UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERISTY

Reinforcement learning in eating disorders: Applications of computational modeling to behavioral and neural data

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Clinical Psychology

by

Carina S. Brown

Committee in charge:

University of California San Diego

Professor Christina Wierenga, Chair
Professor Amanda Bischoff-Grethe
Professor Alan Simmons

San Diego State University

Professor Emily Kappenman
Professor Scott Roesch

2026

The Dissertation of Carina S. Brown is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

Chair

University of California San Diego
San Diego State University

2026

DEDICATION

To my closest friends and family – I love and treasure each of you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge and thank my advisor and mentor, Dr. Christina Wierenga. You have been an invaluable guiding hand in my graduate school experience, and I truly could not have accomplished what I have without your patience, brilliance, and unwavering encouragement. You have given me the freedom and flexibility to grow and challenge myself, while providing structure and support when needed. You have played such an enormous role in shaping who I am as a researcher and clinician, and I am endlessly grateful that I was lucky enough to be your student.

I would also like to thank Dr. Amanda Bischoff-Grethe for your constant support through my graduate school training. You have been such a wonderful mentor for me and such an incredible source of knowledge and skill. I cannot thank you enough for all the ways you have contributed to my development as a researcher.

Special thanks to Audrey Nunez, Katherine Gandee, Megan Martinho, Noriko Coburn, and Angeline Krueger – without you, this work could not have been done. And special thanks to my dissertation committee, Drs. Alan Simmons, Scott Roesch, and Emily Kappenman for your expertise, time, and incredibly valuable guidance.

Finally, I would like to thank my fiancé, Kaustubh Kulkarni, for your constant love, support, insight, and belief in me. I will never be able to put into words how grateful I am to have you in my life. You have spent countless hours listening to my ideas, helping me troubleshoot seemingly endless roadblocks, and championing me when I doubted myself. You are my inspiration and the source of my greatest joy.

Chapter 1 contains unpublished material. I would like to acknowledge and thank: Sean Devine, PhD, Ross Otto, PhD, Amanda Bischoff-Grethe, PhD, and Christina Wierenga, PhD. I

would also like to thank Megan Martinho, Angeline Krueger, the UCSD Eating Disorders Center for Treatment and Research staff and patients for their contributions to this project. This project was supported by National Institutes of Health Grant No. R21MH118409 (awarded to CEW).

Chapter 2 contains unpublished material. I would like to acknowledge and thank: Audrey Nuñez, Amanda Bischoff-Grethe, PhD, and Christina Wierenga, PhD. I would also like to thank Nader Amir, PhD, Noriko Coburn, and the UCSD Eating Disorders Center for Treatment and Research staff and patients for their contributions to this project.

Chapter 3 contains unpublished material. I would like to acknowledge and thank: Katherine Gandee; Megan Martinho, Amanda Bischoff-Grethe, PhD, and Christina Wierenga, PhD. I would also like to thank Angeline Krueger, the UCSD Eating Disorders Center for Treatment and Research staff and patients, Xiaosi Gu, PhD, and Laura Berner, PhD for their contributions to this project. This project was supported by National Institutes of Health Grant No. F31MH133362 (awarded to CSB) and Grant No. R01MH125880 (awarded to CEW).

VITA

2016 Bachelor of Arts in Psychology, Columbia University

2023 Master of Science in Clinical Psychology, San Diego State University

2025 Predoctoral Intern, Yale School of Medicine

2026 Doctor of Philosophy in Clinical Psychology, University of California San Diego/San Diego State University

ABSTRACT OF THE DISSERTATION

Reinforcement learning in eating disorders: Applications of computational modeling to behavioral and neural data

by

Carina S. Brown

Doctor of Philosophy in Clinical Psychology

University of California San Diego, 2026
San Diego State University, 2026

Professor Christina Wierenga, Chair

Rationale: Eating disorders are highly debilitating, and sometimes fatal, psychiatric conditions linked to alterations in reinforcement learning that may maintain illness and interfere with treatment. Computational psychiatry frameworks offer a novel approach for eliciting latent learning and decision-making mechanisms implicated in psychopathology and show promise for

optimizing treatments. This dissertation applied computational methodology (e.g., theory-driven mathematical modeling) and neuroimaging to investigate reinforcement learning across eating disorder diagnoses, contexts, and valences to characterize their role in eating disorder phenomenology.

Study 1 showed that adolescents with anorexia nervosa restricting type (AN-R; $n = 36$) displayed augmented habit-based reward learning compared to age-matched adolescent controls ($n = 28$), with higher goal-directed reward learning associated with greater orbitofrontal cortex-to-nucleus accumbens functional connectivity. This supports theories that suggest dietary restriction in AN-R may be habit-based.

Study 2 investigated valence-dependent learning asymmetry, or learning bias, in young adults with AN-R, AN binge-eating/purging type (AN-BP) and bulimia nervosa (BN; $n = 54$, mean age = 22.4) compared to undergraduates ($n = 40$, mean age = 20.1), using a probabilistic gambling card game. Participants with a binge-eating/purging phenotype (AN-BP, BN) showed slower learning rates and reduced learning bias compared to both participants with a restricting phenotype (AN-R) and to undergraduates. Learning bias was negatively correlated with binge eating severity across phenotypes, further suggesting that attenuated positive learning bias is associated with binge eating.

Study 3 examined aversive interoceptive reversal learning during breathing resistance in adults with AN-R, AN-BP and BN ($n = 67$) compared to age-matched controls ($n = 25$). Participants with AN-R showed faster learning rates to seen and unseen cues, compared to participants with a

binge-eating/purging phenotype (AN-BP, BN). Higher learning rates in this group were associated with lower insula-to-anterior prefrontal cortex functional connectivity, with the control group showing the opposite association. Results support theories suggesting faster aversive interoceptive learning contributes to restrictive eating.

Relevance: Collectively, these transdiagnostic studies demonstrate alterations in latent learning mechanisms across eating disorder phenotypes, developmental stages, valences, and contexts that may critically inform not only conceptualizations of eating disorder pathology, but also efforts to develop idiographic treatments and precision psychiatry practices.

INTRODUCTION

Eating disorders are dangerous psychiatric conditions, characterized by increased risk of mortality and chronicity of illness, debilitating long-term outcomes, medical comorbidity, and impaired psychosocial functioning (Arcelus et al., 2011; Mitchell & Crow, 2006; Solmi et al., 2024). Despite the substantial public health concern these outcomes engender, treatments are often unsuccessful or result in eventual relapse, with one study reporting that a third of participants with either anorexia nervosa (AN) or bulimia nervosa (BN) failed to achieve recovery at 22-year follow-up (Eddy et al., 2017) and others reporting a 35-41% relapse rate within 12-18 months of discharge in AN (Carter et al., 2004, 2012). Importantly, gold-standard eating disorder treatments (e.g., Cognitive Behavioral Therapy – Enhanced) rely on the assumption that learning is intact, with much of the therapeutic protocol taking behaviorist and cognitivist approaches intended to teach patients that behaviors and beliefs about food, weight, and body shape must be re-learned. And yet these approaches may not result in adequate treatment gains if learning is altered (Reilly et al., 2019). It is therefore noteworthy that prior research has identified abnormalities in learning within individuals diagnosed with AN, as well as other eating disorders (Christian & Levinson, 2022; Schaefer & Steinglass, 2021), establishing a need for greater understanding of how aberrant learning processes manifest in multiple domains of eating pathology, for example learning processes that may contribute to repeated dietary restriction or to binge-eating/purging cycles. Such investigations may not only provide insight into maintaining mechanisms that perpetuate illness, but may also inform future treatment development, with a focus on correcting maladaptive or inefficient learning processes contributing to illness (e.g., fear of weight gain).

Classically behaviorist accounts of learning and value-based decision-making stress the impact of learned associations between stimuli, responses, and outcomes, largely linked to Pavlovian and instrumental learning theories that are nested within the broader framework of reinforcement learning (Rangel et al., 2008; Sutton & Barto, 1998). Reinforcement learning, generally speaking, posits that rewarded behaviors will be repeated, although how an agent learns and exploits information it learns will differ. While a Pavlovian framework supports stimulus-outcome learning related to automatic behavioral response (e.g., salivating) to a cue (e.g., bell) as a result of learned expectations about outcomes (e.g., food), outcomes in the context of instrumental learning are contingent upon learned actions (e.g., making lunch) an agent takes in response to a cue (e.g., hunger) in order to receive or avoid a particular outcome (e.g., food), supporting stimulus-response and response-outcome learning.

The progression of eating pathology has been expressed within these frameworks, as choices about food, diet, or exercise can be viewed as value-driven decisions that are learned through sequences of rewarding and punishing outcomes (Pearce et al., 2022). For example, food preferences, and ultimately food choices, may be initially developed through acquired learning that particular food items are associated with positive or negative outcomes. As an individual learns, they build outcome expectancies based on reinforcement and punishment histories that influence their beliefs about the actions they can take and stimuli that predict a particular outcome, shaping and motivating future behavior (Bolles, 1972). Thus, an individual with a history of positive experiences with palatable food, purging, or restriction of food may come to expect future rewarding experiences (Schaefer & Steinglass, 2021). In the case of optimal learning, expectation violations, called prediction errors (PE), result in an update of expectations and are encoded neurally through phasic dopaminergic firing in the striatum; specifically,

positive PEs increase firing, negative PEs decrease firing, and expected outcomes remain at the tonic level of firing (Collins, 2018; Montague et al., 1996). Dopamine signaling is also involved in the modulation of corticostriatal plasticity, which can result in strengthening of action selection pathways that motivate behavior (Adamantidis et al., 2011; Collins, 2018; Hamid et al., 2015; Kravitz et al., 2012; Tai et al., 2012). In the case of non-optimal learning, these processes can become disrupted, leading to aberrant decision-making and behavior.

Research investigating cognition in eating disorders has identified the presence of abnormalities in learning and decision-making across reward and punishment contexts (C. S. Brown et al., 2024; Guillaume et al., 2015; Radzikowska et al., 2025). For example, there is growing evidence that deficits in feedback learning in both underweight and weight-restored AN is fundamental to manifestations of eating disorder pathology (Foerde & Steinglass, 2017; Ritschel et al., 2017; Shott et al., 2012). Indeed, several studies point to the importance of altered reward learning in eating disorders (Guillaume et al., 2015; Wu et al., 2016), including the role of elevated reward PE response in the insula and caudate during Pavlovian learning in individuals with AN compared to controls, even following weight restoration (DeGuzman et al., 2017; Frank et al., 2012, 2018), and attenuated PE response in insula, dorsal striatum, and orbitofrontal cortex in individuals with BN (Frank et al., 2011). Punishment learning has also been highlighted as a critical component of neurocognition in individuals with AN, as prior work has shown increased learning following punishing outcomes and augmented activity in the dorsal anterior cingulate cortex during negative feedback within the context of instrumental reversal learning (Bernardoni et al., 2017; Geisler et al., 2017). Altogether, there is growing evidence to support the hypothesis that learning processes are disrupted in eating disorders, however more research is needed to identify whether abnormalities exist across eating disorder subtypes (i.e.,

restricting versus binge-eating/purging), whether there are valence-specific (i.e., positive versus negative) or domain-specific (i.e., type of learning) deficits, and how learning is linked to symptomatology.

Although there is a long history of utilizing neuropsychiatric testing and behavioral tasks to investigate executive functioning deficits in eating disorders, the field is still beleaguered by inconsistent findings, gaps in the literature, and vague conceptualizations of where alterations exist in cognitive processing and choice behavior. Computational psychiatry offers a novel approach of leveraging mathematical modeling to identify latent behavioral and neural mechanisms that contribute to psychopathology, particularly pertaining to higher-order cognitive functions, such as decision-making and learning (Haynos et al., 2022). This approach aims to map cognition and behavior to neural circuitry by mathematically representing brain-based computations that occur during various psychological states (e.g., decision-making, learning, emotions; Friston, 2016; Huys et al., 2021; Redish & Gordon, 2016). Because algorithmic representation of human behavior can elicit quantification of hidden behavioral or brain states, it has been suggested that these novel approaches will not only provide a platform for more nuanced, more effective psychotherapies, but may also speed the development of idiographic therapies and precision medicine practices (Adams et al., 2016; Moutoussis et al., 2018; Nair et al., 2020). Application of these methods to various psychiatric conditions, including substance use, depression, and anxiety, has been burgeoning (Chen et al., 2015; Gueguen et al., 2021; Pike & Robinson, 2022); and yet, few studies in the eating disorder literature have incorporated this approach, leaving a large gap in collective knowledge of process-level impairments in learning that may contribute to eating pathology. Therefore, this dissertation aimed to apply various computational models, grounded in reinforcement learning theory, to disparate learning

paradigms in individuals with primarily restricting and primarily binge-eating/purging eating disorder presentations to understand if learning and decision-making in this population is impaired, and, if so, how these processes are impacted in several domains, including: habit and goal-directed learning, reward and punishment learning asymmetries, and aversive interoceptive learning.

To that end, **Study 1** examined the relative contributions of habit and goal-directed learning in adolescents with AN restricting subtype (AN-R) and AN binge-eating/purging subtype (AN-BP) compared to adolescent healthy controls (HC). Goal-directed learning is characterized by flexible decision-making and choice behavior that adaptively responds to changes in environmental contingencies, whereas habit learning is characterized by rigidity and inflexibility (Dolan & Dayan, 2013). Habit theories of AN posit that disorder relevant behaviors, such as restriction, begin as goal-directed pursuits of weight loss, but become habitual over time as behavior becomes entrenched, more rigid, and impervious to change (Walsh, 2013). In line with this hypothesis, prior research investigating these constructs using dual-system reinforcement learning have shown links between deficits in goal-directed learning and compulsivity (V. M. Brown et al., 2020; Gillan et al., 2016; Voon, Baek, et al., 2015), supporting the notion that habit is a primary mechanism involved in entrenched behaviors. This same attenuation in model-based learning has been reported in a sample of individuals with AN, a finding that did not resolve at weight restoration and may be suggestive of a trait behavioral marker or a "scar" of the illness (Foerde et al., 2021). Although these constructs have already been explored in a mixed age sample of individuals with AN, less is known about early stages of the disorder. AN tends to onset in adolescence, making this developmental stage a particularly sensitive time to study learning and decision-making processes. Model-based learning begins to

come online at this stage as well (Decker et al., 2016), meaning that alterations in arbitration between goal-directed and habit systems during this stage of development may yield greater insights into progression of the illness. We investigated this arbitration in adolescents with AN using a well-established dual-system reinforcement learning model and behavioral task (Daw et al., 2011), which we correlated with resting-state functional connectivity in neural regions linked to goal-directedness, including orbitofrontal cortex and nucleus accumbens, and habit, including putamen and supplementary motor area (L. S. Morris et al., 2016).

Study 2 applied a reinforcement learning framework to the study of valence-dependent learning asymmetry (i.e., learning bias), or the relative weighting of learning from gain and loss (Cazé & Van Der Meer, 2013; Eldar et al., 2016; Gershman, 2015; Michely et al., 2022), in adults with restricting (i.e., AN-R) and binge-eating/purging (i.e., AN-BP and BN) eating disorder phenotypes, compared to adult undergraduates (UG). Individuals with eating disorders, including AN and BN, show higher trait harm avoidance and sensitivity to punishment, suggesting a possible risk or maintaining factor that propels eating pathology (A. Harrison et al., 2010; Haynos et al., 2020). Higher sensitivity to punishment may suggest overweighting of learning to be particularly responsive to losses; indeed, prior work has found evidence of faster learning in response to negative outcomes in individuals with AN (Bernardoni et al., 2018, 2021; Filoteo et al., 2014; Wierenga, Bischoff-Grethe, et al., 2025), as well as individuals with anxiety and depression (Bishop & Gagne, 2018; Villano et al., 2023). The relative weighting of reward, however, is more debated, with some conceptualizations of binge-eating/purging symptomatology highlighting augmented reward sensitivity as a possible mechanism involved in these behavioral patterns (Chan et al., 2014; A. Harrison et al., 2010; P. Monteleone et al., 2014). **Study 2** therefore explores learning bias in restrictive and binge-eating/purging eating disorder

phenotypes to understand unique roles of valence-dependent learning in different presentations of eating pathology.

