). Situating these and our findings in a developmental context, social interaction is highly rewarding in infancy and thus will be reflected in greater engagement of the necessary systems when viewing social stimuli in an experimental context. In the context of adolescence, although social and emotional situations are arguably more attentionally salient during this developmental stage, youths' enhanced sensitivity to rewards more broadly may be a more important driver in eliciting physiological engagement necessary for such levels of attention (Ciranka & van den Bos, 2019). Youth may also recruit greater physiological resources when exerting more cognitive effort while learning from rewards (Galvan, 2010; 2013). In other words,

perhaps closely updating and monitoring the value of reward-based associations was particularly challenging or motivating for youth. This possibility aligns with past work showing greater cardiac complexity when children are actively challenged by a cognitive task (Berry & Stallworthy, 2018) as well as theory and neurodevelopmental evidence highlighting the adaptive function of heightened striatal reactivity when learning from rewards (Cohen et al., 2010; Telzer, 2016).

The absence of effects for the other computational RL parameters may be due to methodological limitations (see below) or the possibility that the cognitive processes that punishment learning rate and inverse temperature index merely do not require significant psychophysiological resources. Perhaps it is the case that behavioral and cognitive shifts in the face of changing learning demands are all that is required, and further adapting one's physiology to such changing demands could even be too energetically expensive to be useful or sustainable. Nonetheless, a growing body of evidence suggests there are distinct neural correlates of the cognitive processes examined here, including activation of frontostriatal brain networks (e.g., Braams et al., 2015; DePasque & Galvan, 2017; Palminteri et al., 2016; O'Doherty et al., 2001). In theory, our summary cardiac complexity index should have at least roughly captured emergent dynamics of such networks given their close neuroanatomical and functional links with autonomic activity, and thus future work is needed to probe relations examined here.

5.3 Limitations and future directions

This study has notable strengths including the within-person study design, the use of more ecologically-meaningful learning stimuli, and the adoption of relatively novel computational and dynamic systems methods that better capture theoretically-driven

multilevel processes of learning. Nonetheless, several limitations and potential paths forward are worth acknowledging. First, due to the COVID-19 pandemic, this study relied on a small convenience sample of youth and therefore limited the kinds of models that could be used as well as the capacity to draw strong between-person inferences. With over 700 observations per system (~768 task trials; ~800 cardiac observations), per individual, combined with the use of Bayesian modeling previously shown to perform well in small sample studies (Hox et al., 2012; van de Schoot et al., 2014), we were able to increase the robustness of our findings. However, the characteristics of the current sample largely limits the generalizability of our findings to White upper middle class youth. Future work with larger, more diverse samples is needed to replicate and expand upon the findings presented here. For example, one particularly critical future direction is to examine how early experience and adversity adaptively shape cognitive and psychophysiological correlates of learning across ecological contexts. Prior work has primarily demonstrated learning impairments among adverse-exposed youth (e.g., Harms et al., 2018; Hanson et al., 2017), but replicating the within-person design of the current study would afford examination into the intriguing possibility that such youth perform as well or even better than youth not exposed to adversity if given the opportunity to learn in a more ecologically-meaningful context (see Frankenhuis et al., 2020; Young et al., 2022; Humphreys et al., 2015; also see Introduction section).

There are several limitations as well as exciting future directions in complex multilevel learning processes raised by the current study. Neurocognitive computational models are relatively new in developmental science, and it remains an open question as to whether model-derived parameters accurately characterize cognitive processes in youth (Nussenbaum & Hartley, 2019). In line with our pre-registration, we limited our analyses to fitting only one commonly-used computational model adopting an RL

framework, so it is possible that there exist better fitting models with a different set of parameters that similarly explain observed behavior in this task (Eckstein et al., 2022). Relatedly, a bespoke formal theoretical model may have been the more optimal choice. However, the creation of such models requires extensive formal modeling expertise and simulation studies before adoption in an empirical context (Freek, 2021); nonetheless, this is an important future direction for this area of research that holds the potential for answering more specific questions. For instance, one interesting question that our chosen model did not allow us to test is whether youth learned differently when rewards and punishments were (in)congruent with the emotional valence of the stimuli.

Relatedly, while we were able to extract robust within-person cardiac complexity indices for each phase of the task, we lacked the variability and sample size to examine nonlinear relations as well as fixed effects of longitudinal change throughout the task (and thus collapsed across time, as pre-registered). Based on preliminary descriptive examinations of our data (see Figure S25), a preferred method would have been to use piecewise growth modeling, or growth mixture modeling, to capture potential qualitative shifts in psychophysiology when learning demands shift between acquiring new stimulus-outcome associations and reversing old ones (i.e., comparing the acquisition and reversal phases of the task across both versions). It may be the case that cognitive processes indexed by our computational parameters may differentially predict whether and how often psychophysiological engagement shifts between acquisition versus reversal demands.

Furthermore, future work should explore the use of other emerging models from the dynamic systems literature, some of which can accommodate complex relations across multiple levels and time scales (e.g., Cui et al., 2022; Hasselman, 2022; Wijnants, 2014; Heino, et al., 2021). These methods have the potential to illuminate

advanced interaction-dominant dynamics central to the idea of ‘emergence’ in a developmental systems framework, which could draw new insights into how learning unfolds in relation to behavior, physiology, and the brain. Notably, such approaches stand in contrast to computational models of learning, which takes a top-down approach to derive distinct and separable components of cognition. More theoretical and empirical work is sorely needed to advance and/or integrate these two contrasting approaches. Arguably, combining approaches may serve to maximize both explanatory and predictive insights (Yarkoni & Westfall, 2018). One practical step towards integration might be through person-centered latent profile analysis to examine patterns of covariation in various model-derived indices.

Additionally, although age was included in all of our analytical models, the cross-sectional nature of the current study precluded investigation of developmental change in learning processes over time. To our knowledge, no empirical work has prospectively examined the stability and change in cognitive and biological processes of learning examined in the present study. Other cross-sectional studies with samples across a wide age range remain mixed, but emerging evidence suggests growth in cognitive RL parameters may be nonlinear (Eckstein et al., 2022b; Giron et al., 2022; Master et al., 2020). More empirical evidence is needed to establish within- and between-person developmental patterns of change in the relationships presented here.

Finally, although the use of varied socioemotional stimuli in the current study was an improvement to prior learning studies, there are many intriguing experimental manipulations that may illuminate other key drivers of socioemotional salience in learning contexts. Our reliance on static visual images, despite being normed for high arousal, may have limited the depth of attention and arousal we aimed to elicit in the socioemotional version of the task and also are subject to featural confounds that

interact with prior life experience (Barrett et al., 2019; Todd & Manaligod 2018). Relatedly, the stimulus set of sociomemotional scenes depicted a wide range of positive and negative situations and common objects that did not afford investigation into more fine-grained stimulus categorical effects (e.g., facial versus non-facial scenes, positive versus negative scenes) and individual differences in reactions across particular categories. Given mixed prior evidence showing related to task-based performance differences in stimulus category (e.g., Plate et al., 2022; Pollak et al., 2009), decomposing the enhanced learning effects found in the current RL paradigm is an important future direction. In addition to careful consideration of stimulus content, future expansions of the current research could alter how often reinforcement contingencies reverse (Weiss et al., 2021; Yapple & Yu, 2019) to better understand the role of unpredictability in learning from nonemotional versus socioemotional conditions. A pre-task stress induction may also be an interesting future direction, as prior research has shown that stress may enhance learning performance in emotional contexts, sometimes depending upon the cognitive strategy adopted (Schwabe et al., 2007; 2008).

5.4 Conclusions

Taken together, we demonstrated that youth exhibit enhanced performance when learning in a socioemotionally-salient context, a key factor posited to be particularly relevant to the early adolescent transition to independence. This study adopted recent advances in computational and dynamic systems methods to build upon a rapidly growing area of inquiry, offering new multilevel insights into the complex processes that modulate how learning unfolds in early adolescence. Findings revealed that compared to traditionally-used benign or non-emotional stimuli, learning from stimuli

high in socioemotional arousal elicit a behavioral advantage. The use of computational modeling afforded valuable insights into the differential cognitive processes and strategies youth recruited to achieve such a behavioral advantage, demonstrating that socioemotional salience may have elicited faster value-updating processes and qualitative shifts in more exploitative decision-making. Underlying psychophysiological engagement seemed to be particularly modulated by heightened reward sensitivity, pointing to another key factor with specific relevance to this developmental period. Despite the noted limitations above, this study provides an important step in clarifying the contexts and modulatory processes that serve to enhance and support the unique ways youth learn and make decisions. Open questions remain about the adaptive utility of these various patterns of behavior, cognition, and psychophysiology across a variety of learning contexts, how they are shaped by prior lived experiences across development, and how they predict later psychosocial adjustment outcomes. Such work will shed light on how youth learn from—and adapt to—different contextual demands, with the potential to inform programs and policies that support youth’s ability to adjust to their dynamically changing social worlds.

References

- Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry (Cambridge, Mass.)*, 1, 24.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, 20(1), 1-68.
- Beckers, F., Verheyden, B., & Aubert, A. E. (2006). Aging and nonlinear heart rate control in a healthy population. *American Journal of Physiology-Heart and Circulatory Physiology*, 290(6), H2560-H2570.
- Berry, D., Palmer, A. R., Distefano, R., & Masten, A. S. (2019). Autonomic complexity and emotion (dys-) regulation in early childhood across high-and low-risk contexts. *Development and psychopathology*, 31(3), 1173-1190.
- Blair, C., & Raver, C. C. (2016). Poverty, stress, and brain development: New directions for prevention and intervention. *Academic pediatrics*, 16(3 Suppl), S30-S36.
- Bonawitz, E., Denison, S., Griffiths, T. L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: sampling in cognitive development. *Trends in cognitive sciences*, 18(10), 497-500.
- Bolenz, F., Reiter, A. M., & Eppinger, B. (2017). Developmental changes in learning: computational mechanisms and social influences. *Frontiers in Psychology*, 8, 2048.
- Braams, B. R., van Duijvenvoorde, A. C., Peper, J. S., & Crone, E. A. (2015). Longitudinal changes in adolescent risk-taking: a comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *Journal of Neuroscience*, 35(18), 7226-7238.
- Breton-Provencher, V., Drummond, G. T., & Sur, M. (2021). Locus coeruleus norepinephrine in learned behavior: anatomical modularity and spatiotemporal integration in targets. *Frontiers in Neural Circuits*, 46.
- Brown, V. M., Zhu, L., Solway, A., Wang, J. M., McCurry, K. L., King-Casas, B., & Chiu, P. H. (2021). Reinforcement learning disruptions in individuals with depression and sensitivity to symptom change following cognitive behavioral therapy. *JAMA psychiatry*, 78(10), 1113-1122.
- Buritica, J. M. R., Heekeren, H. R., Li, S. C., & Eppinger, B. (2018). Developmental differences in the neural dynamics of observational learning. *Neuropsychologia*, 119, 12-23.
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80, 1-28.
- Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental review*, 28(1), 62-77.
- Ciranka, S., & Van den Bos, W. (2019). Social influence in adolescent decision-making: A formal framework. *Frontiers in psychology*, 10, 1915.
- Ciranka, S., & van den Bos, W. (2021). Adolescent risk-taking in the context of exploration and social influence. *Developmental Review*, 61, 100979.
- Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and psychological maturation of decision-making in adolescence and young adulthood. *Journal of cognitive neuroscience*, 25(11), 1807-1823.
- Coey, C. A., Wallot, S., Richardson, M. J., & Van Orden, G. (2012). On the structure of measurement noise in eye-tracking. *Journal of Eye Movement Research*, 5(4).