Study 3 is similar, in that we again utilized a reinforcement learning framework in the study of distinct eating disorder phenotypes (restricting versus binge-eating/purging); however, here we apply this framework to the concept of aversive interoceptive learning, or the process by which one learns from internal bodily states (Craig, 2002; Seth, 2013; Seth & Critchley, 2013). Disordered eating behaviors are thought to be maintained through altered interoceptive processing that contributes to body image disturbance, impaired ability to recognize hunger/satiety cues, and altered sensory perception (e.g., feeling bloated) (Badoud & Tsakiris, 2017; Berner et al., 2018; Kaye et al., 2009). Thus, disruption in interoception and interoceptive processes has been considered a core feature of eating disorder pathology and has been shown to be a central feature in eating disorder symptom-level networks (T. A. Brown et al., 2020; Levinson & Williams, 2020). Despite this, little is known about latent processes that govern how interoception is shaped and, more specifically, how individuals learn about and process internal sensory signals (Zucker & Bulik, 2020). More recent theories platform the concept of interoceptive predictive coding, which posits a bi-directional encoding of bottom-up sensory information and top-down sensory expectations, the integration of which shapes interoceptive experience (Seth & Critchley, 2013). Disruptions in this process are hypothesized to result in pathology, including contributions to maintenance of psychiatric disorders (Hakamata et al., 2010; Paulus & Stein, 2010; Van den Bergh et al., 2017). Consequently, **Study 3** was conducted as an effort to understand how individuals with eating disorders learn from aversive internal sensory stimuli and represents a first step towards building empirical knowledge about how interoceptive experiences are shaped behaviorally and how they are represented in the brain.

These three studies contribute to collective understanding of latent features involved in reinforcement learning across eating disorder phenotypes, developmental stages, valences, and contexts. Consequently, this body of work extends research investigating eating disorder pathology through the lens of cognitive science, neuroscience, and neuropsychology by adopting a computational psychiatry approach that relies on model-based formulations of cognition and behavior. Computational methods have been heralded as a novel and needed pathway towards development of precision psychiatry treatments (Friston et al., 2017; Redish & Gordon, 2016), for reframing "deficit" theories of psychopathology, and have been specifically proposed as needed framework for the next wave of research and treatment development in AN (Haynos et al., 2022). In line with that perspective, this dissertation represents an initial step towards nuanced, phenotypic understanding of eating disorders, which will aid in development of efficacious treatments for these populations.

Chapter 1 RELATIVE RELIANCE ON MODEL-FREE VERSUS MODEL-BASED LEARNING IN ADOLESCENT ANOREXIA NERVOSA

ABSTRACT

Alterations in learning and decision-making systems are thought to contribute to core features of anorexia nervosa (AN), a psychiatric disorder characterized by persistent dietary restriction and weight loss. Instrumental learning theory identifies a dual-system of habit and goal-directed decision-making, linked to model-free and model-based reinforcement learning algorithms. Difficulty arbitrating between these systems, resulting in an over-reliance on one strategy over the other, has been implicated in compulsivity and extreme goal pursuit, both of which are observed in AN. Characterizing alterations in model-free and model-based systems, and their neural correlates, in AN may clarify mechanisms contributing to symptom heterogeneity (e.g., binge-eating/purging symptoms). This study tested whether adolescents with restricting AN (AN-R; $n = 36$) and binge-eating/purging AN (AN-BP; $n = 20$) differentially utilized model-based and model-free learning systems compared to a healthy control group (HC; $n = 28$) during a Markov two-step decision-making task under conditions of reward and punishment. Associations between model-free and model-based learning and resting-state functional connectivity between neural regions of interest, including orbitofrontal cortex (OFC), nucleus accumbens (NAcc), putamen, and supplementary motor area (SMA) were examined. AN-R showed higher utilization of model-free learning compared to HC for reward, but attenuated model-free and model-based learning for punishment. In AN-R only, higher model-based learning was associated with stronger OFC-to-left NAcc functional connectivity, regions linked to goal-directed behavior. Greater utilization of model-free learning for reward in AN-R may differentiate this group, particularly during adolescence, and facilitate dietary restriction by prioritizing habitual control in rewarding contexts.

INTRODUCTION

Anorexia nervosa (AN) is a debilitating and potentially life-threatening disorder with one of the highest mortality rates among psychiatric conditions (Arcelus et al., 2011; Braun et al., 1994; Harris & Barraclough, 1998; Mitchell & Crow, 2006). AN typically onsets in adolescence, with at least a third of affected individuals developing a chronic course of illness into adulthood (Grange & Loeb, 2007), yet mechanisms contributing to progression of illness are poorly understood. Recent conceptualizations posit a role of habit and goal-directed behavioral systems in the pathogenesis and maintenance of the disorder, although their relative contribution is debated (Steding et al., 2019; J. E. Steinglass & Walsh, 2016; Uniacke et al., 2018; Walsh, 2013). Goal-directed control consists of flexible decision-making, in which the value of specific outcomes can change depending on shifting environmental contingencies. Conversely, habitual control consists of inflexible decision-making, in which choice behavior is difficult to change, even during instances of outcome devaluation (Dolan & Dayan, 2013).

The central AN symptom of dietary restriction can be described as highly goal-directed, with persistent food restriction in service of long-term rewards (i.e., weight loss) indicating an effortful and goal-driven strategy. Alternatively, the rigidity with which individuals with AN uphold their caloric restriction and the cognitive inflexibility that is observed in their decision-making, might indicate greater use of habit-driven strategies. It has been proposed that what begins as a largely goal-directed pursuit of weight loss, driven by action-outcome learning (e.g., restriction becomes associated with weight loss outcomes), ultimately becomes habitual as repeated actions are overtrained and stimulus-response learning results in specific cues being automatically associated with disorder-relevant actions (e.g., palatable food cues become associated with restriction (Walsh, 2013). Many individuals with AN also experience episodes of binge eating and purging, behaviors thought to reflect compulsivity (e.g., inflexible, repetitive

behavior), and may too reflect greater reliance on habit. Thus, there is a question of how individuals with AN arbitrate between these systems and how this corresponds to restriction versus binge-eating/purging symptomatology.

A compelling computational framework for this dual-system theory of goal-directed and habit decision-making exists in a class of reinforcement learning algorithms known as model-based and model-free learning (Daw et al., 2005; Keramati et al., 2011; Pezzulo et al., 2013). These algorithms provide instantiations of neural computations that can be behaviorally modeled using variants of multistep Markov reward-based decision-making tasks that have been validated in both human (Daw et al., 2011) and animal studies (Miller et al., 2022).

Model-based learning algorithms consist of cognitive maps of state-action space that track environmental contingencies, such as transition probabilities when moving from one state to another, and consider these maps when selecting actions (Tolman, 1948). Consequently, this system is behaviorally akin to planning and, although flexible, imposes a considerable computational cost on the decision-maker (Otto, Gershman, et al., 2013). The orbitofrontal cortex (OFC), which tracks value computations in response to feedback or shifting motivational and affective states (J. O'Doherty et al., 2001; Stalnaker et al., 2015; Valentin et al., 2007; Wikenheiser & Schoenbaum, 2016), and the ventromedial prefrontal cortex (vmPFC) are implicated in model-based learning (Balleine & O'Doherty, 2010; Mcdannald et al., 2012; L. S. Morris et al., 2016; Voon, Derbyshire, et al., 2015). Greater intrinsic resting functional connectivity between medial OFC and nucleus accumbens (Nacc), which tracks prediction errors and anticipated reward and punishment outcomes (Groenewegen et al., 1999; Pagnoni et al., 2002; Talmi et al., 2008; Van der Meer & Redish, 2011), have been correlated with greater model-based behavior (L. S. Morris et al., 2016).

In contrast, model-free learning algorithms are linked to temporal difference learning frameworks, where value estimates based on outcomes are stored across trials and selected actions are dependent on the history of reward receipt, without consideration of environmental contingencies (Sutton & Barto, 1998). This system is similar to trial-and-error learning and is not easily adaptable, however it can be quite efficient, akin to habit (Gillan et al., 2015). Although neural signatures of model-based and model-free learning overlap in the medial prefrontal cortex and the ventral striatum, the putamen has also been implicated in habit behavior during sequential decision-making and outcome devaluation (Daw et al., 2011; de Wit et al., 2012; Doll et al., 2015; Mcdannald et al., 2012; Nachev et al., 2008; Patterson & Knowlton, 2018; Wan Lee et al., 2014), with putamen-supplementary motor area (SMA) functional connectivity, as well as increased neurite density in the left posterior putamen and right SMA, associated with model-free learning (L. S. Morris et al., 2016).

Across psychiatric conditions, there is emerging evidence suggesting that disorders of compulsivity (e.g., obsessive-compulsive disorder, substance use) demonstrate deficits in model-based learning during sequential decision-making (V. M. Brown et al., 2020; Gillan et al., 2016; Janssen et al., 2020; Voon, Baek, et al., 2015; Voon, Derbyshire, et al., 2015). Critically, prior work examining model-based and model-free learning in a mixture of community and clinical samples endorsing symptoms of eating pathology (e.g., compulsive eating), also demonstrate attenuation in model-based learning (V. M. Brown et al., 2020; Foerde et al., 2021; Gillan et al., 2016; Janssen et al., 2020; Voon, Derbyshire, et al., 2015). Only one such study was conducted in a sample of individuals with full-threshold AN and found evidence of deficits in model-based learning for food and money that were not remediated at weight restoration, reflecting a potential trait behavioral marker of the disorder.

The current study aimed to characterize reliance on model-free and model-based reinforcement learning in AN by clarifying several key issues. First, because AN typically onsets during adolescence and is associated with high risk of chronicity, it is critical to uncover developmental mechanisms that contribute to maintenance of the disorder. Moreover, this developmental period consists of a time frame when model-based learning begins to come online (Decker et al., 2016), highlighting a possible treatment target if the relative balance between model-free and model-based systems go awry during initial stages of psychiatric illness. Second, there is heterogeneity in the AN phenotype, with some patients exhibiting only dietary restriction (AN-R) and others also exhibiting binge-eating/purging behaviors (AN-BP). Divergence of symptoms may suggest dissociable deficits in decision-making in individuals with AN-R versus AN-BP. Finally, while most studies have focused exclusively on reward learning, recent research suggests avoidance motivation promotes goal-directed learning (Eryilmaz et al., 2017), and in OCD, learning bias shifts towards goal-directed control under conditions of loss (Voon, Baek, et al., 2015). Although reward deficits are well-established in AN (Brooks et al., 2012; Fladung et al., 2013; Keating et al., 2012; A. M. Monteleone et al., 2017; O'Hara et al., 2015; Wu et al., 2016), altered punishment motivation and learning might be particularly salient (Bischoff-Grethe, McCurdy, Grenesko-Stevens, Irvine, et al., 2013; Glashouwer et al., 2014; A. Harrison et al., 2010; Jappe et al., 2011; Matton et al., 2017), as AN is characterized by high levels of harm avoidance, exaggerated neural response to losses, and worse treatment outcomes associated with poorer loss-related learning (Bernardoni et al., 2017; Bischoff-Grethe, McCurdy, Grenesko-Stevens, Irvine, et al., 2013; Wierenga et al., 2022). However, this may differ by subtype as eating disorders characterized by binge-eating/purging symptoms demonstrate increased limbic-striatal reward response (Brooks et al., 2012; Fladung et al., 2013; Keating et al., 2012; O'Hara

et al., 2015; Wierenga et al., 2014; Wu et al., 2016), potentially indicating that both subtype and outcome valence may affect utilization of model-based and model-free learning systems.

To elucidate these questions, we conducted a monetary outcomes version of the two-step sequential decision-making task under conditions of reward and punishment in a sample of female adolescents with AN and demographically-matched healthy controls (HC). Participants also completed a separate resting-state functional magnetic resonance imaging scan (rsfMRI) to ascertain whether corticostriatal goal-directed and habit circuits were differentially associated with model-based and model-free learning in AN compared to HC. We hypothesized that adolescents with AN would have attenuated model-based learning compared to HC (especially AN-BP), consistent with increased compulsivity, and greater reliance on model-free learning (especially AN-R), and with the notion that dietary restriction is habit-based. With regards to neural functioning, we hypothesized that greater utilization of model-free learning would be linked to stronger functional connectivity between putamen and SMA, while greater utilization of model-based learning would be linked to stronger functional connectivity between OFC and NAcc. Finally, we explored whether model-based and model-free learning for reward and punishment were associated with clinical variables.

METHODS

Participants

Participants were all female and adolescent, with 36 meeting DSM-5 criteria for restricting type AN (AN-R; age range = 13.67-18.08), 20 meeting DSM-5 criteria for binge-purge type AN (AN-BP; age range = 13.08-18.17), and 28 being demographically matched healthy controls (HC; age range = 13.42-17.83). Approximately 5% of this sample (n = 4) completed the reinforcement learning task but did not complete the rsfMRI scan. One subject

was excluded due to excessive motion during the rsfMRI scan. Participants included in resting-state analyses included 35 AN-R, 17 AN-BP, and 27 HC. Individuals with AN were recruited from the University of California, San Diego Eating Disorders Center for Treatment and Research (Reilly et al., 2020), and healthy controls were recruited from the community. The study was approved by the University of California San Diego Institutional Review Board, and all participants 18 years old and above were provided written informed consent. Participants younger than 18 years old gave written informed assent and their parents provided written informed consent.

To participate in the study, individuals in the AN group must: (a) have met DSM-5 criteria for AN-R or AN-BP, (b) be medically stable, per American Academy of Pediatrics and the Society of Adolescent Medicine requirements; and (c) be at least 75% of adjusted ideal body weight (IBW). HCs were recruited from the San Diego community and were administered the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children (KSADS; (Kaufman et al., 1997) and the Eating Disorder Examination (EDE; Cooper & Fairburn, 1987) to confirm the absence of symptoms related to any eating disorder or Axis I psychiatric disorder. HCs were included in the study if they had: (a) no stigmata suggestive of an eating disorder, (b) no current or past psychiatric (definitive Axis I disorder) or medical illness; (c) no use of psychoactive drugs; and (d) maintained 90% to 120% ideal body weight since menarche. Exclusion criteria for all subjects included the following: (a) MRI contraindications, (b) psychotic illness/other mental illness requiring hospitalization; (c) current dependence on drugs or alcohol; (d) physical conditions (e.g., diabetes mellitus, pregnancy) known to influence eating or weight; (e) neurological disorder or history of head injury with >30 min loss of consciousness;

(f) neurodevelopmental disorder; (g) primary obsessive compulsive disorder or primary major depressive disorder (within last 3 months as measured by the KSADS).

Measures

The State-Trait Anxiety Inventory – Trait subscale (STAI-T; Spielberger, Gorsuch, & Lushene, 1970) is a 20-item self-report measure that assesses trait anxiety on a 1- 4 Likert scale, with higher values indicating greater severity of symptoms.

The Beck Depression Inventory – II (BDI-II; Beck et al., 1996) is a 21-item self-report measure assessing severity of depression symptoms. Items are rated on a 0 – 3 Likert scale.

The Temperament and Character Inventory (TCI; Cloninger et al., 1994) is a 226-item self-report measure assessing seven dimensions of personality. For the purposes of the current study, we only include the Harm Avoidance subscale, designed to correspond with the magnitude of inhibitory responding towards aversive stimuli.

Sensitivity to Punishment and Reward Questionnaire (SPSRQ; Torrubia et al., 2001) is a 48-item self-report measure, including Punishment and Reward subscales, designed to operationalize constructs from reinforcement learning theory. We included both subscale scores in our analyses. Items are rated as binary (yes/no) responses.

Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS; Carver & White, 1994) is a 24-item self-report measure, which includes a Behavioral Inhibition (BIS) scale and three intercorrelated Behavioral Activation (BAS) subscales: Reward Responsiveness, Drive, and Fun Seeking. We only included BIS and BAS-Reward Responsiveness in our analyses.

Adult Temperament Questionnaire (ATQ; Evans & Rothbart, 2007) is a 77-item self-report questionnaire measuring temperamental constructs, including Effortful Control, Negative Affect, Extraversion, and Orienting Sensitivity. We only included the Effortful Control subscale in our analyses, reflecting self-regulation and executive control.

Procedure

Two-Step Sequential Decision-Making Task

Participants completed a modified two-step sequential decision-making task (Daw et al., 2011), which was previously validated in an adolescent sample (Decker et al., 2016). They performed the task under conditions of reward, where they either won $1 or won nothing, and punishment, where they either lost $1 or lost nothing (Figure 1.1A). The order of the conditions was counterbalanced within each group, with each condition consisting of four blocks and a total of 200 trials.

Participants were tasked with collecting "space treasure" that would amount to a $1 gain in the reward condition or no monetary loss (of $1) in the punishment condition. On each trial, participants made choices at two stages. In the first stage they made a choice between two spaceships that traveled to one of two planets. Each spaceship traveled more frequently to one planet (70%) and less frequently to the other (30%); for example, the blue rocketship had a 70% chance of traveling to the red planet (a common transition) and a 30% chance of traveling to the purple planet (a rare transition), while the green rocketship had the opposite set of probabilities. At the second stage (i.e., when participants are on a planet), they were asked to choose between two aliens. Depending on which alien they chose, they were either presented with space treasure (+$1) or with an empty circle for the reward condition. For the punishment condition, they were either presented with an empty circle or space treasure with a red 'x' over it (-$1). Probability of winning any given trial was determined by a slowly drifting probability bounded between 0.2 and 0.8. During the task, participants were allotted three seconds to make a decision, followed by a one-second animation, a one-second feedback period, and a one-second inter-trial interval. The game itself consisted of four blocks separated by breaks and comprised of 200 trials.

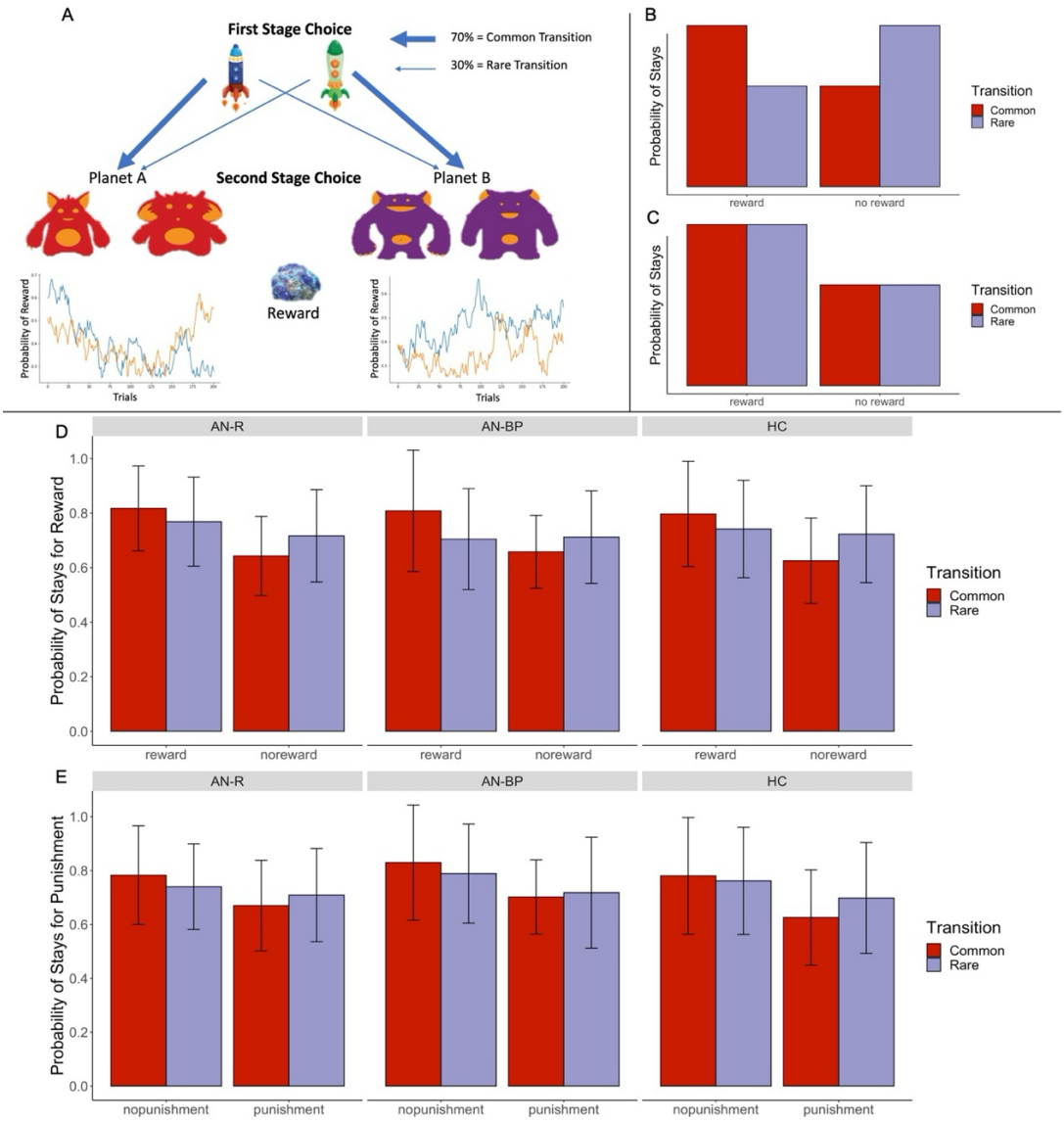


Figure 1.1: Design of the reward version of the sequential spaceship task based on Decker et al. (2016; A) and sample model-based (B) and model-free (C) behavior. On each trial, participants chose between two spaceships (first-stage choice), which was followed by a probabilistic transition (70% or 30%) to Planet A or Planet B. Then participants chose between two aliens (second-stage choice) and were rewarded with space treasure or nothing. The probability of winning space treasure is presented as a function of trial for each alien. The bar graphs show, for idealized model-free and model-based learners, the probability of making the same choice on the next trial (i.e., a first-stage stay) as a function of the outcome and transition type (common or rare) of the previous trial. Probability of first-stage stays for the reward (D) and punishment (E) conditions for each group. We visualized the effect of model-based and model-free learning on raw choice data for reward and punishment conditions by plotting stay/switch behavior as a function of outcomes and transition types (i.e., common and rare transitions). The proportion of first-stage stay choices is graphed as a function of outcome of the previous trial for group, separately for trials following common and rare transitions. The error bars represent ±1 SD.

Functional Magnetic Resonance Imaging (fMRI)

Resting-state fMRI was acquired separately on a 3.0 T GE MR750 scanner equipped with quantum gradients providing echo planar capability, using a Nova Medical 32 channel head coil (maximum gradient strength: 50 mT/m, slew rate: 200 T/m/s). An eyes-open protocol was implemented, with participants instructed to look at fixation crosshair for the duration. The 1 hour session included: a three-plane localizer; sagittally acquired (0.8 mm slice thickness, FOV=256x240 mm) T1-weighted (MPRAGE PROMO, TE=2.3 ms, flip angle=8o, matrix=320, 2x in-plane acceleration, scan time=7:50) and T2-weighted (3D CUBE, TE=60 ms, variable flip angle, matrix=320, 2x in-plane acceleration, scan time=6:51) whole brain scans for alignment and morphometry; and T2* weighted echo-planar imaging (EPI; 749 volumes, TR= 800 ms, TE=37 ms, flip angle=52°, FOV=180 (RL) x 208 (AP) mm, 72 2 mm thick axial slices, multiband factor=8).

Statistical Analysis

Analysis of Raw Choice Data

Consistent with prior work (Daw et al., 2011; Otto, Raio, et al., 2013), mixed-effects logistic regression analysis was performed on the raw choice data using *lme4* package (Version 1.1-8; Bates et al., 2015) for the R software environment (Version 4.0.1; R Development Core Team, 2015). We estimated separate models for reward and punishment conditions. This model predicted first-stage choices (i.e., stay versus switch with respect to the previous response) as a function of previous outcome (i.e., won/lost in previous trial), previous transition type (i.e., rare/common), group, and their interaction. All two- and three-way interactions were defined as

fixed effects, with a random intercept and subject-level adjustments to outcome, transition, and their interaction set as random slopes. The lme4 syntax is as follows:

$$Stay_{Reward} \sim Outcome * Transition * Group + (Outcome * Transition + 1| Subject)$$

From this analysis, we estimated a model-free effect, approximated by a main effect of previous outcome on stay probability, as well as a model-based effect, approximated by a significant outcome-by-transition interaction.

Percentage of optimal choices was calculated for the second stage decision, specifically the choice between two second-stage stimuli. This was done separately for each condition and for every subject. Means for each group, and for each condition, were subjected to one sample t-tests, comparing to chance (i.e., 50%).

Computational Model

We fit participants' choice data to a popular reinforcement learning model previously used to model two-stage choice data (Daw et al., 2011; Decker et al., 2016; Otto, Raio, et al., 2013). In this "hybrid" model, participants' choices are theorized to reflect a weighted combination of model-free and model-based learning. This model has been well-validated in a range of psychiatric disorders, with prior work establishing its reliability in a range of psychiatric disorders (V. M. Brown et al., 2020). The degree to which participants' responses are governed by principles of model-free or model-based learning is controlled, respectively, by β_{MF} and β_{MB} parameters. Both parameters are bounded at zero and larger values reflect increased model-free or model-basedness.

The hybrid reinforcement learning model (Decker et al., 2016) is comprised of two

distinct components: a model-free SARSA (λ) temporal difference algorithm and a model-based "tree-search" reinforcement learning algorithm (Bellman's equation), enabling representation of all possible choice options and associated outcomes (Sutton & Barto, 1998). The task defines three states (i): one state at stage one (s_A) and two states at stage two (s_B, s_C), where each state is associated with a set of actions (a_A, a_B). Both algorithms track state-action value functions, Q(s,a), that represent the expected value for each state-action pair. For every trial, t, the participant makes a stage 1 choice ($a_{1,i,t}$), followed by a transition to the second stage where they again make a choice ($a_{i,2,t}$), that results in a possible reward ($r_{i,t}$).

Model-free algorithm

The model-free temporal difference algorithm updates the state action values according to the following rule:

$$Q_{MF}(s_{i,t}, a_{i,t}) = Q_{MF}(s_{i,t-1}, a_{i,t-1}) + \alpha\delta_{i,t}$$

$$\text{Where, } \delta_{i,t} = r_{i,t} + Q_{MF}(s_{i+1,t}, a_{i+1,t}) - Q_{MF}(s_{i,t-1}, a_{i,t-1})$$

Here, the reward prediction error (RPE) is defined by δ, and the learning rate is defined by α. For stage one, $r_{1,t} = 0$ and the RPE is calculated with respect to the estimate of the second stage action value, $Q_{MF}(s_{2,t}, a_{2,t})$. For the second stage, $r_{2,t} = 0$, for a loss, $r_{2,t} = 1$, for a reward, or $r_{2,t} = -1$ for a punishment and the delta rule simplifies to $\delta_{i,t} = r_{i,t} - Q_{MF}(s_{2,t-1}, a_{2,t-1})$.

Model-based algorithm

The model-based algorithm was applied only at the first stage and updates the state-action pair by incorporating information about the 70/30 transition probability structure:

$$Q_{MB}(s_{1,t}, a_{j,t}) = P(s_B|s_A, a_j) \max_{a\epsilon\{a_A,a_B\}} Q_{TD}(S_B, a) + P(s_C|s_A, a_j) \max_{a\epsilon\{a_A,a_B\}} Q_{TD}(S_C, a)$$

Choice rule

We applied a softmax choice rule as our policy, calculating a probability to each action according to the combination of both Q_{MB} and Q_{MF} values taken from the model-based and model-free algorithms described above. Each value was weighted with its own inverse temperature parameter, β_{MB} and β_{MF}, to ascertain the relative contribution of model-based and model-free valuations:

$$P(a_{i,t} = a|s_{i,t}) = \frac{exp[\beta_{MF} * Q_{MF}(s_{i,t}, a) + \beta_{MB} * Q_{MB}(s_{i,t}, a) + p * rep(a)]}{\sum a' exp[\beta_{MF} * Q_{MF}(s_{i,t}, a) + \beta_{MB} * Q_{MB}(s_{i,t}, a) + p * rep(a')]}$$

Here, rep(a) is an indicator function that takes on a value of 1 for a first-stage action that repeats the action taken in the previous trial. The "stickiness" parameter (p) is applied to measure perseveration (p>0) or switching (p<0) behavior by combining it with the indicator function. Model-based learning is only possible at the first stage, so the softmax function simplifies to only include Q_{MF} with its own inverse temperature, β_2.

Group level modeling

Model parameters were estimated with a hierarchical Bayesian model-fitting technique. To better understand how diagnostic group membership affected model parameters of interest, these parameters were determined on the basis of group membership (akin to regression) estimating coefficients (γ) that encoded how these parameters differed between each group (HC, AN-R, AN-BP):

$$\beta_j \sim \gamma_0 + \gamma_{ANR} is\_R_j + \gamma_{ANBP} is\_BP_j$$

where is_R and is_BP are dummy-coded variables reflecting whether participant j was (or was not) in the AN-R or AN-BP subgroup respectively. Thus, γ_0 encoded model weights in HCs and the other γ coefficients tested for differences between HCs and AN subgroups.

To reduce estimation issues arising from the hierarchical nature of the data, free parameters were reparametrized. Specifically, free parameters were initially drawn from a standard normal distribution (μ^{UT}, UT = untransformed) and then transformed into an appropriate scale based on estimated fixed effects (γ) and independently estimated random variance terms (σ):

$$\beta_{p,j} = \mu_{p,j} + \sigma_{p,j}\mu_{p,j}^{UT}$$

These parameters were further transformed to enforce appropriate bounds. Namely, inverse temperature parameter was exponentiated (to enforce [0,∞] bounds), learning rates were converted to a unit scale, and the stickiness parameter was converted to a unit scale and multiplied by an upper bound of 5. Hyperparameters γ and σ were given normal priors of $N(0,1)$ and $HN(0,.2)$ respectively.

Estimation

The joint distribution of model parameters was estimated separately for reward and punishment conditions. Markov Chain Monte Carlo (MCMC) techniques, specifically the No-U-Turn variant of Hamiltonian Monte Carlo, implemented in the Stan modeling language (Stan Development Team, 2015), was used to obtain samples from the conditional joint distribution. We ran four chains of 4,000 samples each, with the first 1,000 samples of each chain being discarded as burn-in, and no thinning applied.

To assess the statistical reliability of the observed parameter estimates, a 95% credibility interval (CI) was computed for each parameter. If the CI was wholly situated above or below zero, we determined with 95% confidence that the corresponding parameter was principally positive or negative, respectively. Moreover, we computed Bayesian *p*-values (P), which represent one minus the proportion of the posterior that falls above or below zero (depending on the sign of the median posterior value: below zero if $b < 0$ and above if $b > 0$). In line with the traditional interpretation of frequentist *p*-values, Bayesian *p*-values can be interpreted probabilistically as "there is a (P×100) percent chance that the effect is zero or a reversal of the central tendency."

Exploratory analyses with clinical variables

We extracted subject-level random effects from the hierarchical estimation procedure described above to examine the relationship between our computational parameters of interest (β_{MB} and β_{MF}) and clinical variables. These random effects estimates were non-linearly transformed before additional analyses, as they were not normally distributed as indicated by Shapiro-Wilk tests (*p*s<.001). To transform the data, we converted the estimates to percentile scores and applied a probit link function, converting to z-scores, which improved normality of the distributions (*p*s>.05). Transformed estimates were used as dependent variables in a series of

mixed-effects regression analyses with a group-by-condition-by-measure three-way interaction term. All analyses controlled for z-scored BMI, IQ, and age as additional covariates. Analyses were conducted and estimated using the R *lme4* package and family-wise Bonferroni-corrected to account for the exploratory nature of the analyses. Mixed-effects linear regression analyses were conducted with the following measures: Eating Disorder Examination (EDE), State-Trait Anxiety Inventory (STAI), Beck Depression Inventory – II (BDI-II), Temperamental Character Inventory – Harm Avoidance (TCI), Sensitivity to Punishment and Reward Questionnaire (SPSRQ), Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS), and Adult Temperament Questionnaire – Effortful Control (ATQ). The analysis involving the EDE was only conducted in AN-R and AN-BP, as the majority of HC had total scores of zero.

Resting-State Functional Magnetic Resonance Imaging Analysis

Images were processed using scripts developed by the Human Connectome Project freely available online (https://github.com/Washington-University/HCPpipelines). We adapted the HCP quality control metrics (Marcus et al., 2013), and HCP preprocessing requirements (Glasser et al., 2013), prior to using the HCP Workbench. Functional images were processed to remove spatial distortion, motion corrected, aligned to the subject's structural volume, bias field corrected, transformed to standard space, and smoothed using FSL and other tools developed for the HCP protocol. ICA + FIX (Beckmann & Smith, 2004) was applied to remove spatially specific structured noise, such as motion and transient head movement due to swallowing (Glasser et al., 2016). Subjects were excluded if any one parameter's average exceeded 2 standard deviations from the mean to ensure that group differences were not due to motion. One subject was removed from the dataset following preprocessing due to excessive motion.

To investigate whether resting-state functional connectivity is associated with model-based and model-free learning, we conducted an ROI-to-ROI analysis on an *a priori* network. Specifically, we selected regions that have been implicated in both habit and goal-directed learning, including the OFC, NAcc, bilateral putamen, and SMA. Connectivity analyses were conducted in *nilearn*, a python-based brain imaging toolbox (https://nilearn.github.io/index.html). Anatomical ROIs were identified using the Harvard-Oxford Cortical and Subcortical Atlases (Desikan et al., 2006; Frazier et al., 2005; Goldstein et al., 2007; Makris et al., 2006). Timeseries from each ROI were extracted and subjected to ROI-to-ROI analysis. Connectivity values (Fisher z-transformed correlation coefficients) between OFC-NAcc and putamen-SMA dyads were extracted from the connectivity map and entered into mixed-effects linear regression models. All analyses utilized transformed subject-level parameters of interest from the computational model. To examine links between model-based behavior and associated neural regions, model-based estimates (β_{MB}) were set as the dependent variable with OFC-NAcc connectivity, group, condition, a connectivity-by-group-by-condition 3-way interaction, z-scored age, BMI, and IQ entered as independent variables. We specified the same model for model-free estimates (β_{MF}) and putamen-SMA connectivity to examine associations between behavioral and neural regions implicated in habit learning. Because these analyses are linked to *a priori* hypotheses about the relationships and direction of relationships, we did not correct for multiple comparisons. Additionally, ROI-to-ROI connectivity was represented using plotting utilities from *nilearn* (Figure 1.3C).