- Cohen, M. X., & Aston-Jones, G. (2005). Brain norepinephrine: Coordinating regulation of arousal and cognition. *Frontiers in Behavioral Neuroscience*, 9, 120.
- Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). A unique adolescent response to reward prediction errors. *Nature neuroscience*, 13(6), 669-671.
- Coll, C. G., Crnic, K., Lamberty, G., Wasik, B. H., Jenkins, R., Garcia, H. V., & McAdoo, H. P. (1996). An integrative model for the study of developmental competencies in minority children. *Child development*, 67(5), 1891-1914.
- Crone, E. A., & Van Der Molen, M. W. (2007). Development of decision making in school-aged children and adolescents: Evidence from heart rate and skin conductance analysis. *Child development*, 78(4), 1288-1301.
- Cui, J., Lichtwarck-Aschoff, A., Olthof, M., Li, T., & Hasselman, F. (2022). From metaphor to computation: Constructing the potential landscape for multivariate psychological formal models. *Multivariate Behavioral Research*, 1-19.
- de la Torre-Luque, A., Bornas, X., Balle, M., & Fiol-Veny, A. (2016). Complexity and nonlinear biomarkers in emotional disorders: A meta-analytic study. *Neuroscience & Biobehavioral Reviews*, 68, 410-422.
- Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2015). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological Science*, 26(6), 848-858.
- Del Giudice, M., Ellis, B. J., & Shirliff, E. A. (2011). The adaptive calibration model of stress responsivity. *Neuroscience & Biobehavioral Reviews*, 35(7), 1562-1592.
- DeJoseph, M., Finegood, E., Raver, C. C., & Blair, C. B. (2019). Measuring stress reactivity in the home: Preliminary findings from a version of the Trier Social Stress Test (TSST-H) appropriate for field-based research.
- DeJoseph, M. L., Herzberg, M. P., Sifre, R. D., Berry, D., & Thomas, K. M. (2022). Measurement matters: An individual differences examination of family socioeconomic factors, latent dimensions of children's experiences, and resting state functional brain connectivity in the ABCD sample. *Developmental Cognitive Neuroscience*, 53, 101043.
- DeJoseph, M.L., Leneman, K.L., Palmer, A.P., Padrutt, E., Mayo., O., & Berry, D. (under review). Adrenocortical and autonomic cross-system regulation in youth: A meta-analysis.
- DeJoseph, M. L., Sifre, R. D., Raver, C. C., Blair, C. B., & Berry, D. (2021). Capturing environmental dimensions of adversity and resources in the context of poverty across infancy through early adolescence: A moderated nonlinear factor model. *Child Development*, 92(4), e457-e475.
- DePasque, S., & Galván, A. (2017). Frontostriatal development and probabilistic reinforcement learning during adolescence. *Neurobiology of Learning and Memory*, 143, 1-7.
- Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. (2022b). Reinforcement learning and Bayesian inference provide complementary models for the unique advantage of adolescents in stochastic reversal. *Developmental Cognitive Neuroscience*, 55, 101106.
- Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. (2022). The interpretation of computational model parameters depends on the context. *Elife*, 11, e75474.
- Eckstein, M. K., Wilbrecht, L., & Collins, A. G. (2021). What do reinforcement learning models measure? Interpreting model parameters in cognition and neuroscience. *Current Opinion in Behavioral Sciences*, 41, 128-137.

- Ellwood-Lowe, M. E., Whitfield-Gabrieli, S., & Bunge, S. A. (2021). Brain network coupling associated with cognitive performance varies as a function of a child's environment in the ABCD study. *Nature communications*, *12*(1), 7183.
- Ellis, B. J., Sheridan, M. A., Belsky, J., & McLaughlin, K. A. (2022). Why and how does early adversity influence development? Toward an integrated model of dimensions of environmental experience. *Development and Psychopathology*, *34*(2), 447-471.
- Everaert, J., Koster, E. H., & Joormann, J. (2020). Finding patterns in emotional information: Enhanced sensitivity to statistical regularities within negative information. *Emotion*, *20*(3), 426.
- Favela, L. H. (2019). Emergence by way of dynamic interactions. *Southwest Philosophy Review*, *35*(1), 47-57.
- Fields, A., Bloom, P. A., VanTieghem, M., Harmon, C., Choy, T., Camacho, N. L., ... & Tottenham, N. (2021). Adaptation in the face of adversity: Decrements and enhancements in children's cognitive control behavior following early caregiving instability. *Developmental Science*, *24*(6), e13133.
- Finger, E. C., Mitchell, D. G., Jones, M., & Blair, R. J. R. (2008). Dissociable roles of medial orbitofrontal cortex in human operant extinction learning. *Neuroimage*, *43*(4), 748-755.
- Frank, M. J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Current opinion in neurobiology*, *21*(3), 381-386.
- Frankenhuis, W. E., & Nettle, D. (2020). The strengths of people in poverty. *Current Directions in Psychological Science*, *29*(1), 16-21.
- Frankenhuis, W. E., Panchanathan, K., & Barto, A. G. (2019). Enriching behavioral ecology with reinforcement learning methods. *Behavioural Processes*, *161*, 94-100.
- Frankenhuis, W. E., Young, E. S., & Ellis, B. J. (2020). The hidden talents approach: Theoretical and methodological challenges. *Trends in cognitive sciences*, *24*(7), 569-581.
- Fuhrmann, D., Knoll, L. J., & Blakemore, S. J. (2015). Adolescence as a sensitive period of brain development. *Trends in Cognitive Sciences*, *19*(10), 558-566.
- Fusaroli, R., & Cox, C. (2022, January 13). Workshop on bayesian inference: Priors and workflow. Retrieved September 10, 2022, from <https://4ccoxau.github.io/PriorsWorkshop/>
- Garrison, K. E., & Schmeichel, B. J. (2019). Effects of emotional content on working memory capacity. *Cognition and Emotion*, *33*(2), 370-377.
- Gatzke-Kopp, L. M., Benson, L., Ryan, P. J., & Ram, N. (2020). Cortical and affective regulation of autonomic coordination. *Psychophysiology*, *57*(5), e13544.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper).
- Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190502.
- Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2017). When younger learners can be better (or at least more open-minded) than older ones. *Current Directions in Psychological Science*, *26*(3), 194-200.
- Greenough, W. T., Black, J. E., & Wallace, C. S. (2002). Experience and brain development.
- 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. *Journal of Child Psychology and Psychiatry*, 58(7), 770-778.
- Harms, M. B., Shannon Bowen, K. E., Hanson, J. L., & Pollak, S. D. (2018). Instrumental learning and cognitive flexibility processes are impaired in children exposed to early life stress. *Developmental Science*, 21(4), e12596.
- Hasselmann, F. (2022). Early warning signals in phase space: Geometric resilience loss indicators from multiplex cumulative recurrence networks. *Frontiers in Physiology*, 13.
- Heino, M. T., Knittle, K., Noone, C., Hasselmann, F., & Hankonen, N. (2021). Studying behaviour change mechanisms under complexity. *Behavioral Sciences*, 11(5), 77.
- Hox, J. J., van de Schoot, R., & Matthijsse, S. (2012, July). How few countries will do? Comparative survey analysis from a Bayesian perspective. In *Survey Research Methods* (Vol. 6, No. 2, pp. 87-93).
- Humphreys, K. L., Telzer, E. H., Flannery, J., Goff, B., Gabard-Durnam, L., Gee, D. G., ... & Tottenham, N. (2016). Risky decision making from childhood through adulthood: Contributions of learning and sensitivity to negative feedback. *Emotion*, 16(1), 101.
- Ihlen, E. A. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *Frontiers in physiology*, 3, 141.
- Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., Holmes, A., & Levine, B. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, 345, 12-26.
- Javadi, A. H., Schmidt, D. H., & Smolka, M. N. (2014). Adolescents adapt more slowly than adults to varying reward contingencies. *Journal of cognitive neuroscience*, 26(12), 2670-2681.
- Jenson, J. M., Porter, B., & Wolf, L. (2017). The effects of poverty on children. *Pediatrics in Review*, 38(6), 262-274.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, 90(430), 773-795.
- Kello, C. T., Anderson, G. G., Holden, J. G., & Van Orden, G. C. (2008). The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. *Cognitive Science*, 32(7), 1217-1231.
- Keramati, M., & Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife*, 3, e04811.
- Kraus, M. W., Côté, S., & Keltner, D. (2023). Social class, empathy, and prosocial behavior. *Current Opinion in Psychology*, 46, 23-28.
- Krough, J. P., Sherman, S. J., & Shimamura, A. P. (2013). Cognitive models of decision making: Integrating insights from learning and reasoning. *The Psychology of Learning and Motivation*, 58, 233-267.
- Kurz, S. (2022) Doing Bayesian Data Analysis in brms and the tidyverse. Available at: <https://bookdown.org/content/3686/> (Accessed: October 2, 2022).
- Kurz, S. (2022, October 11). Statistical rethinking with brms, ggplot2, and the tidyverse. Retrieved November 10, 2022, from <https://bookdown.org/content/3686/>
- Kurz, S. (2023, March 20). *Introduction to Bayesian statistics with brms and the tidyverse (03-2023)*. OSF. <https://osf.io/bfvs4/>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (pp. A-8). Gainesville, FL: NIMH, Center for the Study of Emotion & Attention.

- Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate behavioral research*, 39(4), 653-686.
- Lin, W. C., Delevich, K., & Wilbrecht, L. (2020). A role for adaptive developmental plasticity in learning and decision making. *Current Opinion in Behavioral Sciences*, 36, 48-54.
- Martínez, C. A. G., Quintana, A. O., Vila, X. A., Touriño, M. J. L., Rodríguez-Liñares, L., Presedo, J. M. R., & Penín, A. J. M. (2017). Heart rate variability analysis with the R package RHRV.
- Masten, A. S. (2018). Resilience theory and research on children and families: Past, present, and promise. *Journal of Family Theory & Review*, 10(1), 12-31.
- Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., & Collins, A. G. (2020). Disentangling the systems contributing to changes in learning during adolescence. *Developmental cognitive neuroscience*, 41, 100732.
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press.
- McLaughlin, K. A., DeCross, S. N., Jovanovic, T., & Tottenham, N. (2019). Mechanisms linking childhood adversity with psychopathology: Learning as an intervention target. *Behaviour research and therapy*, 118, 101-109.
- Marlow, C. A., Viskontas, I. V., Matlin, A., Boydston, C., Boxer, A., & Taylor, R. P. (2015). Temporal structure of human gaze dynamics is invariant during free viewing. *PloS one*, 10(9), e0139379.
- Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, 39, e200.
- McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 750-773.
- Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S., & Young, E. S. (2015). Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology*, 109(4), 604-621.
- Nalborczyk, L., Batailler, C., Løevenbruck, H., Vilain, A., & Bürkner, P. C. (2019). An introduction to Bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language, and Hearing Research*, 62(5), 1225-1242.
- Nashiro, K., Sakaki, M., Nga, L., & Mather, M. (2012). Differential brain activity during emotional versus nonemotional reversal learning. *Journal of Cognitive Neuroscience*, 24(8), 1794-1805.
- Nelson, C. A., de Haan, M., & Thomas, K. M. (2012). *Neuroscience of cognitive development: The role of experience and the developing brain*. John Wiley & Sons.
- Niv, Y. (2021). The primacy of behavioral research for understanding the brain. *Behavioral Neuroscience*, 135(5), 601.
- Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research?. *Developmental cognitive neuroscience*, 40, 100733.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature neuroscience*, 4(1), 95-102.
- Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for

- autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2), 265-286.
- Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S. J. (2016). The computational development of reinforcement learning during adolescence. *PLoS computational biology*, 12(6), e1004953.
- Palminteri, S., & Lebreton, M. (2022). The neural computation of valence-based learning rates: An empirical and theoretical investigation. *NeuroImage*, 245, 118706.
- Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, 52(1), 111-139.
- Peng, C. K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: an interdisciplinary journal of nonlinear science*, 5(1), 82-87.
- Peverill, M., Dirks, M. A., Narvaja, T., Herts, K. L., Comer, J. S., & McLaughlin, K. A. (2021). Socioeconomic status and child psychopathology in the United States: A meta-analysis of population-based studies. *Clinical psychology review*, 83, 101933.
- Plate, R. C., Schapiro, A. C., & Waller, R. (2022). Emotional Faces Facilitate Statistical Learning. *Affective Science*, 3(3), 662-672.
- Pollak, S. D. (2005). Early adversity and mechanisms of plasticity: Integrating affective neuroscience with developmental approaches to psychopathology. *Development and Psychopathology*, 17(3), 735-752.
- Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion recognition. *Cognition*, 110(2), 242-247.
- Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional bias for positive emotional stimuli: A meta-analytic investigation. *Psychological bulletin*, 142(1), 79.
- Reeb-Sutherland, B. C., Levitt, P., & Fox, N. A. (2012). The predictive nature of individual differences in early associative learning abilities: A latent variable growth curve model. *Developmental Science*, 15(4), 437-447.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, 64-99.
- Rosen, M. L., Sheridan, M. A., Sambrook, K. A., Dennison, M. J., Jenness, J. L., Askren, M. K., ... & McLaughlin, K. A. (2018). Saliency network response to changes in emotional expressions of others is heightened during early adolescence: relevance for social functioning. *Developmental Science*, 21(3), e12571.
- Rosenblum, G. D., & Lewis, M. (2003). Emotional development in adolescence.
- Roy, A. L., Raver, C. C., Masucci, M. D., & DeJoseph, M. (2019). "If they focus on giving us a chance in life we can actually do something in this world": Poverty, inequality, and youths' critical consciousness. *Developmental Psychology*, 55(3), 550.
- Schwabe, L., Bohringer, A., Chatterjee, M., & Schachinger, H. (2008). Effects of pre-learning stress on memory for neutral, positive and negative words: Different roles of cortisol and autonomic arousal. *Neurobiology of learning and memory*, 90(1), 44-53.
- Schwabe, L., Oitzl, M. S., Philippson, C., Richter, S., Bohringer, A., Wippich, W., & Schachinger, H. (2007). Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Learning & memory*, 14(1-2), 109-116.
- Shiu, L.-P., & Chan, C. C. H. (2006). Associative learning of visual and spatial

- information in a spatial environment with a probabilistic landmark. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 116-134.
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in cognitive sciences*, 7(8), 343-348.
- Somerville, L. H., Sasse, S. F., Garrad, M. C., Drysdale, A. T., Abi Akar, N., Insel, C., & Wilson, R. C. (2017). Charting the expansion of strategic exploratory behavior during adolescence. *Journal of experimental psychology: general*, 146(2), 155.
- Stallworthy, I. C., Sifre, R., Berry, D., Lasch, C., Smith, T. J., & Elison, J. T. (2020). Infants' gaze exhibits a fractal structure that varies by age and stimulus salience. *Scientific reports*, 10(1), 17216.
- Stern, H. (Host). (2012). Snoop Dogg (Calvin Cordozar Broadus Jr.) Interview [Interview quote]. The Howard Stern Show.
<https://www.artsmanagementmagazine.com/article/amm-inspiration-snoop-dogg/>
- Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. MIT Press.
- Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science*, 22(6), 430-436.
- Telzer, E. H. (2016). Dopaminergic reward sensitivity can promote adolescent health: A new perspective on the mechanism of ventral striatum activation. *Developmental cognitive neuroscience*, 17, 57-67.
- Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of behavioral medicine*, 37(2), 141-153.
- Thomas, K. M., & de Haan, M. (2012). Development of human brain functions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(3), 261-277.
- Todd, R. M., & Manaligod, M. G. (2018). Implicit guidance of attention: The priority state space framework. *Cortex*, 102, 121-138.
- Todd, R. M., Talmi, D., Schmitz, T. W., Susskind, J., & Anderson, A. K. (2012). Psychophysical and neural evidence for emotion-enhanced perceptual vividness. *Journal of Neuroscience*, 32(33), 11201-11212.
- Tooley, U. A., Bassett, D. S., & Mackey, A. P. (2021). Environmental influences on the pace of brain development. *Nature Reviews Neuroscience*, 22(6), 372-384.
- Ugarte, E., & Hastings, P. (2023). Assessing Unpredictability in Caregiver-Child Relationships: Insights from Theoretical and Empirical Perspectives. *Development and Psychopathology*.
- Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature reviews neuroscience*, 10(6), 397-409.
- Van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral cortex*, 22(6), 1247-1255.
- Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & Van Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child development*, 85(3), 842-860.
- Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of experimental psychology: General*, 132(3), 331.
- Van Orden, G. C., Kloos, H., & Wallot, S. (2011). Living in the pink: Intentionality, wellbeing, and complexity. In *Philosophy of complex systems* (pp. 629-672). North-Holland.
- Veenman, M., Stefan, A., & Haaf, J. M. (2022). Bayesian Hierarchical Modeling: An Introduction and Reassessment.

- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, 27, 1413-1432.
- Vogel, S., & Schwabe, L. (2016). Learning and memory under stress: Implications for the classroom. *npj Science of Learning*, 1, 16011.
- Vuilleumier, P., & Huang, Y. M. (2009). Emotional attention: Uncovering the mechanisms of affective biases in perception. *Current Directions in Psychological Science*, 18(3), 148-152.
- Walle, E. A., Reschke, P. J., & Knothe, J. M. (2017). Social referencing: Defining and delineating a basic process of emotion. *Emotion Review*, 9(3), 245-252.
- Weiss, E. O., Kruppa, J. A., Fink, G. R., Herpertz-Dahlmann, B., Konrad, K., & Schulte-Rüther, M. (2021). Developmental differences in probabilistic reversal learning: a computational modeling approach. *Frontiers in neuroscience*, 14, 536596.
- Wijnants, M. L. (2014). A review of theoretical perspectives in cognitive science on the presence of scaling in coordinated physiological and cognitive processes. *Journal of Nonlinear Dynamics*, 2014.
- Wirz, L., Bogdanov, M., & Schwabe, L. (2018). Habits under stress: mechanistic insights across different types of learning. *Current opinion in behavioral sciences*, 20, 9-16.
- Wu, Y., Schulz, L. E., Frank, M. C., & Gweon, H. (2021). Emotion as information in early social learning. *Current Directions in Psychological Science*, 30(6), 468-475.
- Yaple, Z. A., & Yu, A. J. (2019). Reinforcement learning in the human brain: A review of recent progress. *Current Opinion in Behavioral Sciences*, 29, 167-174.
- Yaple, Z. A., & Yu, R. (2019). Fractionating adaptive learning: A meta-analysis of the reversal learning paradigm. *Neuroscience & Biobehavioral Reviews*, 102, 85-94.
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100-1122.
- Yoo, A. H., & Collins, A. G. (2022). How working memory and reinforcement learning are intertwined: A cognitive, neural, and computational perspective. *Journal of cognitive neuroscience*, 34(4), 551-568.
- Young, E. S., Frankenhuis, W. E., DelPriore, D. J., & Ellis, B. J. (2022). Hidden talents in context: Cognitive performance with abstract versus ecological stimuli among adversity-exposed youth. *Child development*, 93(5), 1493-1510.
- Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience*, 15(6), 695-707.

Supplemental Material

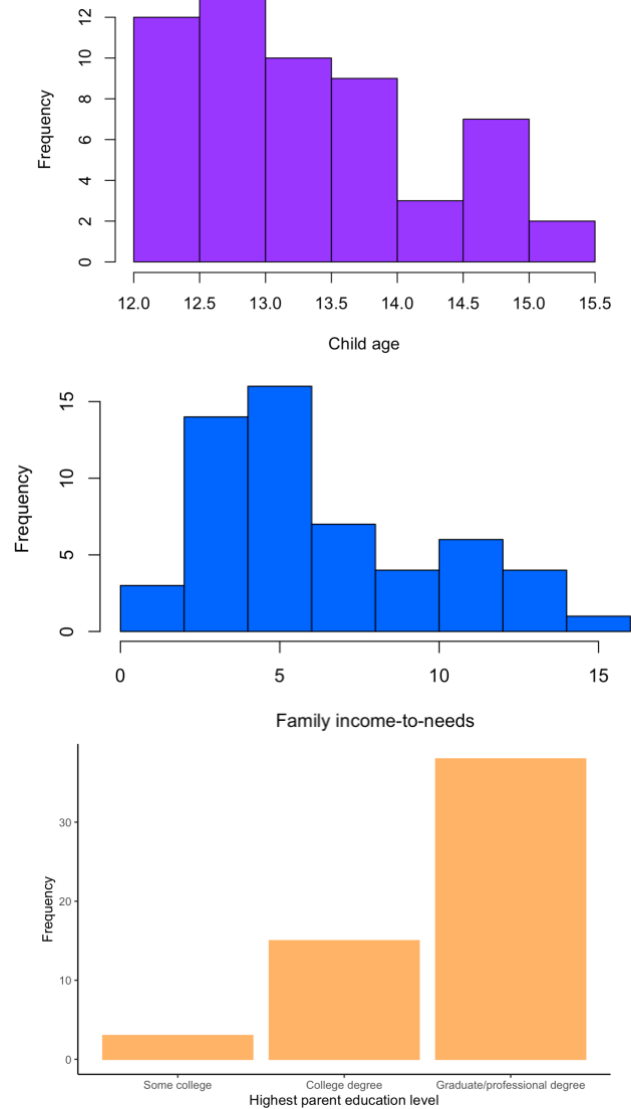


Figure S1. Histograms of additional sociodemographic variables from the current sample.

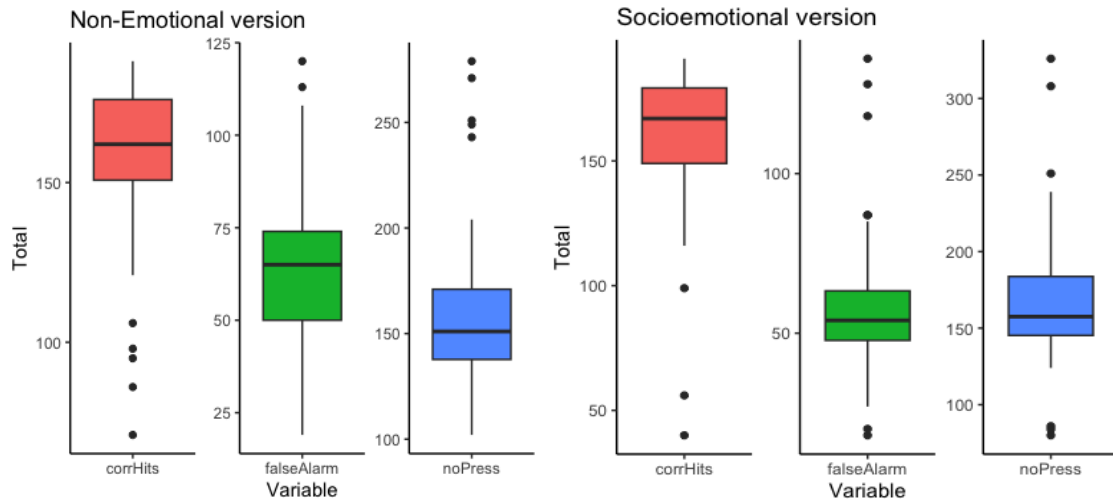


Figure S2. Boxplots of raw behavior variables: total correct hits (i.e., clicks on rewarded stimuli), false alarms (i.e., clicks on punished stimuli), and no press (i.e., did not click and let stimulus pass without feedback). These variables were summed to create the ‘total points’ outcome variable used for Aim 1a.