RESULTS

Sample Characteristics

Groups did not differ on age, education, race, ethnicity, IQ, or length of illness (for AN-R and AN-BP only); however, they did differ on BMI, lowest lifetime BMI (for AN-R and AN-BP only), anxiety (STAI-T and STAI-S), depression (BDI), and eating disorder severity (EDE; Table 1.1).

Table 1.1: Demographic characteristics of each group. EDE – Eating Disorder Examination; WASI – Weschler Abbreviated Scale of Intelligence; TCI – Temperament and Character Inventory; BIS – Behavioral Inhibition System; BAS – Behavioral Activation System; SPSRQ – Sensitivity to Punishment and Sensitivity to Reward Questionnaire; ATQ – Adult Temperament Questionnaire; BDI – Beck Depression Inventory; STAI-T – State Trait Anxiety Inventory – Trait; STAI-S – State Trait Anxiety Inventory – State. AN-R – restricting type anorexia nervosa; AN-BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

Variable	AN-R (n = 36) *M (SD)/n* (%)	AN-BP (n = 20) *M (SD)/n* (%)	HC (n = 28) *M (SD)/n* (%)	$t/F/X^2$	*p*
Age	16.07 (1.36)	16.21 (1.48)	15.94 (1.33)	0.23	.80
BMI at time of study (kg/m^2)	19.94 (1.63)	21.70 (3.43)	21.10 (1.80)	4.50	.01
Lowest BMI	16.10 (1.27)	16.94 (1.14)	--	2.55	.01
Length of illness (years)	2.22 (1.25)	2.41 (1.87)	--	0.40	.69
Education (years)	9.28 (1.45)	9.35 (1.46)	9.18 (1.44)	0.09	.92
Race				9.00	.34
Caucasian	31 (86.11)	17 (85.00)	20 (71.43)		
Asian	2 (5.56)	2 (10.00)	3 (10.71)		
African American	0 (0.00)	0 (0.00)	2 (7.14)		
Other	3 (8.33)	1 (5.00)	3 (10.71)		
Ethnicity				3.25	.20
Hispanic	3 (8.33)	5 (25.00)	6 (21.43)		
Non-Hispanic	33 (91.67)	15 (75.00)	22 (78.57)		
EDE Global Score	2.93 (1.26)	3.10 (1.35)	0.08 (0.09)	68.61	<.001
EDE Eating Concerns	2.08 (1.07)	2.09 (1.40)	0.05 (0.27)	37.89	<.001
EDE Shape Concerns	4.03 (1.58)	4.21 (1.58)	0.16 (0.19)	83.16	<.001
EDE Weight Concerns	3.71. (1.58)	3.88 (1.60)	0.13 (0.20)	70.68	<.001
EDE Restraint	1.90 (1.54)	2.21 (1.54)	0.00 (0.00)	23.2	<.001
WASI Full Scale IQ	112.39 (10.96)	113.15 (13.2)	108.21 (7.72)	1.67	.20
TCI Harm Avoidance	25.22 (5.78)	22.13 (6.10)	13.91 (7.73)	23.45	<.001
BIS Punishment	25.33 (2.44)	23.83 (4.06)	19.78 (3.79)	22.06	<.001
BAS Reward	15.81 (2.62)	16.17 (2.66)	17.07 (1.80)	2.22	.12
BAS Drive	10.08 (2.59)	10.39 (2.91)	10.59 (2.10)	0.33	.72
BAS Fun	10.61 (2.72)	11.50 (2.33)	11.85 (1.8)	2.22	.12
SPSRQ Reward	10.64 (5.21)	12.18 (3.92)	9.26 (3.63)	2.25	.11
SPSRQ Punishment	16.81 (4.88)	15.71 (4.95)	9.48 (6.21)	15.26	<.001
ATQ Effortful Control	4.26 (0.82)	4.38 (0.51)	5.06 (0.75)	9.16	<.001
ATQ Activation Control	4.33 (1.03)	4.55 (0.65)	5.38 (0.81)	10.74	<.001
ATQ Attentional Control	3.68 (1.15)	3.91 (1.02)	4.98 (1.16)	10.29	<.001
ATQ Inhibitory Control	4.60 (0.87)	4.55 (0.73)	4.80 (0.81)	0.59	.56
BDI (n=83)	24.00 (13.49)	28.74 (13.58)	2.61 (3.46)	40.43	<.001
STAI-T (n=83)	56.81 (10.48)	57.79 (14.35)	29.50 (6.85)	64.40	<.001
STAI-S (n=83)	53.08 (12.53)	54.32 (13.15)	29.00 (7.25)	44.52	<.001

Raw Choice Data

Model-based and model-free learning were approximated using raw choice data by examining probabilities of first-stage stay behavior (Figure 1.1D – E). Specifically, increased probabilities of stay versus switch following second-stage wins served as a measure of model-free behavior. Increased probabilities of stay versus switch following wins during common transitions, and losses during rare transitions, served as a measure of model-based behavior. To visualize differences in these systems (Foerde et al., 2021), we computed difference scores based on prior work by (Eppinger et al., 2017). The model-based difference score was taken as: (common/win + rare/lose) – (common/lose + rare/win). The model-free difference score was taken as: (common/win+ rare/win) – (common/lose + rare/lose). These values are plotted as a function of group and condition (Supplemental Figure 1.1).

Results from the mixed-effects logistic regression analysis indicate that for reward, there was a significant main effect of outcome (B = 1.27, SE = 0.27, p = <.001) and an outcome-by-transition interaction (B = -1.16, SE = 0.32, p = <.001), but no significant group differences (Table 1.2). Punishment was similar, with a main effect of outcome (B = 1.15, SE = 0.23, p = <.001) and an outcome-by-transition interaction (B = -0.69, SE = 0.24, p = .003), but again no significant group differences (Table 1.2). An omnibus mixed-effects logistic regression analysis also including condition revealed similar results, with no group differences (Supplemental Table 1.1). Results therefore support mixed usage of model-free and model-based learning across groups, with computational model results below offering a more granular understanding of group differences in the relative balance of these systems.

Table 1.2: Raw choice data mixed effects logistic regression (by condition). AN-R – restricting type anorexia nervosa; AN-BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

Variable	Odds Ratio	Standard Error	Statistic	*p*
Reward				
Intercept [AN-R]	0.64	0.11	5.66	<.001
Last won	1.32	0.23	5.83	<.001
Last transition	0.45	0.13	3.55	<.001
AN-BP	0.07	0.19	0.36	.71
HC	-0.07	0.17	-0.43	.67
Last won x last transition	-1.00	0.28	-3.60	<.001
Last won x AN-BP	0.04	0.38	0.12	.91
Last won x HC	-0.06	0.34	-0.16	.87
Last transition x AN-BP	-0.11	0.21	-0.52	.60
Last transition x HC	0.10	0.19	0.51	.61
Last won x last transition x AN-BP	-0.41	0.46	-0.88	.38
Last won x last transition x HC	-0.14	0.42	-0.32	.75
Punishment				
Intercept [AN-R]	0.81	0.14	5.88	<.001
Last won	0.96	0.20	4.70	<.001
Last transition	0.22	0.12	1.80	.07
AN-BP	0.14	0.23	0.61	.54
HC	-0.23	0.21	-1.11	.27
Last won x last transition	-0.72	0.21	-3.49	<.001
Last won x AN-BP	0.33	0.34	0.96	.34
Last won x HC	0.18	0.30	0.60	.55
Last transition x AN-BP	-0.03	0.21	-0.15	.88
Last transition x HC	0.19	0.19	1.02	.31
Last won x last transition x AN-BP	-0.06	0.35	-0.17	.86
Last won x last transition x HC	0.02	0.31	0.07	.94

Computational Model

The reinforcement learning model consisted of a combination of model-based and model-free weights (β_{MB} and β_{MF}, respectively) that control the relative contribution of model-free and model-based updating on choice behavior. Differences between HC and AN subgroups were encoded in regression coefficients (γ_0 , γ_{ANR}, and γ_{ANBP}) for each freely estimated parameter, with HC being set as the reference group (γ_0): if these coefficients differed from zero, then diagnostic subgroups differed on these parameters from HC. To compare AN subgroups to each other, we reconstructed subgroup posteriors and directly compared these posterior distributions to each other: AN-R = $\gamma_0 + \gamma_{ANR}$ and AN-BP = $\gamma_0 + \gamma_{ANBP}$. Posterior distributions of transformed γ weights are plotted in Figure 1.2. Below, we report average γ values, 95% credibility intervals, and Bayesian p-values (see Method).

In the reward condition, AN-R had significantly higher model-free learning compared to HC (γ_{ANR} = 0.92, P = .03, 95%CI[-0.02, 1.82]) and lower model-based learning at a trend level (γ_{ANR} = -0.31, P = .05, 95%CI[-0.70, 0.07]). In the punishment condition, AN-R displayed significantly attenuated model-free learning (γ_{ANR} = -1.26, P = .002, 95%CI[-2.38, -0.38]), and model-based learning compared to HC (γ_{ANR} = -0.58, P = .004, 95%CI[-1.01, -0.16]). AN-BP also had attenuated model-free learning for punishment compared to HC at a trend level (γ_{ANBP} = -0.72, P = .07, 95%CI[-1.99, 0.21]). Overall, AN-R showed greater reliance on model-free learning for the reward condition, but deficits in both model-free and model-based learning for the punishment condition (Figure 1.2). Conversely, AN-BP only displayed trend-level deficits in model-free learning for punishment relative to HC. These results are displayed in Table 1.3. Estimated γ weights for the remaining freely estimated model parameters (α, β_2, p) are included in Supplemental Tables 1.2-1.4 and shown in Supplemental Figures 1.2-1.4.

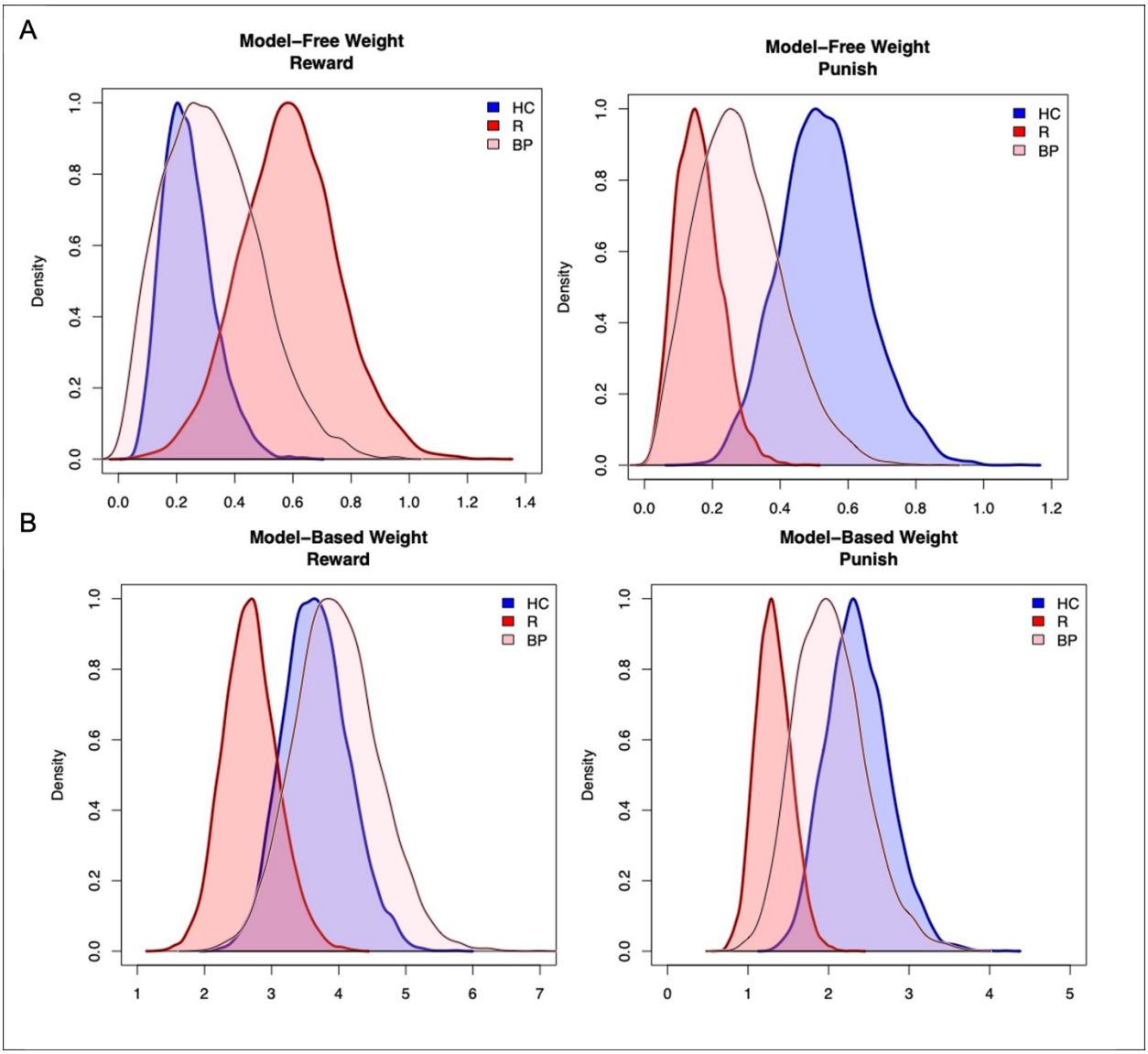


Figure 1.2: Group-level posterior distributions (A) of model-free weights (γ_0, γ_1, γ_2) for reward (left) and punishment (right), and (B) model-based weights (γ_0, γ_1, γ_2) for reward (left) and punishment (right). Values have been log transformed to display estimates on the correct scale. R – restricting type anorexia nervosa; BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

Table 1.3: Latent regression coefficients. AN-R – restricting type anorexia nervosa; AN-BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

Parameter	Mean Posterior	*p*	CI[2.5%]	CI[97.5%]
Reward – Model-free Weights				
γ_0 [HC]	-1.49	0	-2.30	-0.83
γ_1 [AN-R vs HC]	0.92	.03	-0.02	1.82
γ_2 [AN-BP vs HC]	0.20	.34	-1.24	1.30
Punishment – Model-free Weights				
γ_0 [HC]	-0.67	<.001	-1.23	-0.23
γ_1 [AN-R vs HC]	-1.26	.002	-2.38	-0.38
γ_2 [AN-BP vs HC]	-0.72	.07	-1.99	0.22
Reward – Model-based Weights				
γ_0 [HC]	1.29	0	1.02	1.54
γ_1 [AN-R vs HC]	-0.31	.05	-0.70	0.07
γ_2 [AN-BP vs HC]	0.08	.34	-0.32	0.47
Punishment – Model-based Weights				
γ_0 [HC]	0.85	<.001	0.52	1.15
γ_1 [AN-R vs HC]	-0.58	.004	-1.01	-0.16
γ_2 [AN-BP vs HC]	-0.17	.25	-0.68	0.34

Associations with Resting-State Functional Connectivity

We extracted rsfMRI functional connectivity coefficients from subject-level connectivity maps for OFC-NAcc and SMA-putamen pairs. There were no group differences in connectivity between these ROIs, controlling for z-scored age, BMI, and IQ. There were no significant effects between SMA-putamen connectivity and β_{MF}, including no main effect of connectivity, no connectivity-by-group interaction, and no connectivity-by-condition interaction. For OFC-left NAcc connectivity and β_{MB}, there was a significant main effect of connectivity for AN-R (B = 2.15, SE = 0.92, p = .02), but no significant main effects for any other group. There were also no

significant connectivity-by-group interactions, nor were there connectivity-by-condition interactions. Because reward was the reference level for the condition variable, a main effect of connectivity indicated that higher reward β_{MB} was associated with stronger resting-state connectivity between the OFC-left NAcc in AN-R only (Figure 1.3A – B). To explore this relationship within the AN-R group, we conducted a post-hoc linear regression analysis predicting reward β_{MB} and specifying an OFC-left NAcc connectivity and EDE Global Score interaction, controlling for the same covariates. The main effect of connectivity (B = 7.53, SE = 2.07, p = .001) and the interaction term (B = -1.92, SE = 0.70, p = .01) were significant, showing that the more positive the relationship between reward β_{MB} and connectivity, the lower the EDE Global Score (Supplemental Figure 1.5, Supplemental Table 1.5).

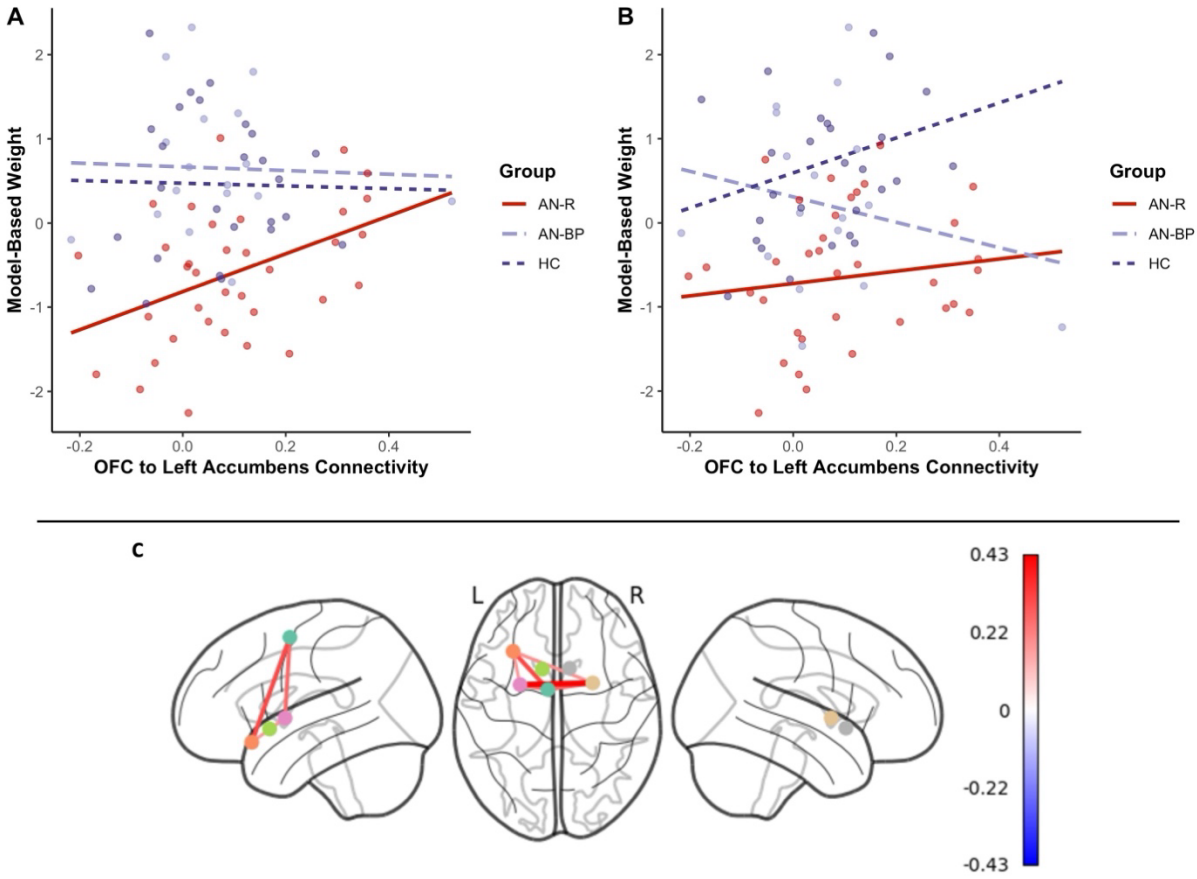


Figure 1.3: Resting-state functional connectivity results, including associations between model-based weights and OFC-to-left NA connectivity for reward (A) and punishment (B). Covariance matrix for ROI-to-ROI connectivity are visualized as a functional connectome, with edges representing the strength of the correlation between nodes (C). AN-R – restricting type anorexia nervosa; AN-BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

Associations with Clinical Variables

Consistent with prior work (Foerde et al., 2021), there were no significant associations between β_{MB} or β_{MF} and symptom measures (BMI, EDE, BDI, or STAI) after family-wise correction. We also did not find significant associations with TCI Harm Avoidance, SPSRQ Reward, SPSRQ Punishment, BIS, or BAS Reward Responsivity. However, when examining β_{MB} and ATQ Effortful Control, there was a significant Effortful Control-by-condition-by-group interaction, suggesting differences in the directionality of both the effect between AN-BP and

AN-R and the effect between conditions for AN-BP (Supplemental Table 1.6, Supplemental Figure 1.6).

DISCUSSION

This study investigated, amongst adolescents with subtypes of AN compared to HC, alterations in model-based and model-free learning under conditions of reward and punishment, as well as associated links to corticostriatal functional connectivity. Overall, the AN-R group showed greater reliance on model-free learning for reward, but decreased utilization of both model-free and model-based learning for punishment compared to HC. Moreover, in adolescents with AN-R only, higher model-based learning for reward was associated with stronger functional connectivity between regions implicated in goal-directedness, and this was associated with lower eating disorder severity. In contrast, adolescents with AN-BP displayed trend-level deficits in model-free learning for punishment, as well as a negative relationship between effortful control and model-based learning for reward, but no other significant behavioral or neural differences when compared to HC or AN-R. Together, results suggest substantive differences in the balance between model-free and model-based learning between AN-R and HC across conditions, as well as differentiation between AN subtypes primarily under reward. Such alterations in adolescent AN-R may reflect over-reliance on inflexible, habit-based reward learning and inadequate utilization of flexible, goal-directed learning.

Parameter estimates from the reinforcement learning model indicate that adolescents with AN-R displayed patterns of model-free and model-based learning similar to those observed in individuals with OCD (Voon, Baek, et al., 2015). There is accumulating evidence that disorders characterized by compulsivity (e.g., OCD, substance use) rely less on model-based learning during sequential reward-based decision-making tasks (V. M. Brown et al., 2020; Foerde et al.,

2021; Gillan et al., 2016; Janssen et al., 2020; Voon, Baek, et al., 2015; Voon, Derbyshire, et al., 2015), indicating that this bias may be a transdiagnostic feature of psychiatric conditions characterized by cognitive or behavioral perseveration. Of note, greater utilization of model-free learning in our AN-R sample may be aligned with etiological models purporting that dietary restriction is initially maintained by operant conditioning (e.g., reinforced by reward of weight loss and praise) and becomes overtrained (Godier et al., 2016; J. E. Steinglass & Walsh, 2016; J. Steinglass & Walsh, 2006), raising the question of whether a tendency towards reward-based model-free learning may reflect a premorbid susceptibility factor that predisposes individuals to early development of dietary habits. This hypothesis is aided by the positive relationship between model-based learning and OFC-to-left NAcc functional connectivity in AN-R only, potentially indicating that heightened neural synchrony in reward circuitry is connected to greater reward-based goal pursuit. Prior studies have reported disturbances in reward-related functional architecture in AN (Cha et al., 2016; Haynos et al., 2019), particularly in ventral and dorsal frontostriatal networks, that support etiological theories implicating alterations in reward functioning and over-reliance on habit (Kaye et al., 2013; J. Steinglass & Walsh, 2006; Walsh, 2013; Wierenga et al., 2014). Our post-hoc analysis further clarified that a stronger, positive association between OFC-to-left NAcc functional connectivity and model-based learning was linked to less severe eating pathology, providing potential evidence of a brain-based protective marker, though this was exploratory. If true, this would suggest that learning systems contributing to decision-making under reward and corresponding neural circuitry may be of particular clinical importance (Dang et al., 2020).

While behavioral patterns present for the AN-R group are relatively consistent with prior work in psychiatric samples (V. M. Brown et al., 2020; Foerde et al., 2021; Gillan et al., 2016;

Janssen et al., 2020; Voon, Baek, et al., 2015; Voon, Derbyshire, et al., 2015), the lack of difference between AN-BP and HC groups was unexpected. A prior study showed that individuals with AN displayed attenuated model-based learning compared to HC, with this effect being largely driven by individuals with AN-BP (Foerde et al., 2021). Developmental factors and/or illness stage may have contributed to our findings; indeed, prior research demonstrates that utilization of model-based learning can change over the course of development, with model-based systems coming online during adolescence and strengthening into adulthood (Bolenz & Eppinger, 2022; Decker et al., 2016; Potter et al., 2017; Scholz et al., 2023; Vaghi et al., 2020). Further, higher compulsivity at baseline in healthy adolescents is associated with diminished strengthening of model-based control at a later time point, suggesting that maturation of this system can be hijacked by compulsive psychological features (Vaghi et al., 2020). It is therefore possible that adolescent stages of AN-R reflect delayed maturation of these systems and adolescents with AN-BP may not show signs of deficiency at this stage, but may fail to improve as a function of age. Longitudinal studies are needed to resolve this issue.

Our results also suggest that adolescents with AN-R and AN-BP are behaviorally dissociable, particularly in their responses to reward learning. The AN-R group utilized greater model-free learning for reward, while the AN-BP group did not significantly differ from HC in either learning system. Exploratory analyses with clinical variables supported the distinction between subtypes as well, with greater reward-related model-based learning in the AN-BP group linked to lower levels of adaptive effortful control (i.e., self-regulation), while AN-R displayed a weaker, positive relationship. One possible explanation for this distinction may be differences in saliency of reward. It is postulated that AN-BP is characterized by greater sensitivity to reward compared to AN-R, with some studies corroborating this hypothesis (Beck et al., 2009; Claes et

al., 2006; A. Harrison et al., 2010), and others contradicting it (Glashouwer et al., 2014; Jappe et al., 2011). In the current study, self-reported sensitivity to reward is intact in the AN sample and is not differentiated by subtype, but utilization of model-free/model-based learning to make decisions in a rewarding context is. Model-based control has high computational expense, making it an effortful choice. If individuals with AN-BP are indeed more sensitive to reward, it is possible that positive outcomes (e.g., monetary gain, palatable food consumption) have higher motivational and incentive salience, leading to use of a cognitively demanding strategy even in the context of lower trait self-regulatory control. This is consistent with findings that adults with binge eating symptoms are more willing to work for reward as demonstrated by a higher breakpoint during a progressive ratio task (Keel et al., 2022), and prior research demonstrating that model-based learning can be modulated by stress and cognitive control (Otto et al., 2015; Otto, Raio, et al., 2013).

It is worth noting that the results we observed in the punishment condition were not consistent with our hypotheses. Despite greater self-reported sensitivity to punishment (i.e., subjective responsivity and/or avoidance of aversive consequences), the AN-R group displayed reductions in both model-based and model-free learning compared to HC, raising a question of how this group ultimately made choices if they were relatively deficient in both systems during punishment. Notably, this is in contrast to prior work suggesting increased model-based learning for punishment, which has been observed in individuals with OCD (Voon, Baek, et al., 2015), serotonin-depletion (Worbe et al., 2016), and in healthy adolescents and adults (Bolenz & Eppinger, 2022). We previously reported lower learning rates for both reward and punishment in a predominantly young adult sample of women with AN-R; however, worse learning on punishment trials was associated with poorer outcome, suggesting that these individuals may

have particular difficulty modifying expectations when punishment is possible. Whether this relates to active avoidance of aversive outcomes or cognitive inflexibility is difficult to determine, although in the current study, the punishment inverse temperature parameter (β_2) was significantly lower in adolescents with AN-R compared to HC, indicating more exploration of choices and less exploitation of choice valuation during the second stage (Supplemental Figure 1.3). Alternatively, the AN-R group may have utilized another learning system or strategy (e.g., successor representation (Dayan, 1993; Gershman, 2018) for the punishment condition, although more research is needed to explore this hypothesis.

This study was the first to investigate model-based/model-free learning in adolescent samples of AN-R, AN-BP, and HC. Strengths of this work include a well-characterized adolescent AN sample, use of a well-validated two-step decision-making task with monetary instead of disorder-specific (e.g., food) outcomes to avoid symptom provocation, modulation of outcome valence, use of a well-validated computational model to elicit latent learning features, and the inclusion of rsfMRI to examine relationships between cognitive and neural features of learning. Results add to growing literature on the arbitration between these learning systems in psychopathology and provide greater insight into not only earlier stages of development and disease progression, but also valence-specific effects that may be of particular importance in development and maintenance of AN.

Limitations

Despite these strengths, limitations of this study highlight directions for future work. Our sample included fewer participants with AN-BP compared to AN-R and HC, and both weight-restored and underweight adolescents with AN were included. Additionally, neural data only consisted of rsfMRI scans, with a limited ability to examine task-specific neural functioning that

would correspond to these constructs. Specifically, we found no relationship between learning and connectivity in AN-BP and HC, possibly indicating that these learning systems might not be encoded in the connectivity signatures investigated here but may be more distributed, particularly as our sample is adolescent. Future work should replicate AN-BP findings, examine impact of illness state, and utilize task-based neuroimaging to reveal corresponding neural functioning. Lastly, we recognize that the dual-system framework may have limitations. (Collins & Cockburn, 2020) have argued that the forced dichotomy of the model-based/model-free framework artificially limit the dimensionality of learning processes. Moreover, model-based and model-free systems may not be as dissociable as this framework suggests. Consequently, future work, as suggested by (Collins & Cockburn, 2020), should explore different dimensions of learning systems, as well as the constructs that may influence components of those systems (e.g., uncertainty).

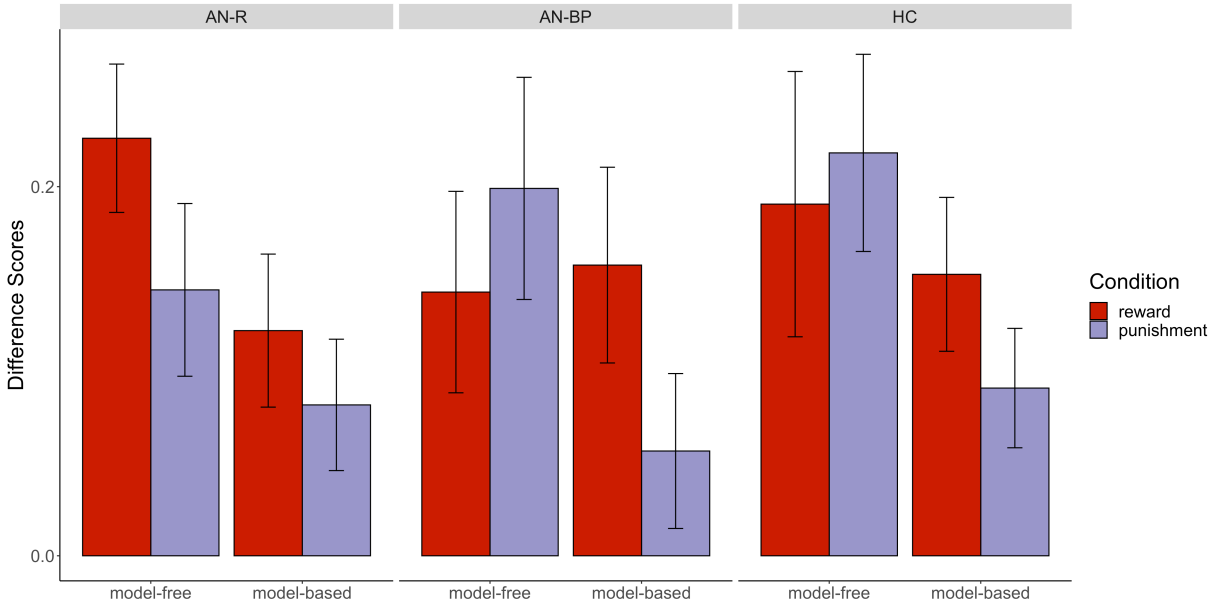
Conclusions and clinical implications

This is the first study to evaluate the contribution of model-based and model-free learning for reward and punishment in adolescents with AN-R and AN-BP. Using a well-validated two-step decision-making task and rsfMRI, we observed important distinctions primarily between the AN-R and HC groups, suggesting increased reliance on model-free learning for reward (e.g., habit) may contribute to the ability to consistently restrict food in individuals with AN-R. Exploratory results also suggest that a stronger relationship between model-based learning for reward and neural connectivity in reward circuitry may serve as a protective factor for AN-R. Finally, there is preliminary evidence of distinctions between AN subtypes that may result in divergence of mechanisms contributing to symptom expression. Consequently, interventions that

reduce model-free learning and/or improve model-based learning may facilitate recovery in adolescence, though this may be subtype-dependent.