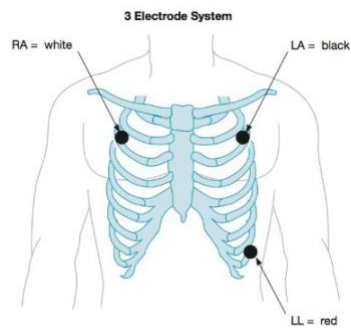


Figure S3. ECG electrode placement.

Bayesian Data Analysis

The Bayesian data analytic framework offers an intuitive approach for statistical modeling by incorporating prior knowledge that can then be updated with empirical data to obtain posterior beliefs about model parameters (Jeffreys, 1961; Veenman et al., 2022). The Bayesian approach differs from frequentist-based analysis in several ways (see McElreath et al., 2020; Nalborczyk et al., 2018; van de Schoot et al., 2014 for a detailed overview).

In Bayesian analysis, each parameter of the model is considered a random variable; in frequentist, they are unknown and fixed quantities. This method of treating parameters as random variables is particularly important in the context of the substantive models used in the current study, whereby the number of random variables in the model is greater than the number of observations—which would have been an unidentified model in a frequentist framework. The increased flexibility of Bayesian modeling addresses these model identification and overfitting issues that can occur in frequentist approaches. In particular, the use of prior distributions in Bayesian modeling helps to regularize the parameter estimates, preventing overfitting by shrinking estimates towards prior beliefs. This improves the stability and robustness of the model and can allow for estimation of complex models with a large number of parameters and a relatively small sample size. The prior distributions can help to constrain the estimates of the random effects, preventing overfitting and allowing the model to converge. However, in the context of the models in the current study, results are slightly more sensitive to the priors. The rigorous prior and posterior sensitivity checks shown in the figures below increase confidence in their utility.

Bayesians also explicitly use probability to model uncertainty, whereas frequentists use probability as the limit of a relative frequency. In Bayesian terms, probability refers to the experience of uncertainty; in frequentist, it is the relative frequency of an event when the number of trials approaches infinity. Bayesian data analysis centers on the probability of a parameter, θ , given a set of data y :

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

Where $p(\theta|y)$ is a probability distribution, derived via a *posterior distribution* that reflects knowledge about the parameter that can be used for statistical inference. The term $p(\theta)$ corresponds to the prior distribution, or the prior information about the parameters. The term $p(y|\theta)$ represents the function through which the data affect the posterior distribution, denoting the likelihood of the data to appear for each possible value of θ . The term $p(y)$ is the marginal likelihood, which scales the probability of the data summed over all values of θ , captured by $p(y) = \sum_{\theta} p(\theta) p(y|\theta)$ for discrete parameters and by $p(y) = \int p(\theta) p(y|\theta) d\theta$ in the case of continuous parameters.

Taken together, any posterior distribution $p(\theta|y)$ from a Bayesian data analysis is given by the product of our priors (i.e., what we already knew) and the likelihood (i.e., what the

data tell us), to optimally maximize and combine our prior knowledge and empirical evidence.

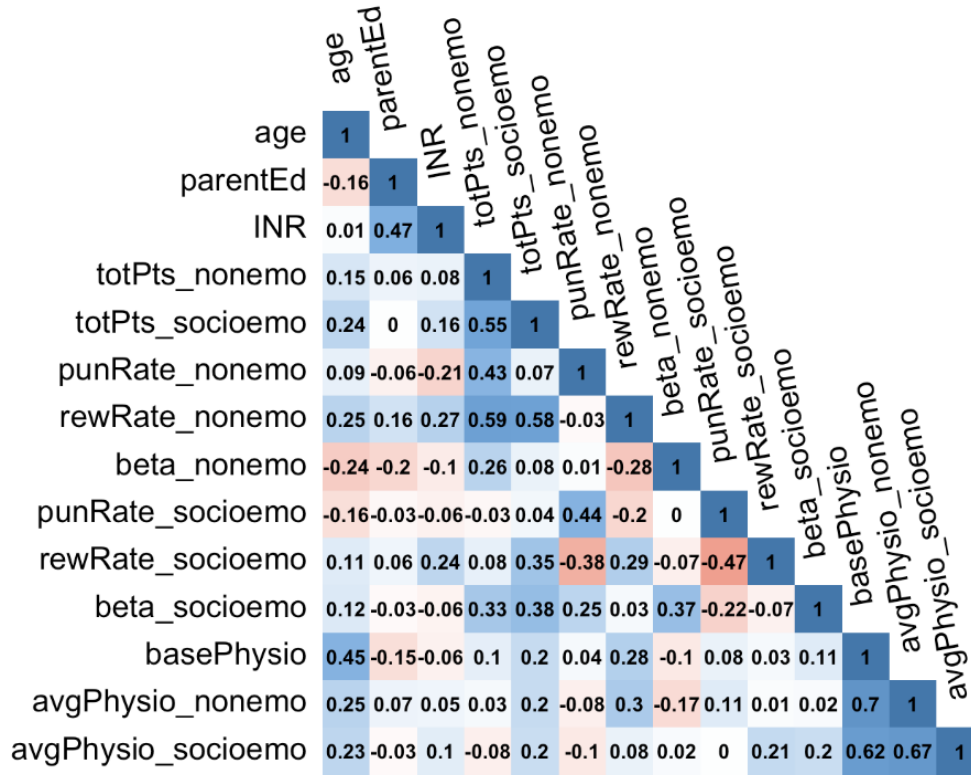


Figure S4. Bivariate correlations between study variables. Color denotes the direction of effect (blue = positive; red = negative) and opacity denotes the magnitude of correlation.

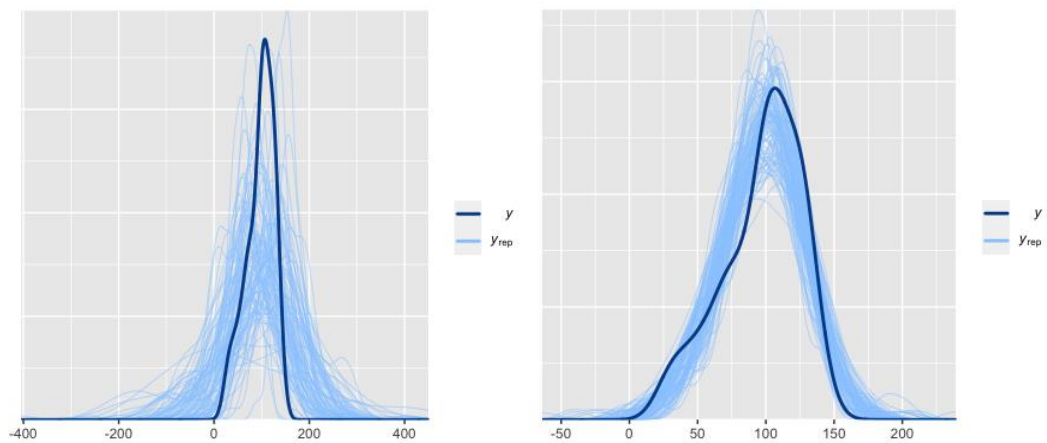


Figure S5. Prior predictive check (left) and posterior predictive check (right) for the total points outcome.

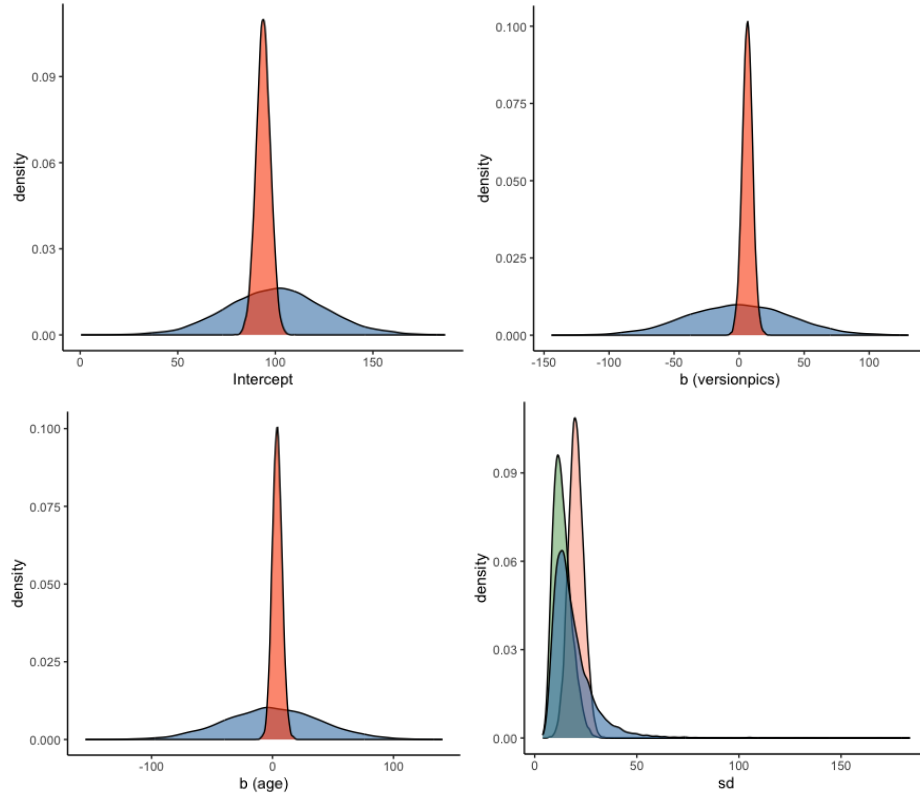


Figure S6. Prior-posterior update plots for key parameters from the total points outcome model. Red = posterior, blue = prior.

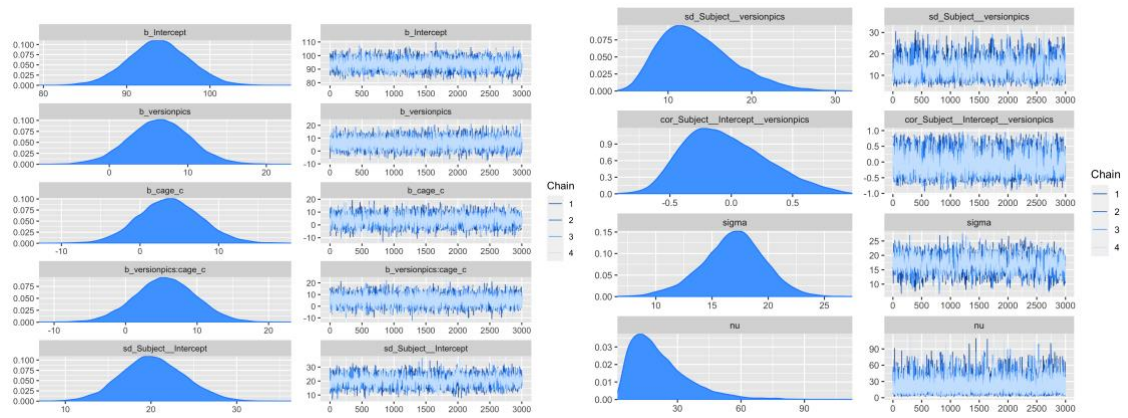


Figure S7. Trace plots depicting convergence across chains for the total points outcome model.