ACKNOWLEDGEMENTS

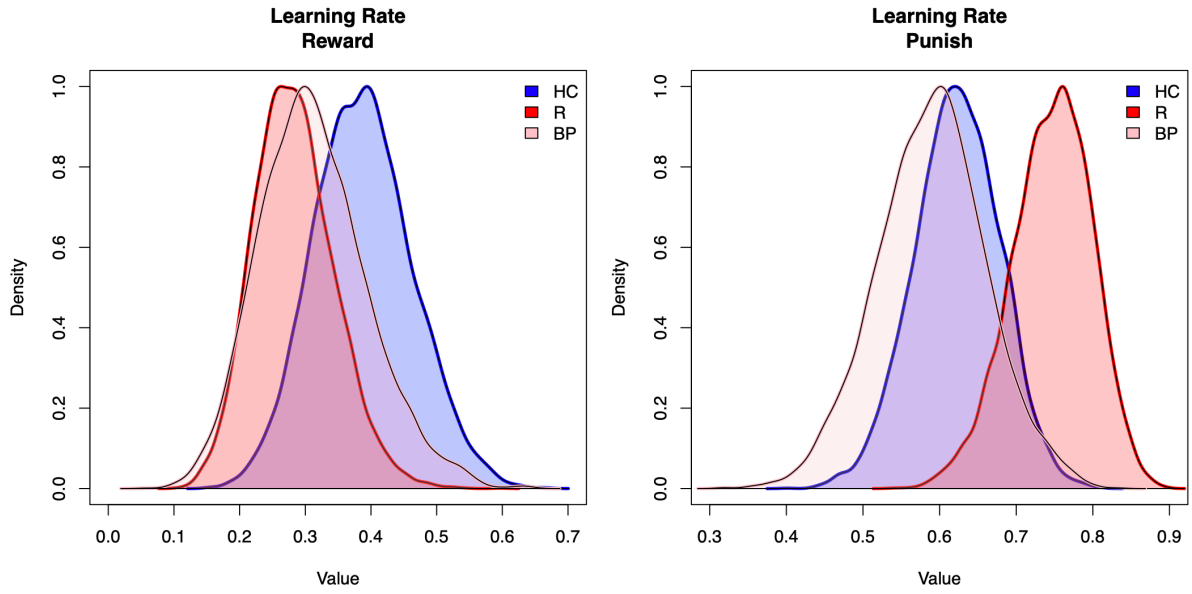
Chapter 1 contains unpublished material. I would like to acknowledge and thank: Sean Devine, PhD, Ross Otto, PhD, Amanda Bischoff-Grethe, PhD, and Christina Wierenga, PhD. I would also like to thank Megan Martinho, Angeline Krueger, the UCSD Eating Disorders Center for Treatment and Research staff and patients for their contributions to this project. This project was supported by National Institutes of Health Grant No. R21MH118409 (awarded to CEW).



Supplemental Figure 1.1: Index of the model-based and model-free effects by condition and group based on first-stage stay probabilities. The model-based difference score was taken as: (common/win + rare/lose) – (common/lose + rare/win). The model-free difference score was taken as: (common/win+ rare/win) – (common/lose + rare/lose). Error bars represent ±1 SEM. AN-R – restricting type anorexia nervosa; AN-BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

Supplemental Table 1.1: Raw choice data mixed effects logistic regression – four-way interaction. AN-R – restricting type anorexia nervosa; AN-BP – binge-eating/purging type anorexia nervosa; HC – healthy controls.

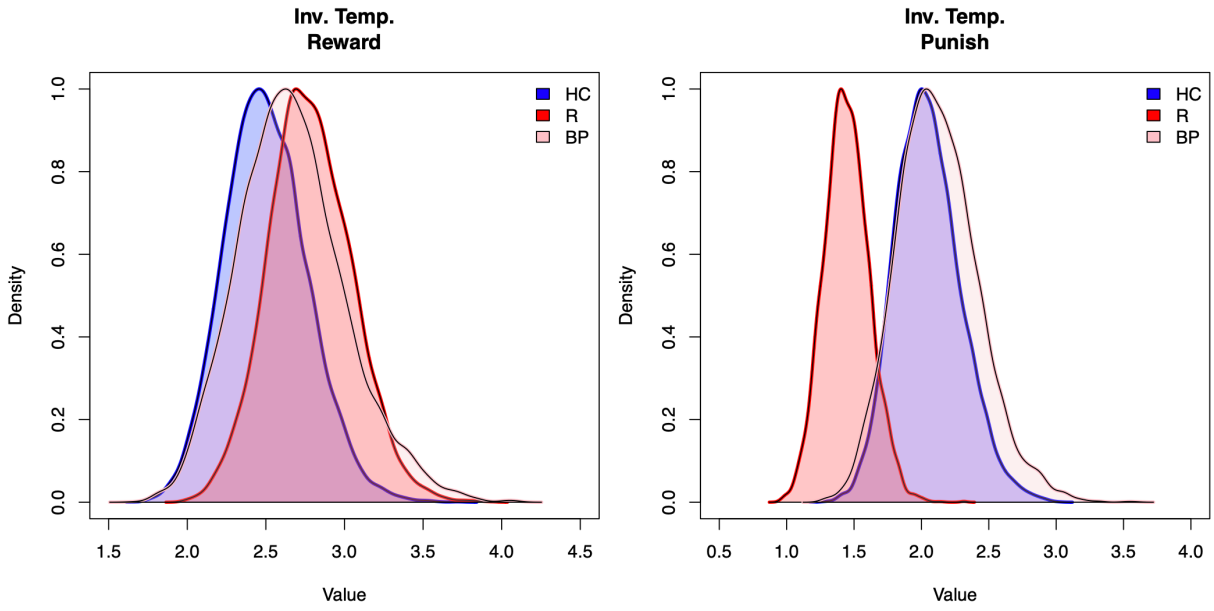
Variable	Odds Ratio	Standard Error	Statistic	*p*
Intercept [AN-R]	0.81	0.14	5.83	**<.001**
Last won	0.95	0.2	4.83	**<.001**
Last transition	0.22	0.13	1.75	.08
Condition	-0.16	0.1	-1.67	.10
AN-BP	0.15	0.23	0.65	.52
HC	-0.21	0.21	-1.03	.30
Last won x last transition	-0.71	0.2	-3.46	**<.001**
Last won x condition	0.35	0.17	2.03	**.04**
Last transition x condition	0.21	0.13	1.62	.10
Last won x AN-BP	0.33	0.33	0.98	.33
Last won x HC	0.16	0.3	0.54	.59
Last transition x AN-BP	-0.04	0.21	-0.21	.83
Last transition x HC	0.18	0.19	0.94	.35
Condition x AN-BP	-0.09	0.16	-0.56	.57
Condition x HC	0.12	0.15	0.81	.42
Last won x last transition x condition	-0.28	0.26	-1.08	.28
Last won x last transition x AN-BP	-0.05	0.35	-0.16	.88
Last won x last transition x HC	0.02	0.31	0.08	.94
Last won x condition x AN-BP	-0.23	0.29	-0.8	.42
Last won x condition x HC	-0.23	0.26	-0.9	.37
Last transition x condition x AN-BP	-0.04	0.22	-0.17	.87
Last transition x condition x HC	-0.06	0.2	-0.29	.78
Last won x last transition x condition x AN-BP	-0.4	0.43	-0.93	.35
Last won x last transition x condition x HC	-0.18	0.39	-0.46	.64



Supplemental Figure 1.2: Group-level posterior distributions of learning rates (γ_0, γ_1, γ_2) for reward (left) and punishment (right). Values have been transformed using the inverse cumulative distribution function to display estimates on the correct scale. R – restricting type anorexia nervosa, BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

Supplemental Table 1.2: Latent regression coefficients on learning rate. AN-R – restricting type anorexia nervosa, AN-BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

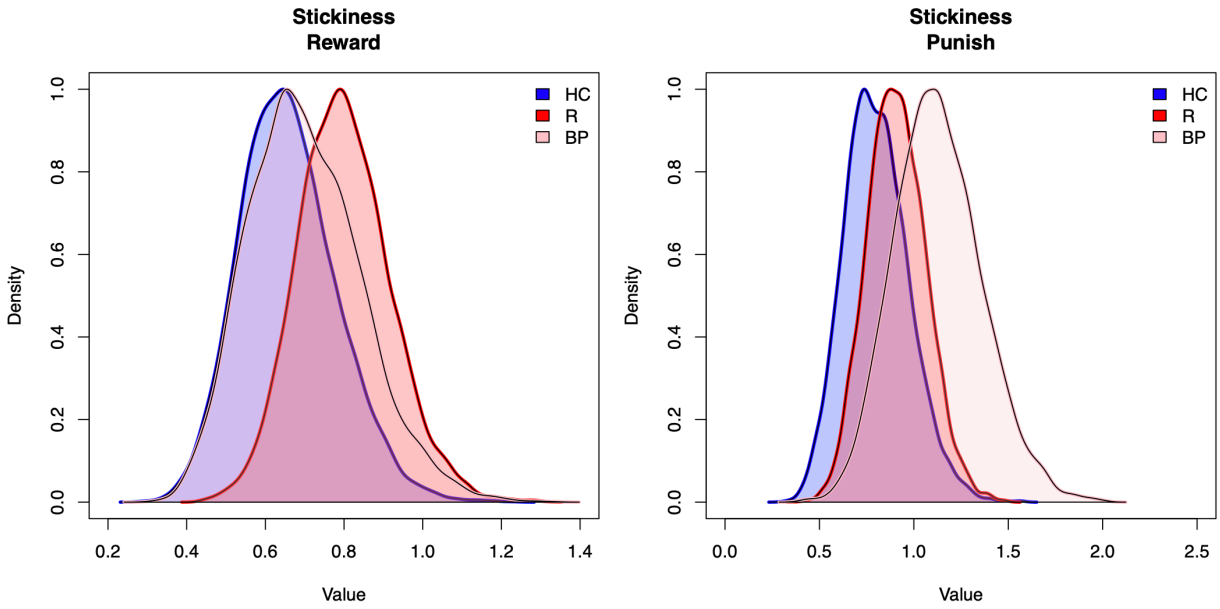
Parameter	**Mean Posterior**	***p***	**CI[2.5%]**	**CI[97.5%]**
Reward				
γ_0 [HC]	-0.29	.07	-0.68	-0.42
γ_1 [AN-R vs HC]	-0.29	.14	-0.81	0.23
γ_2 [AN-BP vs HC]	-0.21	.24	-0.81	0.37
Punishment				
γ_0 [HC]	0.32	.02	0.03	0.61
γ_1 [AN-R vs HC]	0.35	.05	-0.08	0.79
γ_2 [AN-BP vs HC]	-0.09	.36	-0.55	0.39



Supplemental Figure 1.3: Group-level posterior distributions of second-stage inverse temperature parameters ($\gamma_0, \gamma_1, \gamma_2$) for reward (left) and punishment (right). Values have been log transformed to display estimates on the correct scale. R – restricting type anorexia nervosa, BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

Supplemental Table 1.3: Latent regression coefficients on inverse temperature. AN-R – restricting type anorexia nervosa, AN-BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

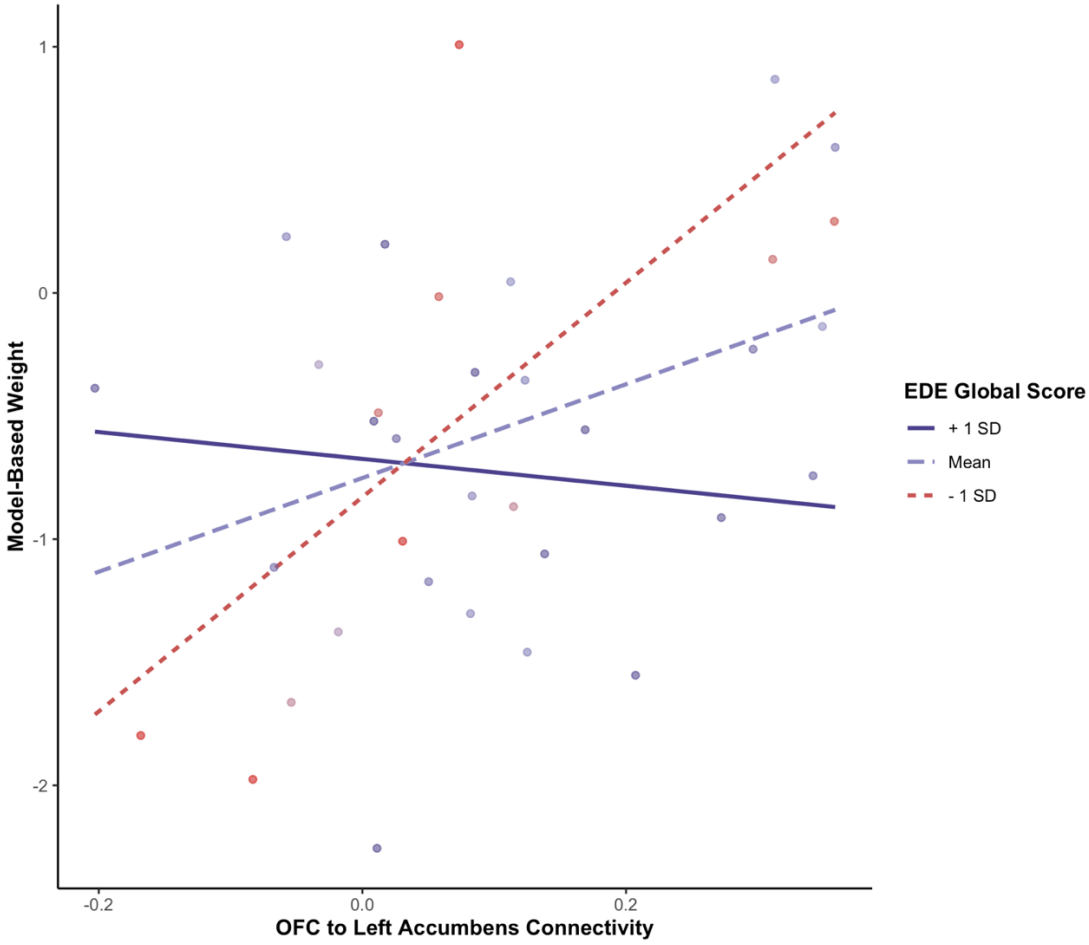
Parameter	**Mean Posterior**	***p***	**CI[2.5%]**	**CI[97.5%]**
Reward				
γ_0 [HC]	0.91	0	0.71	1.11
γ_1 [AN-R vs HC]	0.11	.22	0.17	0.38
γ_2 [AN-BP vs HC]	0.06	.36	-0.26	0.38
Punishment				
γ_0 [HC]	0.71	0	0.48	0.94
γ_1 [AN-R vs HC]	-0.35	.01	-0.67	-0.04
γ_2 [AN-BP vs HC]	0.03	.43	-0.33	0.37



Supplemental Figure 1.4: Group-level posterior distributions of stickiness parameters (γ_0, γ_1, γ_2) for reward (left) and punishment (right). Values have been log transformed using the inverse cumulative distribution function to display estimates on the correct scale. R – restricting type anorexia nervosa, BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

Supplemental Table 1.4: Latent regression coefficients on stickiness parameter. AN-R – restricting type anorexia nervosa, AN-BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

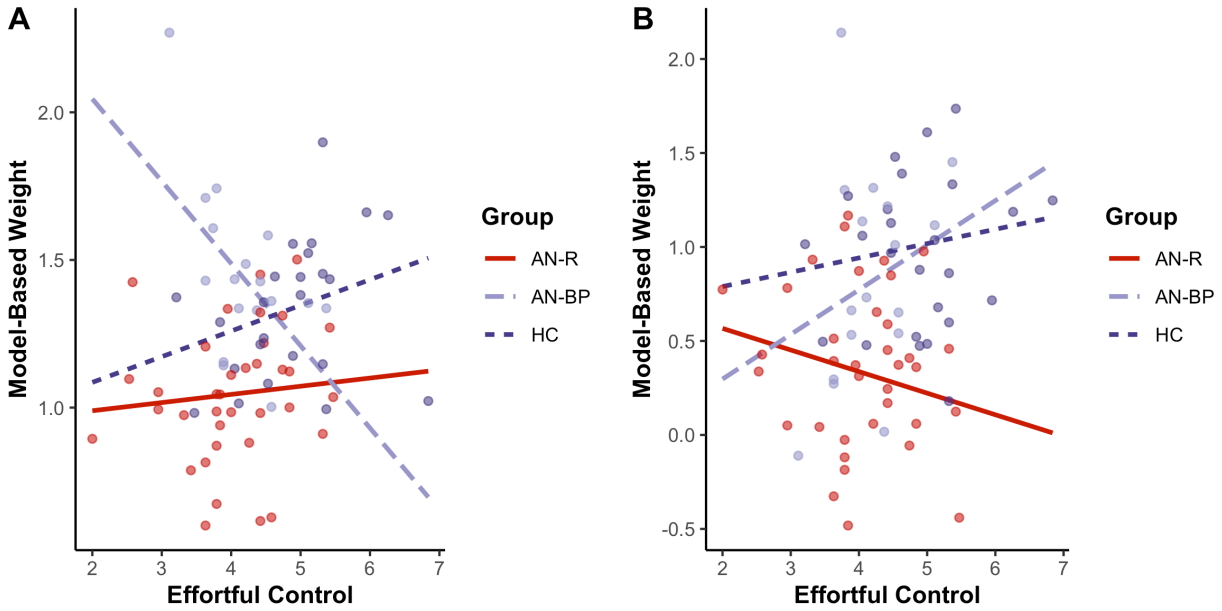
Parameter	**Mean Posterior**	***p***	**CI[2.5%]**	**CI[97.5%]**
Reward				
γ_0 [HC]	-1.13	0	-1.35	-0.90
γ_1 [AN-R vs HC]	0.13	.20	-0.16	0.42
γ_2 [AN-BP vs HC]	0.04	.42	-0.29	0.37
Punishment				
γ_0 [HC]	-1.00	<.001	-1.28	-0.73
γ_1 [AN-R vs HC]	0.09	.32	-0.28	0.44
γ_2 [AN-BP vs HC]	0.24	.13	-0.15	0.63



Supplemental Figure 1.5: Visualization of the interaction between EDE Global Score and OFC-to-left NAcc connectivity on reward-related model-based learning in AN-R only.