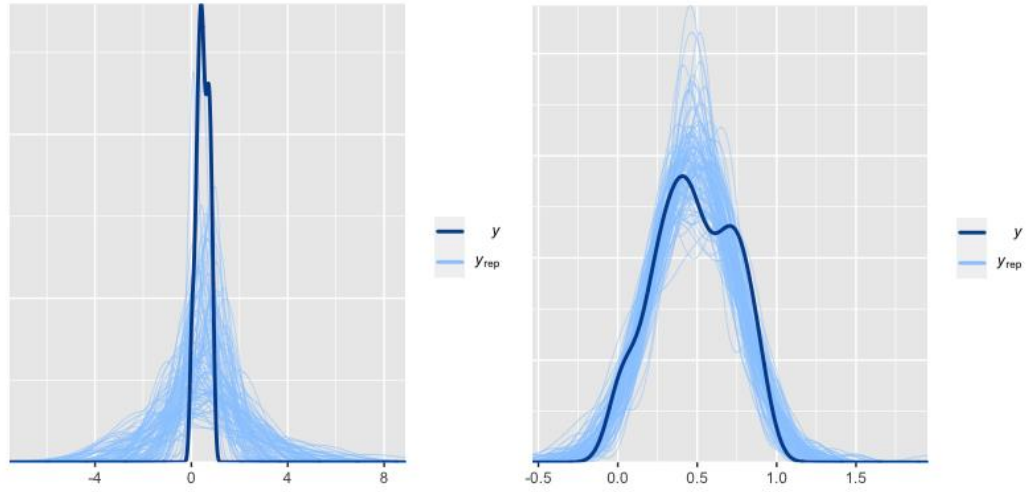


Figure S8. Prior predictive check (left) and posterior predictive check (right) for the punishment learning rate outcome.

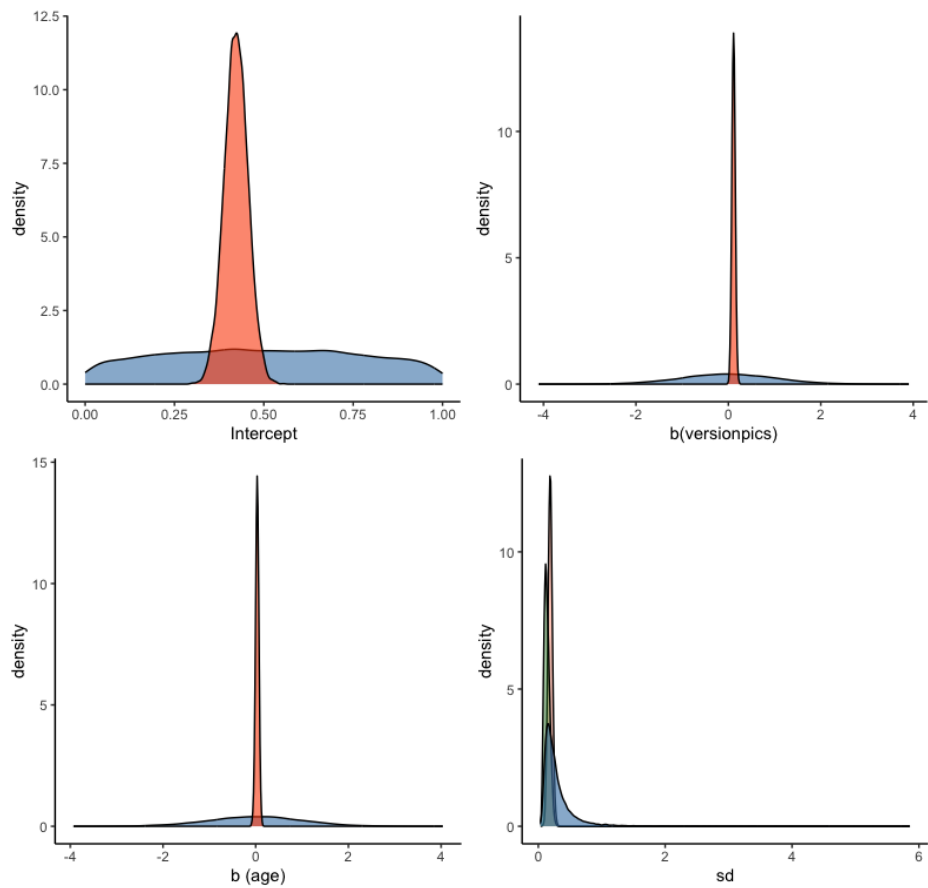


Figure S9. Prior-posterior update plots for key parameters from the punishment learning rate outcome model. Red = posterior, blue = prior.

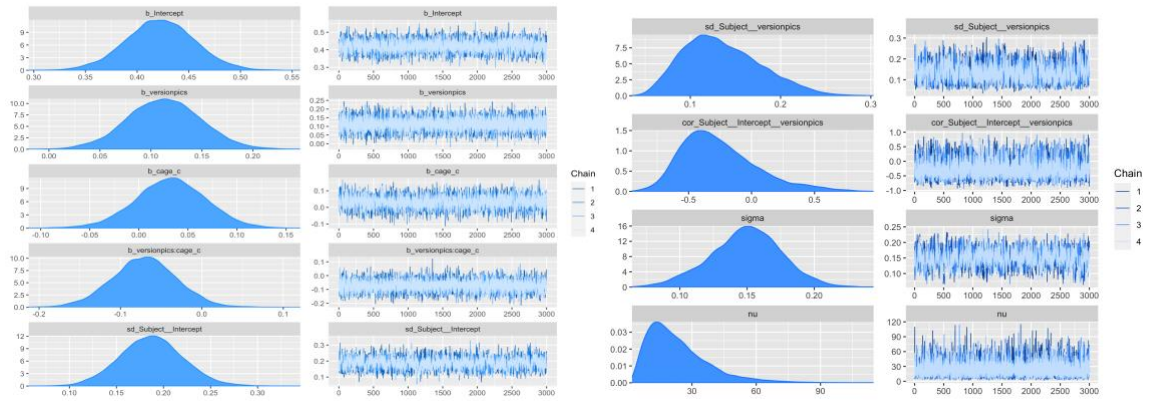


Figure S10. Trace plots depicting convergence across chains for the punishment learning rate outcome model.

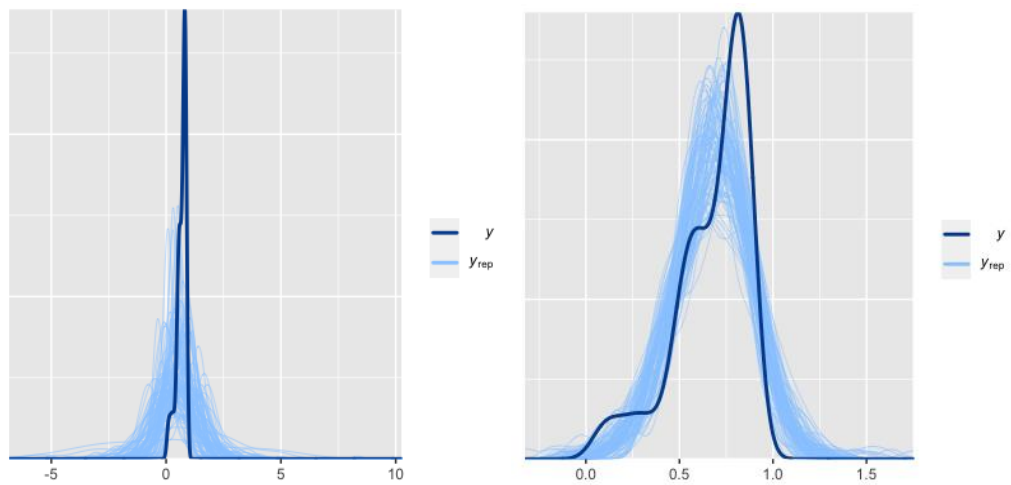


Figure S11. Prior predictive check (left) and posterior predictive check (right) for the reward learning rate outcome.

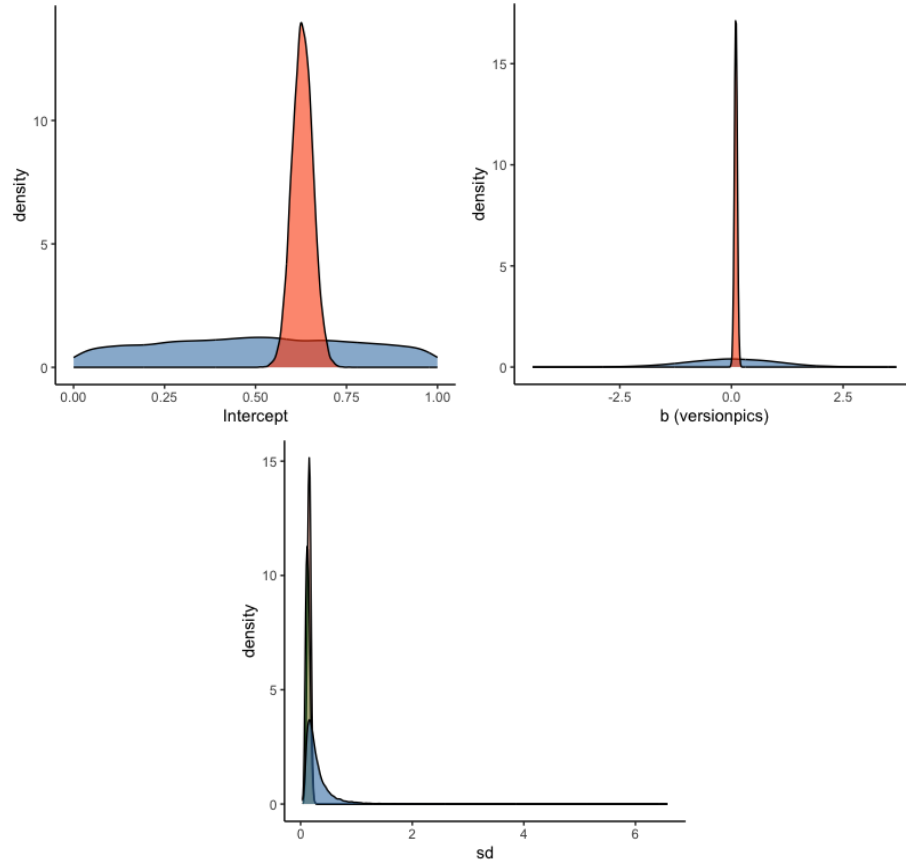


Figure S12. Prior-posterior update plots for key parameters from the reward learning rate outcome model. Red = posterior, blue = prior.