Supplemental Table 1.5: Reward model-based weights and OFC-to-left NAcc-by-EDE total score linear regression in AN-R. EDE – Eating Disorder Examination; WASI – Weschler Abbreviated Scale of Intelligence.

Variable	Estimate	Standard Error	Statistic	*p*
Intercept	-0.99	0.33	-2.99	**.006**
OFC-NAcc connectivity	7.53	2.07	3.64	**.001**
EDE Total Score	0.06	0.11	0.56	.579
Age	-0.07	0.14	-0.51	.616
BMI	-0.24	0.19	-1.23	.230
WASI Full Scale IQ	0.19	0.12	1.53	.138
OFC-NAcc connectivity x EDE Total Score	-1.92	0.70	-2.74	**.011**



Supplemental Figure 1.6: Association between model-based weights and ATQ Effortful Control scale for reward (A) and punishment (B). AN-R – restricting type anorexia nervosa, AN-BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

Supplemental Table 1.6: Model-based weights and effortful control mixed effects regression. AN-R – restricting type anorexia nervosa, AN-BP – binge-eating/purging type anorexia nervosa, HC – healthy control.

Variable	Estimate	Standard Error	Statistic	$p_{corrected}$
Intercept [AN-R]	-1.10	0.71	-1.54	1.00
AN-BP	5.41	1.68	3.22	.017
HC	-0.25	1.23	-0.21	1.00
Condition [punishment]	1.36	0.92	1.48	1.00
ATQ Effortful Control	0.11	0.17	0.62	1.00
Age	-0.02	0.07	-0.27	1.00
BMI	0.20	0.07	2.82	.070
WASI Full Scale IQ	-0.03	0.11	-0.29	1.00
AN-BP x Condition [punishment]	-7.94	2.19	-3.63	.006
HC x Condition [punishment]	0.03	1.60	0.02	1.00
AN-BP x ATQ Effortful Control	-0.98	0.40	-2.45	.169
HC x ATQ Effortful Control	0.26	0.27	0.98	1.00
Condition [punishment] x ATQ Effortful Control	-0.32	0.22	-1.42	1.00
AN-BP x Condition [punishment] x ATQ Effortful Control	1.80	0.52	3.45	.010
HC x Condition [punishment] x ATQ Effortful Control	0.08	0.35	0.23	1.00

Chapter 2 VALENCED LEARNING IN RESTRICTING AND BINGE-EATING/PURGING EATING DISORDER PHENOTYPES

ABSTRACT

Learning from both positive and negative feedback is crucial for adaptive behavior, but individuals may show valence-specific sensitivities to feedback that asymmetrically influence speed of learning (i.e., learning biases). Overweighting learning from negative feedback compared to positive, for example, has been linked to anxiety disorders and may represent an anxiogenic mechanism that perpetuates anxiety symptoms. Eating disorders show high levels of trait harm avoidance and sensitivity to punishment and may therefore be characterized by learning asymmetries that overweight negative feedback. However, sensitivity to reward has been hypothesized to differentiate primarily restrictive from primarily binge-eating/purging eating disorder phenotypes and may suggest distinct learning biases differentiate these clinical presentations. The current study tested whether adults diagnosed with a predominantly restricting eating disorder, specifically anorexia nervosa restricting subtype, (R; $n_R = 29$) and adults diagnosed with a predominantly binge-eating/purging eating disorder, specifically anorexia nervosa binge-eating/purging subtype and bulimia nervosa (BP; $n_{BP} = 25$) showed greater positive or negative learning bias compared to a sample of adult undergraduates (UG; $n_{UG} = 40$). Participants completed a probabilistic gambling task and a battery of self-report questionnaires. Behavior was analyzed with model-based, computational approaches, including a series of reinforcement learning models. Results showed significant group differences in learning bias (operationalized as the difference between positive and negative learning rates estimated by the winning computational model), where the R group had a higher, more positive learning bias and the BP group had lower learning bias compared to the UG group. Learning bias for R and BP groups was also negatively correlated with binge eating severity, suggesting that asymmetrical

learning that does not favor learning from positive outcomes may be associated with binge eating. The distinctions in learning processes observed between restricting versus binge-eating/purging phenotypes suggests transdiagnostic mechanisms that may be differentially associated with behavioral symptoms and could inform future treatment research.

INTRODUCTION

Adaptation of behavior following rewarding and punishing outcomes is considered core to efficient learning, however prior research indicates that gain and loss are represented as distinct categorical events (Kubanek et al., 2015; Yechiam & Hochman, 2013), associated with disparate neural activation (Wächter et al., 2009), and conferring independent influences on approach/avoidance behavior (Guitart-Masip et al., 2014; Huys et al., 2011). Reward and punishment learning are not binary/orthogonal constructs, but may reflect biases that overweight the importance of one type of learning over the other. This has been labeled valence-dependent learning asymmetries, which have been reported consistently in human instrumental learning (Cazé & Van Der Meer, 2013; Eldar et al., 2016; Gershman, 2015; Michely et al., 2022) and have been shown to change as a function of age (Rosenbaum et al., 2022), value context dependence (Palminteri et al., 2015), and perceived risk (Niv et al., 2012).

Although it is posited that humans will inherently value punishment-related information to a greater degree and will display what is known as a negativity bias (Baumeister et al., 2001), or in utility models, loss aversion, most studies have shown the opposite, with otherwise healthy participants displaying an optimism bias that discounts or fails to incorporate worse-than-expected outcomes during learning (Lefebvre et al., 2017; Palminteri & Lebreton, 2022; Villano et al., 2023). Consequently, researchers have theorized that learning asymmetries favoring punishment learning will not be present for incentive structures with similar levels of gain and

loss, but will still be marked by mobilization of attentional resources following loss outcomes (Yechiam & Hochman, 2013). Others posit an adaptive explanation for this optimism bias, suggesting that those who more readily learn from rewarding outcomes are better able to favorably exploit their valuation of choices (Cazé & Van Der Meer, 2013; Lefebvre et al., 2017).

However, some individuals engage in a "defensively pessimistic" learning style that prioritizes punishment, manifesting as larger updates to choice valuation, or beliefs, following a negative outcome (Norem & Cantor, 1986). This phenomenon has been conceptualized as a strategy to minimize negative PEs, seemingly related to theories of anxiety suggesting that worry can be utilized as a method for maintaining negative expectancies in an effort to minimize unexpected negative shifts or contrasts (Newman & Llera, 2011). Indeed, this learning bias towards punishment has been observed in participants with, or at risk of, elevated anxiety and depression (Bishop & Gagne, 2018; Villano et al., 2023), although with some contradictory findings in healthy samples (Aberg et al., 2016). Individuals with high anxiety also demonstrate faster valence-agnostic learning rates compared to those with low anxiety, as well as difficulties adjusting their learning rates to match changing environmental volatilities, suggesting deficits in optimizing the speed of their belief updating (Browning et al., 2015; Huang et al., 2017). These patterns of behavior are believed to support the notion that aversion-related decision-making characterizes some psychiatric conditions, including anxiety disorders, and may serve as a method for engaging in either active or passive forms of avoidance (Paulus, 2020).

More recently, applications of anxiety frameworks to conceptualizations of eating disorder symptoms have begun to emerge (Christian & Levinson, 2022; Schaumberg et al., 2021), all citing the substantial overlap in phenomenology and the high comorbidity of eating disorders and anxiety disorders (Aspen et al., 2014), particularly in anorexia nervosa (AN;

Schaumberg et al., 2021) and bulimia nervosa (BN; Kaye et al., 2004). AN and BN are characterized by fear (e.g., fear of weight gain), negative affect, worry, avoidance, and compulsory actions (e.g., ritualized eating, bingeing, purging, exercise), which are shared with anxiety and obsessive-compulsive disorders. Prior work has shown that anxiety-related repetitive thinking and worry predict future disordered eating behaviors (Sala et al., 2019; Sala & Levinson, 2016) and a recent treatment study has highlighted the reciprocity between anxiety and eating disorder symptoms longitudinally (Reilly et al., 2025). AN also shares certain temperamental and personality traits with anxiety disorders, including high harm avoidance, intolerance of uncertainty, and perfectionism (Atiye et al., 2015; Bardone-Cone et al., 2007; M. Brown et al., 2017; Klump et al., 2004). Given these similarities, it is possible that some neurocognitive features in eating disorders may also be mutual, including learning biases.

Importantly, punishment and loss sensitivity have been shown to be elevated in AN and BN, even in adolescent samples, providing evidence of a possible risk and/or maintaining factor (A. Harrison et al., 2010; Haynos et al., 2020), although this has not been entirely consistent in the literature (Verharen et al., 2019). It is therefore possible that asymmetries in learning favoring punishment (i.e., negative learning bias) may be a core component of some eating disorders that maintain fear and anxiety associated with food, weight, and body stimuli. While punishment and loss sensitivity have been proposed as possible transdiagnostic features of eating disorders, reward sensitivity has been shown to differ between primarily restrictive and primarily binge-eating/purging eating disorder phenotypes (Chan et al., 2014; A. Harrison et al., 2010; P. Monteleone et al., 2014). This difference has primarily manifested as greater sensitivity to reward in individuals with binge-eating/purging eating disorder diagnoses; for example, Hagan & Forbush (2021) reported greater reward learning in a sample of individuals with BN during a

probabilistic reward learning task compared to healthy controls. However, Harrison et al. (2010) reported elevated scores on harm avoidance and punishment sensitivity measures in both individuals with BN and individuals with anorexia nervosa binge-eating/purging subtype (AN-BP) compared to healthy controls.

Although studies applying computational modeling approaches to eating disorder research is relatively nascent, some have already reported disruptions in latent learning and decision-making processes that may contribute to the presence and severity of these symptoms (C. S. Brown et al., 2024; Radzikowska et al., 2025). For example, Wierenga et al. (2022) reported that participants with AN displayed attenuated learning rates for both better-than-expected (i.e., positive PEs) and worse-than-expected (i.e., negative PEs) outcomes, while Bernardoni et al. (2018) reported higher learning rates for punishment outcomes and Verharen et al. (2019) reported greater insensitivity to losses. Studies in BN that incorporate computational modeling approaches are fewer, but demonstrate alterations in the ability to flexibly update beliefs (Berner, Fiore, et al., 2023).

To our knowledge, no study has specifically examined learning asymmetries in restrictive and binge-eating/purging eating disorder phenotypes. Therefore, this study aimed to determine whether negative learning rates are higher than positive learning rates in an eating disorder sample, including AN and BN, which may provide initial evidence for valence-dependent weighting of feedback in this population. Given that prior work suggests dissociable characteristics between restrictive and binge-eating/purging eating disorder phenotypes, we compared adults with anorexia nervosa restricting subtype (R) and a combined sample of adults with anorexia nervosa binge-eating/purging subtype and bulimia nervosa (BP) with an adult undergraduate comparator group (UG) to test differences in learning bias by clinical presentation

during a probabilistic gambling task (Eldar et al., 2016; Michely et al., 2022). We hypothesized that the R group would show attenuated learning bias compared to UG, favoring negative learning rates, while BP would show augmented learning bias compared to UG, favoring positive learning rates. Alternatively, given that binge eating and purging are associated with negative urgency, or the tendency to act impulsively in response to negative affect, we considered the possibility that more negative learning bias would be associated with higher severity of binge eating and/or purging. Specifically, as positive learning rates become less favored or underweighted, we hypothesized that severity of symptoms would increase. Finally, we conducted an exploratory canonical correlation analysis to examine associations between eating disorder symptom scores and other computational model parameters.

METHODS

Participants

A sample of acutely ill and partially remitted adults diagnosed with AN and BN were recruited from the UCSD Eating Disorder Center for Treatment and Research Partial Hospitalization Program ($n_{AN} = 43$, $n_{BN} = 11$). Participants completed eating disorder and other self-report measures within 14 days of treatment admission. Diagnoses were assigned using a semi-structured interview based on DSM-5 criteria (Structured Clinical Interview for DSM-5; First et al., 2015). Inclusion criteria included: 1) at least 18 years of age, 2) assigned AN or BN diagnosis. Exclusion criteria include: 1) diagnosis of an autism spectrum disorder, learning disorder, bipolar disorder, or psychosis, 2) over 35 years of age. To address our question about differences in eating disorder phenotype, we created a restricting group composed of only individuals diagnosed with AN restricting subtype (R; $n_R = 29$), and a binge-eating/purging group composed of both AN binge-eating/purging subtype and BN (BP; $n_{BP} = 25$).

A comparator group of adult undergraduate participants (n = 40) were recruited at San Diego State University through an online recruitment platform (SONA). Participants completed self-report questionnaires within 7 days of completing the behavioral task, but were not assessed with a semi-structured clinical interview. Therefore, clinical diagnostic information could not be obtained in this group. Inclusion criteria included: 1) at least 18 years of age. Exclusion criteria were minimal in order to capture a representative undergraduate sample; these included: 1) over 35 years of age, 2) diagnosis of a learning disorder.

The study was approved by the University of California, San Diego Institutional Review Board and the San Diego State University Institutional Review Board, and all participants provided written informed consent.

Measures

*Anthropometrics***.** BMI was calculated (kg/m^2) for individuals diagnosed with AN or BN using height and weight information. Height and weight information is available for all participants except undergraduate participants collected at San Diego State University.

Eating disorder severity. The Eating Pathology Symptoms Inventory (EPSI; Forbush et al., 2013) is a 45-item questionnaire used to assess eating disorder symptoms and is made up of eight subscales: Body Dissatisfaction, Binge Eating, Cognitive Restraint, Purging, Restricting, Excessive Exercise, Negative Attitudes toward Obesity, and Muscle Building. Items were rated on a 0–4 Likert scale, with higher values indicating higher frequency of eating disorder symptoms. The EPSI has demonstrated good construct and external validity within heterogeneous samples of clinical and non-clinical subjects (Coniglio et al., 2018).

Trait anxiety. The State-Trait Anxiety Inventory – Trait subscale (STAI-T; Spielberger et al., 1983) is a 20-item self-report measure that assesses trait anxiety on a 1-4 Likert scale, with higher values indicating greater severity of symptoms. The STAI-T has demonstrated excellent reliability and validity (Spielberger & Vagg, 1984).

Behavioral Inhibition/Activation. The Behavioral Inhibition/Behavioral Activation System Scales (BIS/BAS; Carver & White, 1994) is a 24-item measure that assesses the temperamental tendency to react to punishment (behavioral inhibition) and to react to reward (behavioral activation). The BIS/BAS factor structure has been reproduced in a large community sample and shows good validity and reliability (Jorm et al., 1998).

Harm Avoidance. The Temperament and Character Inventory (TCI) – Harm Avoidance subscale is a 35-item measure that assesses fear of uncertainty, shyness around others, fatiguability, and anticipatory worry about future harmful events (Cloninger et al., 1994).

Procedure

For the R and BP groups, participants first completed the self-report questionnaires and the semi-structured clinical interview. During a separate study visit, participants completed a computerized gambling task, described below. For the UG group, the procedures were identical, except that no semi-structured clinical interview was conducted. Additionally, for participants in the UG group, all study procedures were completed on the same day.

Gambling task. To examine differences in learning from success and failure, participants completed a modified version of a gambling card game (Eldar et al., 2016; Michely et al., 2022),

in which they were tasked with maximizing monetary wins and minimizing monetary losses (Figure 1). The task was coded in MATLAB and presented using PsychToolbox.

The game consisted of three 60-trial blocks (180 trials total). On each trial, participants encountered one of three decks distinguished by unique colors and patterns. Following a brief interval lasting 2 to 5 seconds (uniformly distributed), the computer randomly generated a number ranging from 1 to 9. Participants then had ~2.5 seconds to decide whether they wished to take a gamble that the number they draw will be higher than the computer's number. Opting for the gamble resulted in a $5 win if the participant's drawn number was higher, and a $5 loss if it was lower (as well as in half of the trials where the numbers were equal). However, participants also retained the choice to abstain from the gamble, leading to neither a win nor a loss, signifying a "no-risk" decision. In such instances, the outcome remained undisclosed to the participants. Feedback was presented 700 milliseconds after each decision, indicated by a visual symbol of '+$5' for a win, '−$5' for a loss, or '=' for no win or loss. The drawn number itself was not revealed during feedback.