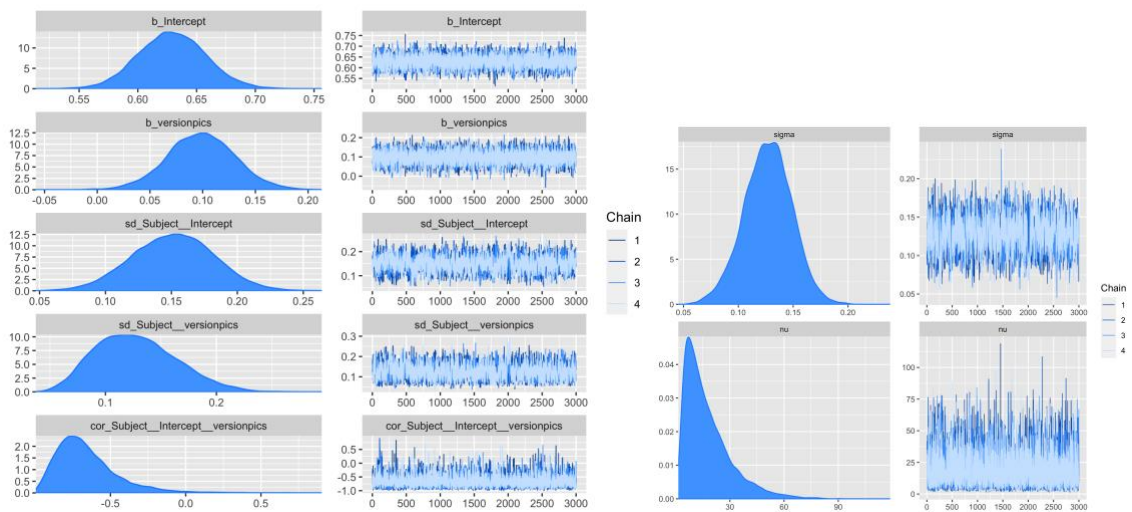


Figure S13. Trace plots depicting convergence across chains for the reward learning rate outcome model.

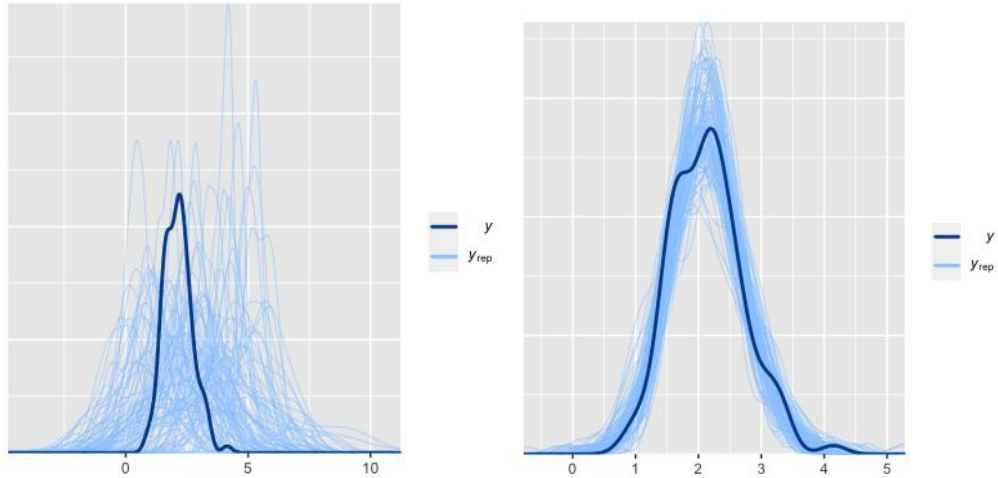


Figure S14. Prior predictive check (left) and posterior predictive check (right) for the inverse temperature outcome.

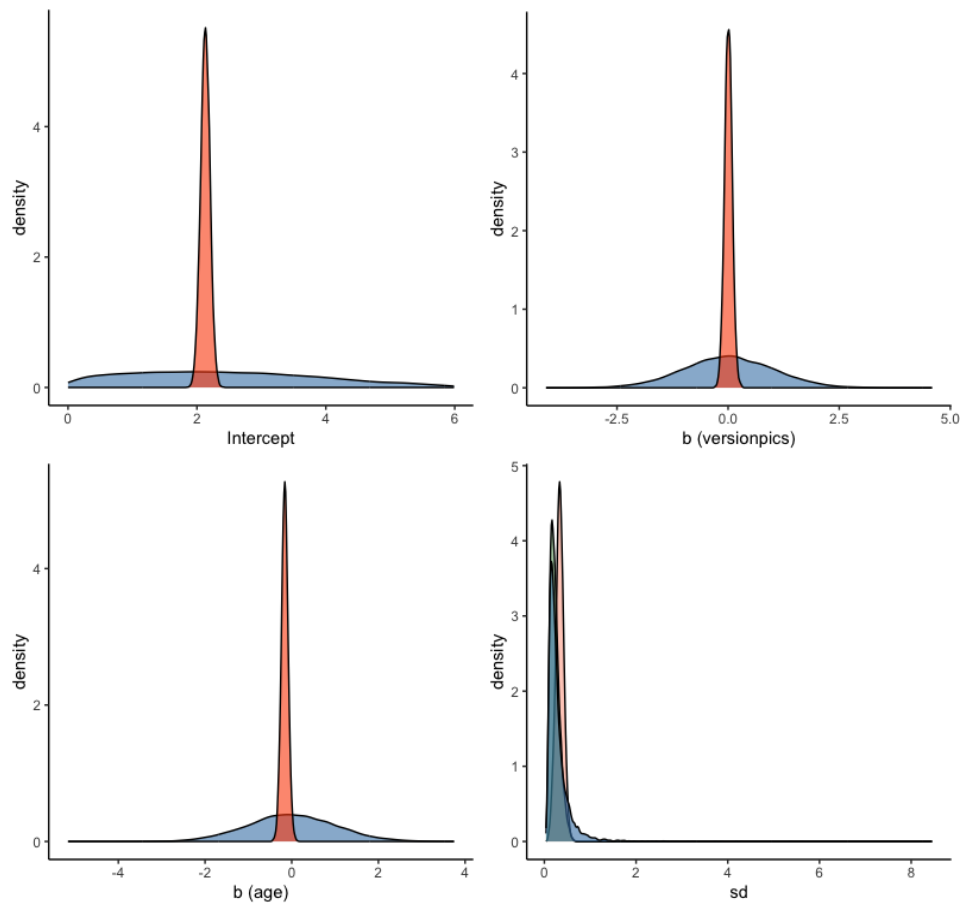


Figure S15. Prior-posterior update plots for key parameters from the inverse temperature outcome model. Red = posterior, blue = prior.

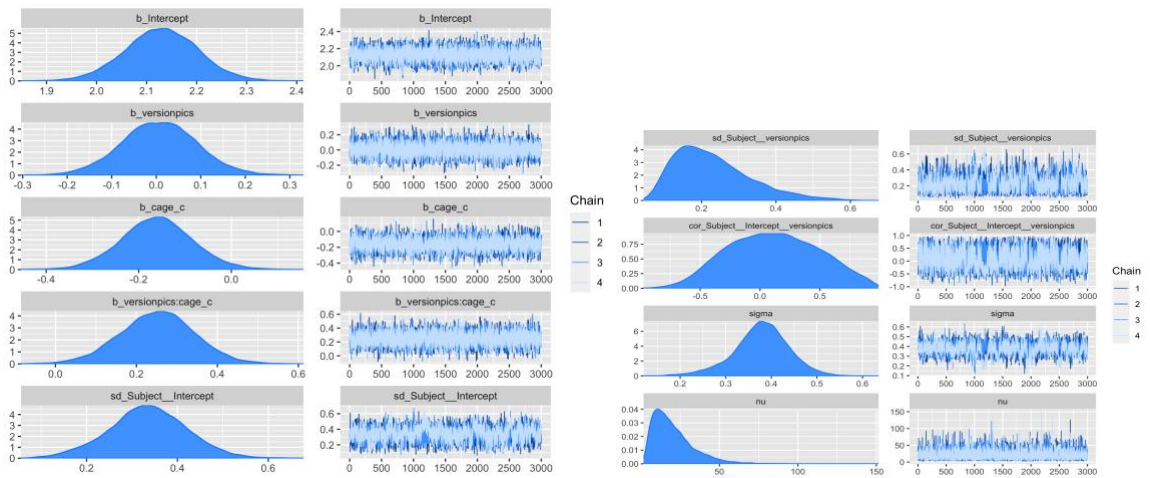


Figure S16. Trace plots depicting convergence across chains for the inverse temperature outcome model.

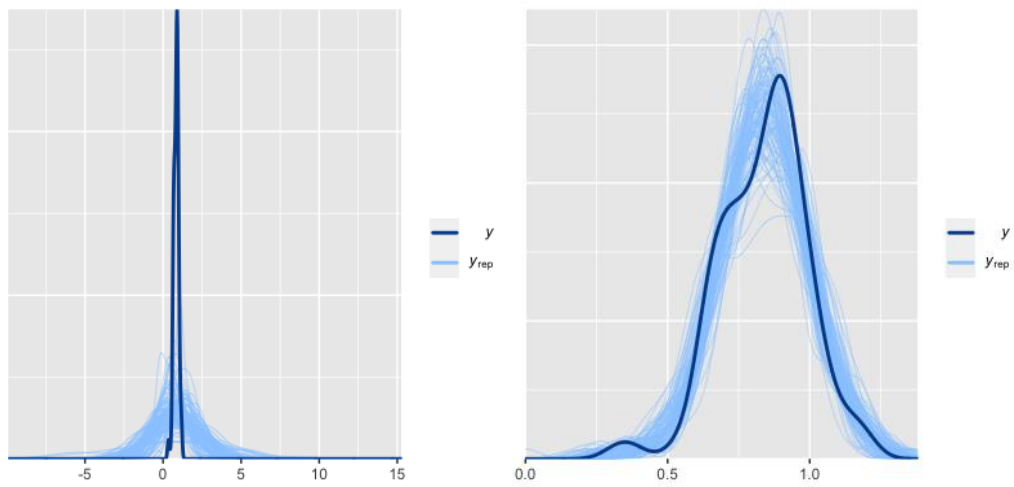


Figure S17. Prior predictive check (left) and posterior predictive check (right) for the alpha (psychophysiology) outcome.

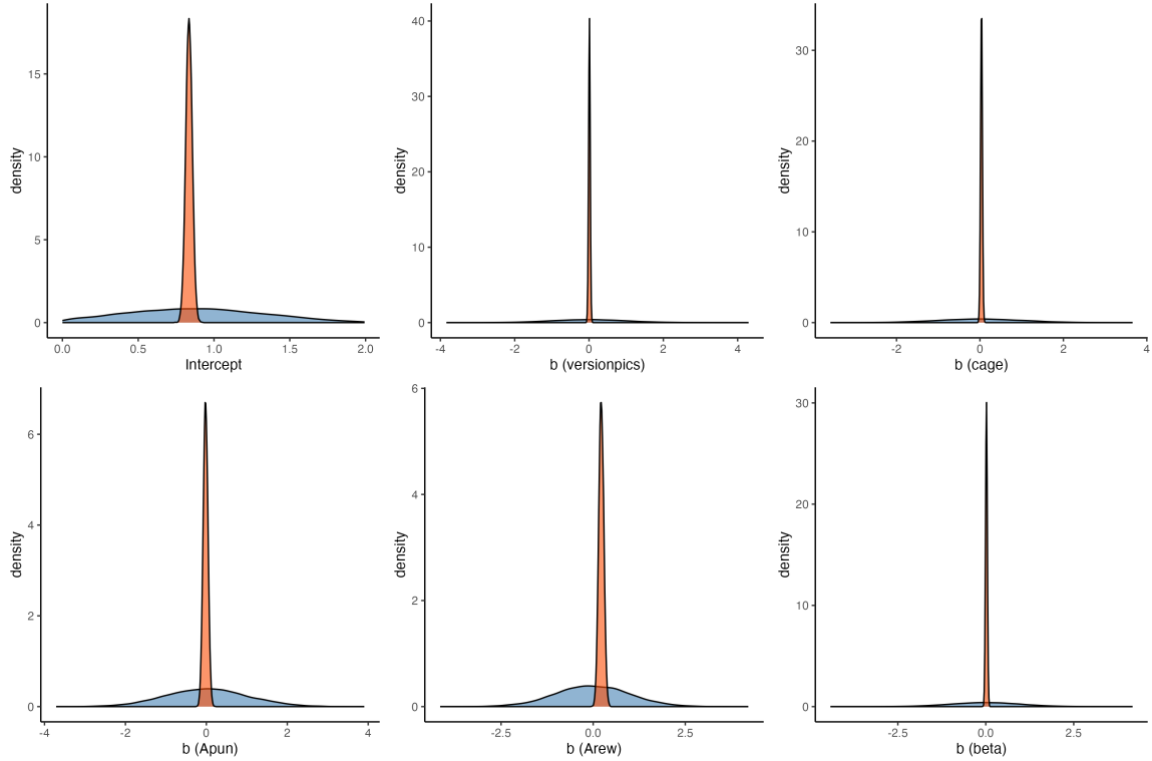


Figure S18. Prior-posterior update plots for key parameters from the alpha (psychophysiology) outcome model. Red = posterior, blue = prior.

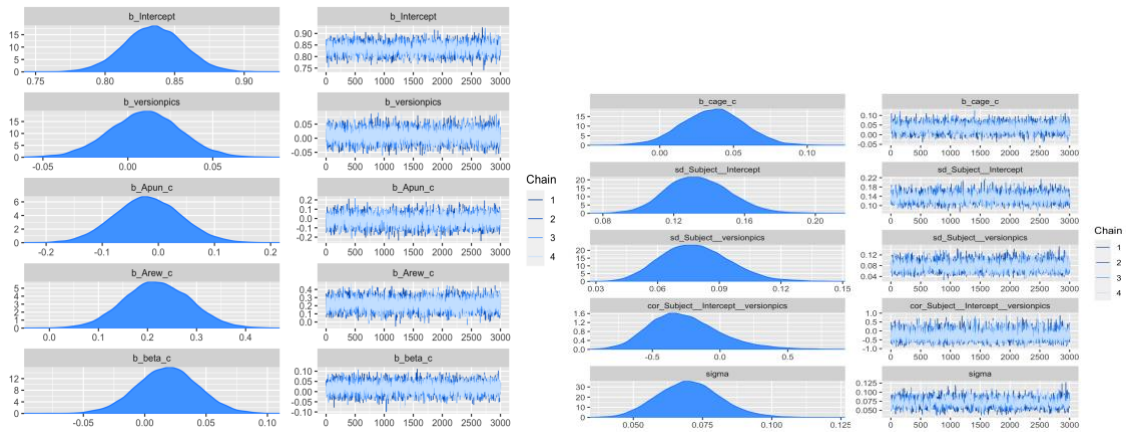


Figure S19. Trace plots depicting convergence across chains for the alpha (psychophysiology) outcome model.

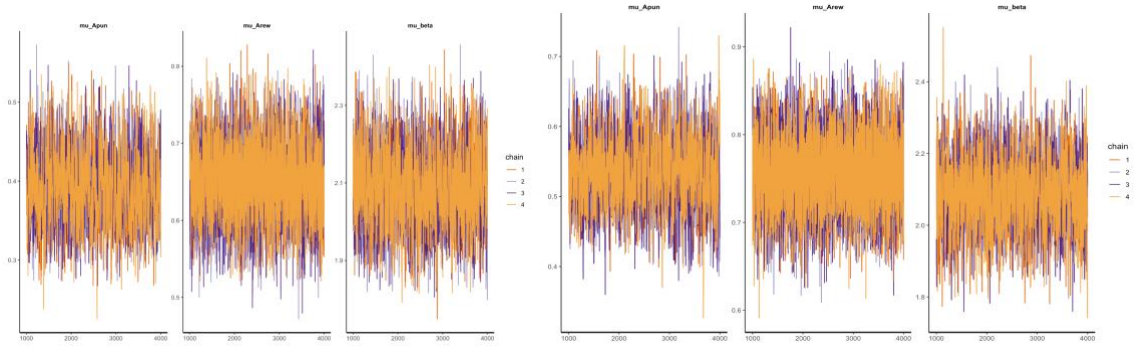


Figure S20. Trace plots depicting convergence across chains for non-emotional (left) and socioemotional (right) computational models.

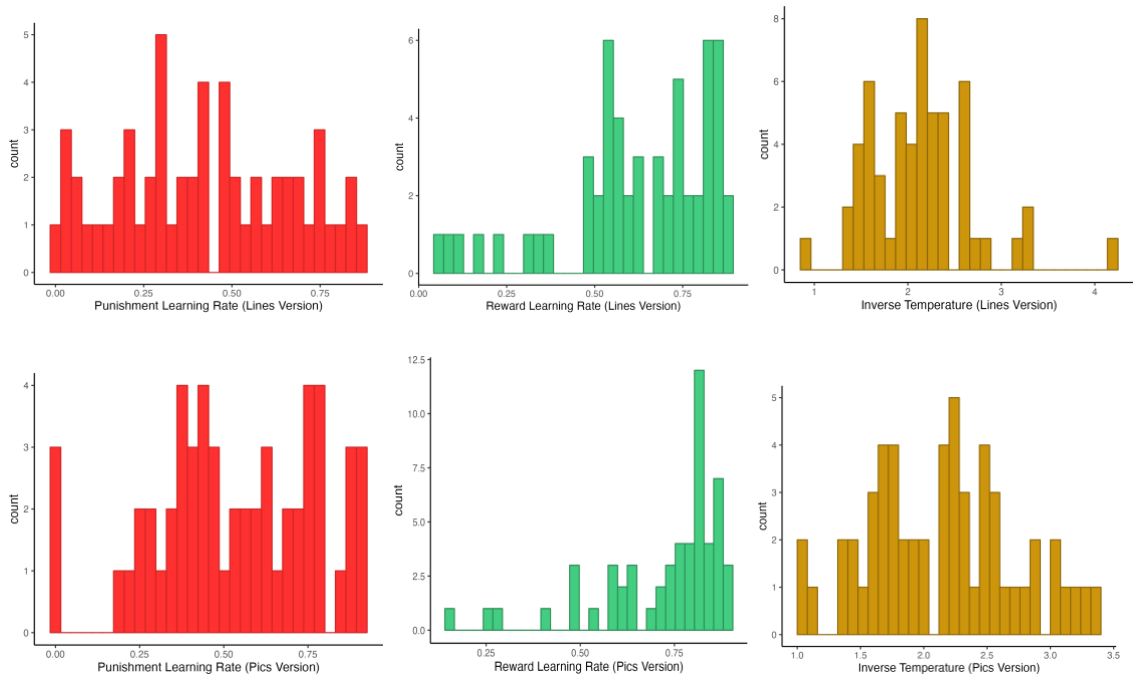


Figure S21. Histograms of computational model-derived parameters for non-emotional (lines) version (top row) and socioemotional (pics) version (bottom row).

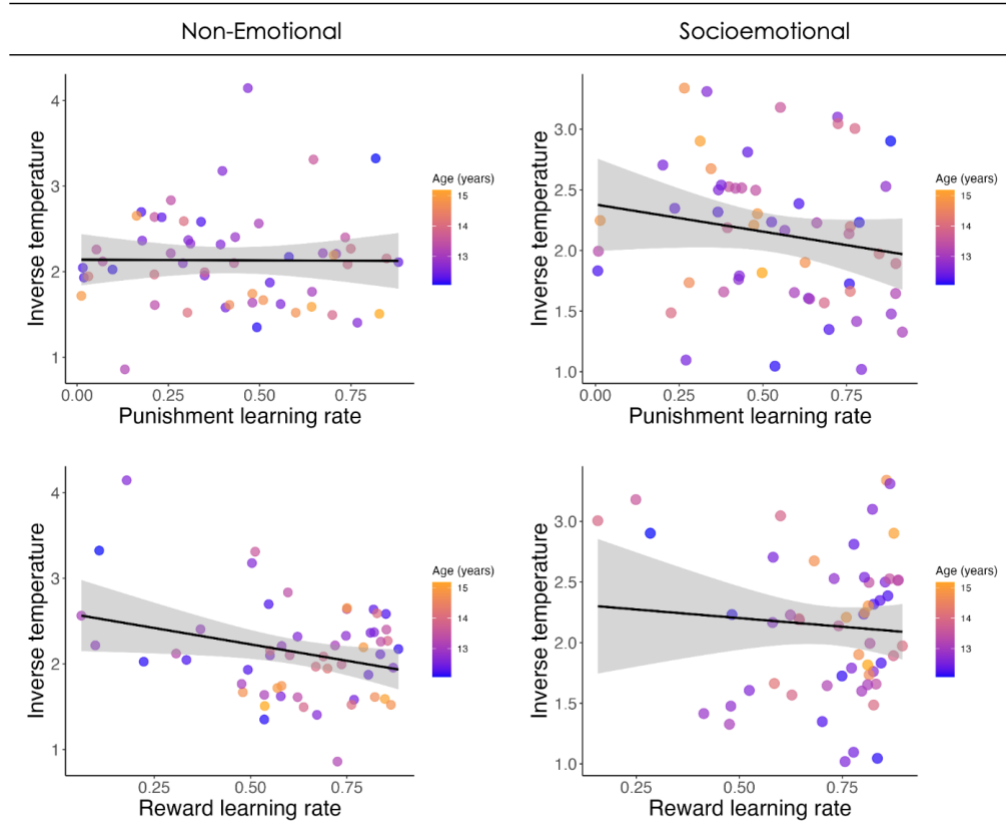


Figure S22. Within-version (left: non-emotional; right: socioemotional) correlations across learning rates and inverse temperature parameters colored by age.

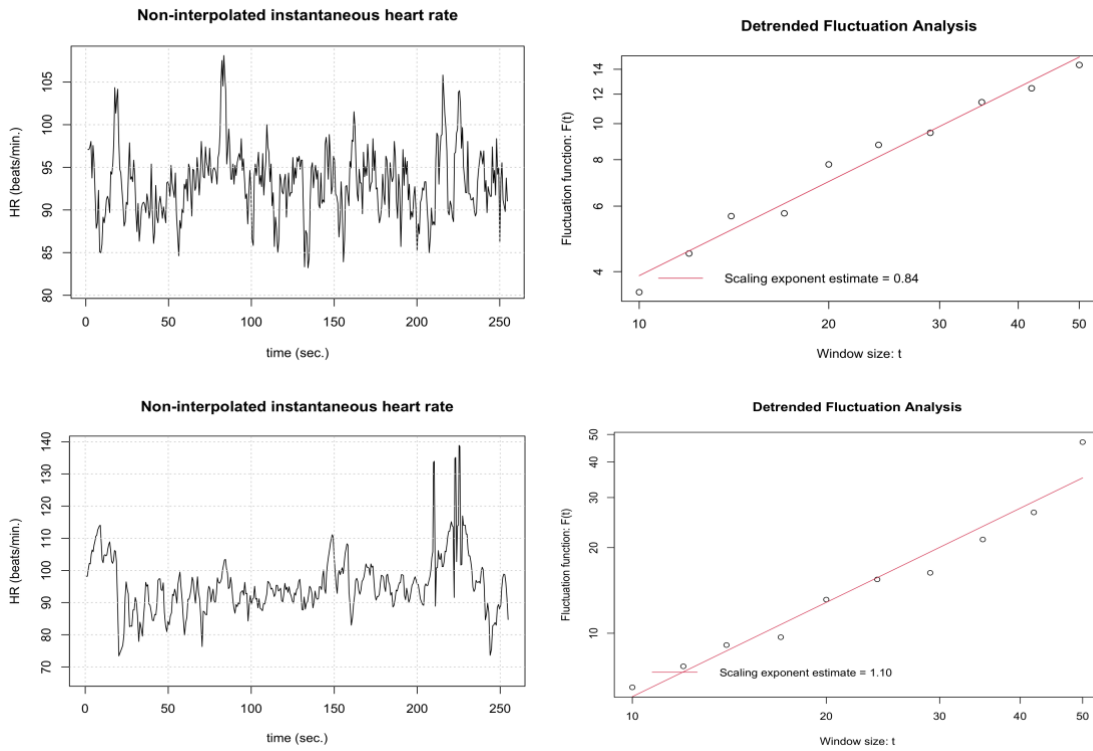


Figure S23. Example IBI segments and DFA regression lines from two randomly selected participants for randomly selected portions of the task.

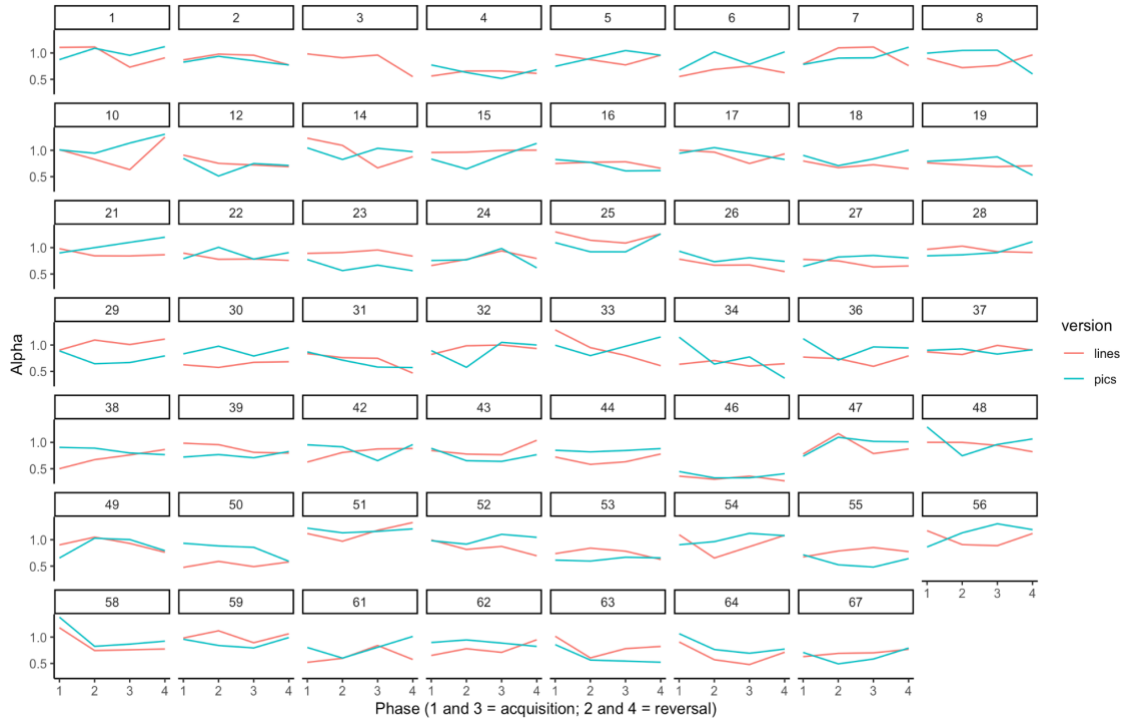


Figure S24. Raw DFA-derived alpha values across the acquisition and reversal phases of the learning task, colored by version (red = non-emotional; blue = socioemotional) and separated by participant.

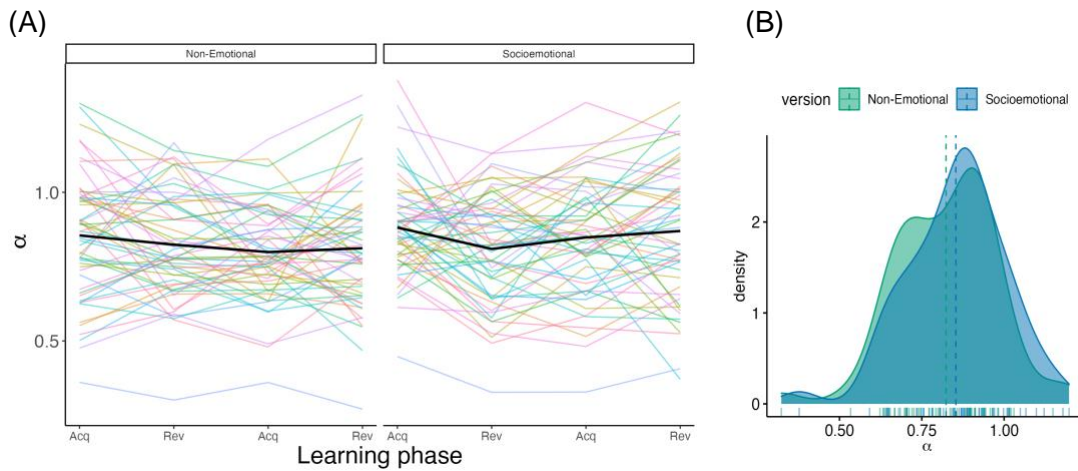


Figure S25. (A) Raw DFA-derived alpha values across the acquisition and reversal phases of the learning task, separated by version (left: non-emotional; right: socioemotional) and colored by participant. Acq = acquisition phase; Rev = reversal phase. (B) Density plot of task average alpha scores collapsed across the task and colored by version.

Table S1. *Final Bayesian mixed effect model predicting psychophysiological engagement (alpha) without the inclusion of the outlier.*

	Posterior Mean	SE	95% Credible Interval	Rhat
<i>Fixed effects</i>				
(Intercept)	0.84	0.02	[0.80, 0.88]	1.00
Version (socioemotional)	0.01	0.02	[-0.03, 0.05]	1.00
Punishment learning rate	-0.03	0.06	[-0.14, 0.08]	1.00
Reward learning rate	0.21	0.07	[0.08, 0.34]	1.00
Inverse temperature	0.01	0.02	[-0.03, 0.06]	1.00
Age	0.03	0.02	[-0.01, 0.07]	1.00
<i>Random effects</i>				
(Intercept)	0.12	0.02	[0.08, 0.15]	1.00
Version (socioemotional)	0.08	0.02	[0.02, 0.05]	1.00
Cor (Intercept, Version)	-0.28	0.27	[-0.73, 0.35]	1.00

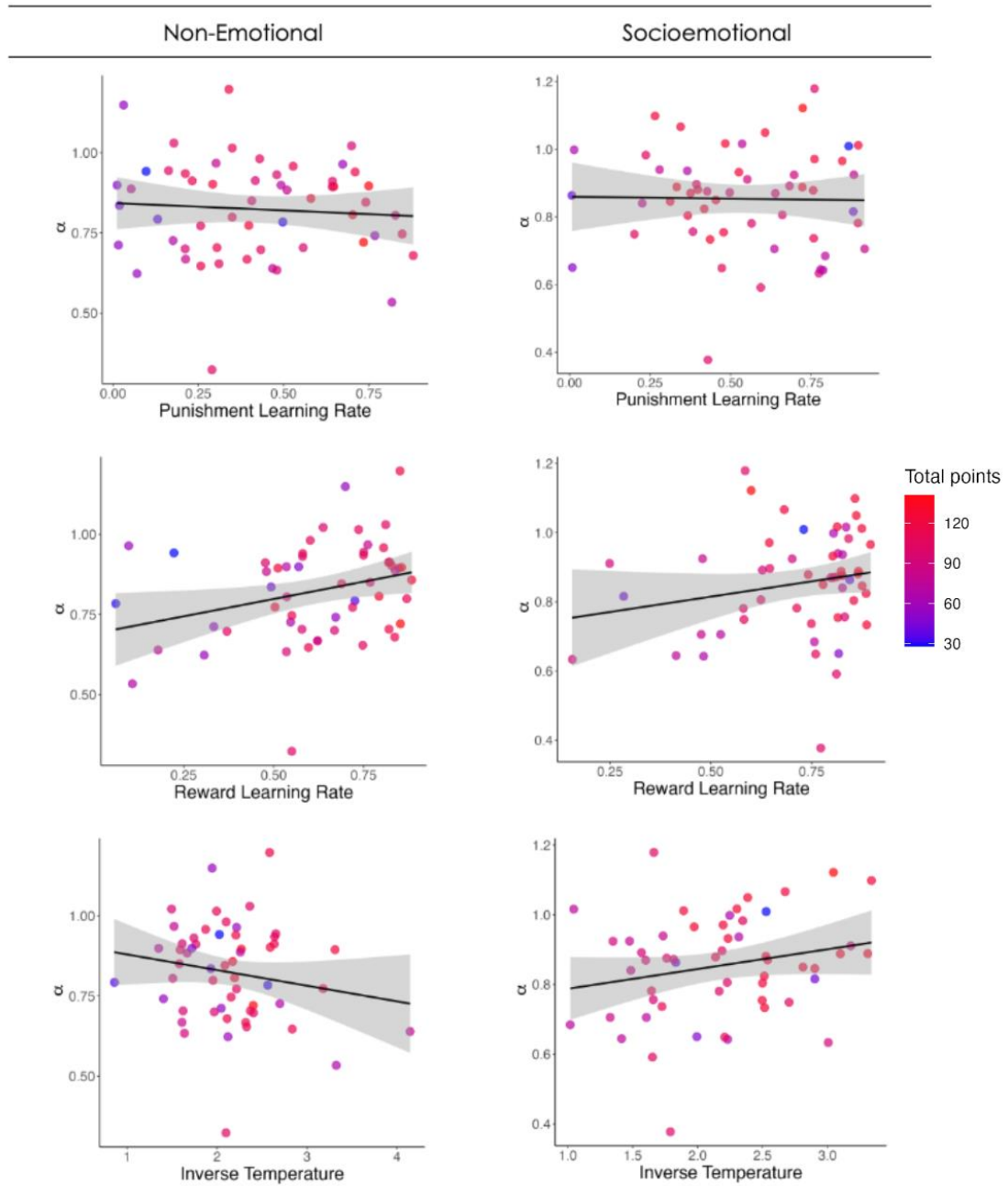


Figure S26. Within-version (left: non-emotional; right: socioemotional) correlations across computational parameters and psychophysiology (alpha) colored by total points.