Participants were informed that the three decks they drew from consisted of varying proportions of high and low numbers. However, they remained unaware that one deck featured a consistent distribution of numbers between 1 and 9 (referred to as the 'even deck' or 'moderate risk deck'), another deck had an increased frequency of 1's compared to other numbers (referred to as the 'low deck' or 'high risk deck'; 30% reduction in wins), and a third deck had an elevated frequency of 9's compared to other numbers (referred to as the 'high deck' or 'low risk deck'; 30% increase in wins). During the initial 15 trials, the computer generated the numbers 4, 5, and 6 three times each, while the remaining numbers appeared only once. To maintain an approximately 50% gamble participation rate for all participants, the frequencies of the numbers

drawn three times each were adjusted in subsequent sets of 15 trials. Specifically, if participants chose to gamble against these numbers ≥66.7% of the time in the preceding 15 trials, the numbers were increased by one (e.g., from [4, 5, 6] to [5, 6, 7]). Conversely, if participants engaged in ≤33.3% gambles, the numbers were decreased by one. The order of the decks were pseudorandomized, ensuring that the three decks were consistently matched against comparable computer-generated numbers and to prevent any deck from appearing consecutively more frequently than the other decks.

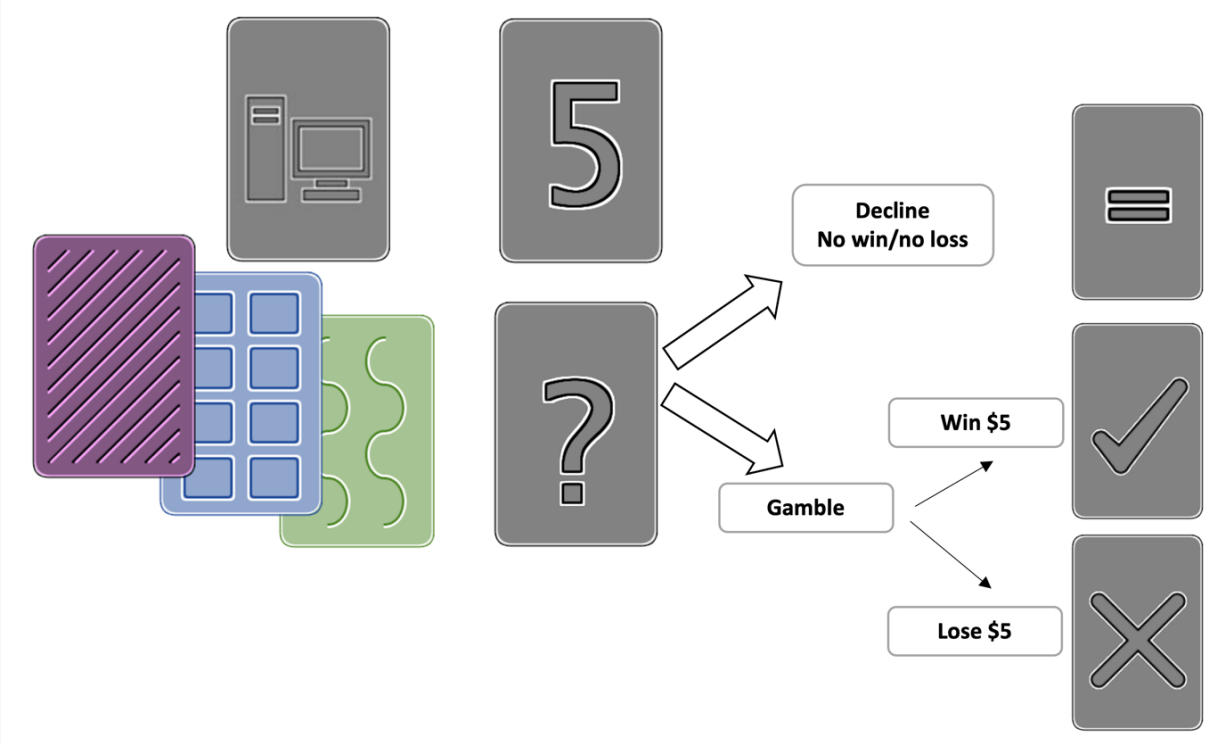


Figure 2.1: Gambling card game adapted from Michely et al. (2022).

Analysis

We used two methods to determine whether subjects were learning primarily via wins or losses. The first, model-agnostic method utilized a logistic regression approach to model task choices (Michely et al., 2022). To examine latent learning and decision-making mechanisms, we

also employed a series of computational models, including *Q*-learning models that approximate trial-and-error learning (Eldar et al., 2016; Michely et al., 2022).

Raw choice data

We first used a trial-by-trial logistic regression to model participants' decisions to either take a gamble (c_t = 1) or decline (c_t = 0) on trial *t*. The model is defined below (Michely et al., 2022):

$$p(c_t = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t})}}$$

Where, x_{1t} was the computer number that was shown to the participant on trial *t*, scaled between -1 (card number 9) and 1 (card number 1), so that the values were reversed and constrained. Cumulative success was defined by x_{2t}, or the sum of previous wins for the current deck, computed as +1 multiplied by the computer's number. Cumulative failure was defined by x_{3t}, or the sum of previous losses for the current deck, computed as -1 multiplied by the computer's number. The win or loss outcomes were multiplied by the computer's number to account for the magnitude of evidence against or for a deck, given a win or loss at a higher or lower number. To compare coefficients, x_{2t} and x_{3t} was scaled similarly to x_{1t}, such that the ranges of values were between -1 and 1.

β_0 was defined as the intercept. A positive coefficient for the initial predictor (β_1) signified that participants exhibited a greater inclination to gamble against lower computer-generated numbers. Likewise, a positive coefficient for the second predictor (β_2) suggested that participants were more prone to gamble when using a deck that yielded higher cumulative

success. Similarly, a positive coefficient for the third predictor (β_3) indicated that participants were more inclined to decline a gamble when confronted with a deck associated with higher cumulative failure.

The logistic regression was estimated in Matlab (R2024b) to be consistent with Michely et al. (2022). The model was estimated per subject, producing weights that were then compared by group with Wilcoxon rank sum tests. Raw accuracy per deck was tested with mixed effects linear regression, specifying an interaction term between deck and group and a subject-level random intercept. Similarly, propensity/probability of gambling was tested with a mixed effects linear regression, with the same interaction term and subject-level random intercept. These analyses were conducted in R Studio (Version 1.3.959) using the *lme4* package.

Computational modeling

We compared ten algorithms to identify the one that best explains participants' choices (Eldar et al., 2016; Michely et al., 2022). The first three are considered "null" models and do not include update functions relevant to *Q*-learning models. The null models were utilized as a validity check to ensure that participants were engaging in trial-by-trial feedback learning and not relying solely on a heuristic-dependent strategy to make choices (e.g., using only the computer's number). In all algorithms, the probability of taking each gamble (i.e., the action policy) was estimated using a logistic function applied to a term that represented all available evidence, as has been defined in our model-agnostic analysis.

Algorithm 1 ("no learning model")

Evidence: $\beta + \beta' N_t$

Where, β is defined as a decision bias parameter that influences the participant's static propensity to gamble, β' is an inverse temperature that influences likelihood to gamble based on the computer's number, and N_t is the computer's number (scaled between -1 and 1).

Algorithm 2 ("no learning model + general persistence")

Update function: $if\ a_t = gamble$: $\begin{cases} p_t^{gamble} = 1 \\ p_t^{decline} = \lambda p_t^{decline} \end{cases}$

$if\ a_t = decline$: $\begin{cases} p_t^{decline} = 1 \\ p_t^{gamble} = \lambda p_t^{gamble} \end{cases}$

Evidence: $\beta + \beta' N_t + \beta'' \Delta p_t$

Where, β'' controlled persistence strength and p^a was defined as a persistence variable for each action a (i.e., "gamble" or "decline"). For each trial, p_t^a was set to 1 if the associated action was taken, and decayed by free parameter λ otherwise. Δp_t was then defined as p_t^{gamble} - $p_t^{decline}$.

Algorithm 3 ("no learning model + deck-specific persistence")

Update function: $if\ a_t^d = gamble$: $\begin{cases} p_t^{d_t,gamble} = 1 \\ p_t^{d_t,decline} = \lambda p_t^{d_t,decline} \end{cases}$

$if\ a_t^d = decline$: $\begin{cases} p_t^{d_t,decline} = 1 \\ p_t^{d_t,gamble} = \lambda p_t^{d_t,gamble} \end{cases}$

Evidence: $\beta + \beta' N_t + \beta'' \Delta p_t^{d_t}$

Where, the update function and computed evidence were similar to Algorithm 2, but specific to the presented deck on trial *t.* Thus, the persistence variable p_t^a was calculated for each deck-action pair and $\Delta p_t^{d_t} = p_t^{d_t,gamble} - p_t^{d_t,decline}$.

Algorithm 4 ("Q-learning")

Update function: $Q_{t+1}^{d_t} = Q_t^{d_t} + \eta \delta_t$

Evidence: $\beta + \beta' N_t + \beta'' Q_t^{d_t}$

Where, $\delta_t = r_t - Q_t^{d_t}$ or outcome minus expected value for the shown deck ($Q_t^{d_t}$), also known as the prediction error (PE). r_t = 1 for win outcomes and r_t = -1 for lose outcomes. The learning rate was defined as η and β'' represented a stochasticity parameter that controlled the weight of the expected value for the shown deck on choice.

Algorithm 5 ("adjusted Q-learning")

Update function: $Q_{t+1}^{d_t} = Q_t^{d_t} + \eta \delta_t$

Evidence: $\beta + \beta' N_t + \beta'' Q_t^{d_t}$

Where, $\delta_t = r_t - Q_t^{d_t} - \frac{\beta'}{\beta''} N_t$ or PE minus a ratio of the weights multiplied by the computer's number at trial *t.* Thus, the algorithm learned more from outcomes that were more surprising (i.e., from win outcomes of gambles taken against higher numbers, and from loss outcomes of gambles taken against lower numbers).

Algorithm 6 ("adjusted Q-learning + associability")

Update function: $Q_{t+1}^{d_t} = Q_t^{d_t} + \alpha_t^{d_t} \eta \delta_t$

Evidence: $\beta + \beta' N_t + \beta'' Q_t^{d_t}$

Where, $\alpha_t^{d_t}$ is an associability variable that modulated learning and was calculated as the running average of the absolute values of PEs for each deck, $\alpha_{t+1}^{d_t} = \alpha_t^{d_t} + \eta'(|\delta_t| - \alpha_t^{d_t})$. η' was defined as a free parameter controlling the speed of associability variable update. The initialized associability variable, $\alpha_{t=0}^{d_t}$, was defined as a freely estimated parameter.

Algorithm 7 ("adjusted Q-learning + general persistence")

Update function: $Q_{t+1}^{d_t} = Q_t^{d_t} + \eta \delta_t$

Evidence: $\beta + \beta' N_t + \beta'' Q_t^{d_t} + \beta''' \Delta p_t$

Where, the current algorithm was a combination of algorithm 2 and algorithm 5.

Algorithm 8 ("adjusted Q-learning + deck-specific persistence")

Update function: $Q_{t+1}^{d_t} = Q_t^{d_t} + \eta\delta_t$

Evidence: $\beta + \beta' N_t + \beta'' Q_t^{d_t} + \beta''' \Delta p_t^{d_t}$

Where, the current algorithm was a combination of algorithm 3 and algorithm 5.

Algorithm 9 ("adjusted Q-learning + associability + deck-specific persistence")

Update function: $Q_{t+1}^{d_t} = Q_t^{d_t} + \alpha_t^{d_t}\eta\delta_t$

Evidence: $\beta + \beta' N_t + \beta'' Q_t^{d_t} + \beta''' \Delta p_t^{d_t}$

Where, the current algorithm was a combination of algorithm 3 and 6.

Algorithm 10 ("learning bias")

Algorithms 4 – 9 were compared with Algorithms 1 – 3 to ensure that the Q-learning models fit the data better than the no-learning models. The best fitting algorithm was then modified to include a learning bias component where separate learning rates were specified for win (η^+) and loss (η^-) outcomes.

Model fitting

To fit model parameters to participant choices, we used a hierarchical Bayesian estimation procedure in the Stan modeling language (Stan Development Team, 2023). This

procedure utilizes Markov Chain Monte Carlo techniques, in this case the No-U-Turn variant of Hamiltonian Monte Carlo, which has been shown to efficiently estimate highly correlated parameters and multilevel models (Gelman et al., 2015). The algorithms described above represent the mathematical representation of learning and decision-making for a single subject, but were embedded in a multi-level random effects model with this estimation approach. To that end, the free parameters were taken as random subject-level effects that were drawn from a larger group distribution.

Normal distributions were used for the priors of the group-level means and half-normal distributions were used for the priors of the group-level standard deviations (i.e., hyperparameters and associated hyper-priors; Gelman, 2006). Because we recruited smaller sample sizes for each group, we specified weakly informative priors to minimize the influence of those priors on posterior distributions.

To test whether model parameters varied as a function of diagnostic group, group membership was incorporated as a hierarchical regression at the group level. Specifically, the group-level mean of each parameter, p, was modeled as:

$$\mu_{p,j} \sim \gamma_0 + \gamma_R is\_R_j + \gamma_{BP} is\_BP_j$$

where is_R and is_{BP} are dummy-coded indicators denoting whether participant, j, belonged to the R or BP diagnostic group, respectively, and γ parameters represented group-level regression coefficients.

To improve sampling efficiency in the hierarchical model, parameters were estimated using a non-centered parameterization. Latent subject-level parameters were first drawn from a unit normal distribution (μ^{UT}, UT = untransformed) and subsequently scaled using the

regression-derived location parameters (γ) and independently estimated variance terms (σ), yielding subject-level parameters:

$$\beta_{p,j} = \mu_{p,j} + \sigma_{p,j}\mu_{p,j}^{UT}$$

Learning rates (η), associability variable updates (η'), decay parameters (λ), and initialized associability variables ($\alpha_{t=0}$) were bounded between 0 and 1 (e.g., learning rates $\in (0,1)$); consequently, we used an inverse logit transformation to convert unconstrained values into this range. Our hyper-priors and priors are detailed below for the learning rate, as has been done with similar work (Homan et al., 2019):

$$\mu_\eta \sim Normal(0, 1)$$

$$\mu_\eta^{UT} \sim Normal(0, 1)$$

$$\sigma_\eta \sim half - Normal(0, 0.2)$$

$$\eta_{untranformed} \sim \mu_\eta + \sigma_\eta\mu_\eta^{UT}$$

$$\eta = Logit^{-1}(\eta_{untransformed})$$

Model weights for the computer's number, persistence, and deck Q values were bounded between 0 and an upper limit (e.g., weight for computer's number $\in (0,20)$). To do this, we adapted the declaration rule for [0,1] bounded parameters by multiplying the upper limit (*UL*) with the inverse probit-transformed parameter, as has been done in prior work (Ahn et al., 2017). Our hyper-priors and priors are detailed below for the computer number weight:

$$\mu_{\beta'} \sim Normal(0, 1)$$

$$\mu_{\beta'}^{UT} \sim Normal(0, 1)$$

$$\sigma_{\beta'} \sim half - Normal(0, 0.2)$$

$$\beta'_{untransformed} \sim \mu_{\beta'} + \sigma_{\beta'}\mu_{\beta'}^{UT}$$

$$\beta' = Probit^{-1}\left(\beta'_{untransformed}\right) * UL$$

Decision bias parameters (β) were unbounded and no further transformation was implemented.

Three-thousand samples were drawn after 1,000 burn-in samples for each of four chains, with no thinning, leaving a total of 12,000 samples. We assessed convergence of the chains using the Gelman-Rubin test, in which the test statistic $\hat{R}$ should remain close to 1.00 (Gelman & Rubin, 1992) and by inspecting trace plots that visualize the sequence of parameter values sampled for every chain. We also assessed sampling efficiency with estimates of effective sample size (ESS), which should be at least 100x the number of chains specified (Baribault & Collins, 2023; Vehtari et al., 2021).

Model comparison

We compared models using two methods: leave-one-out information criterion (LOOIC) and widely available information criterion (WAIC) as implemented in the *loo* package (Vehtari et al., 2017). Pointwise log-likelihoods were computed within Stan and passed to *loo* to obtain LOOIC and WAIC estimates. Lower LOOIC and WAIC values indicated superior predictive accuracy. Standard errors and expected log pointwise predictive density differences (ΔELPD) were used to quantify uncertainty related to comparisons between models.

The algorithm with the lowest LOOIC and WAIC values, and thus the best-fitting, was used to build and estimate algorithm 10. Parameter estimates from this model were used in all remaining statistical analyses. Parameter recovery was conducted on algorithm 10 to ensure that

the model was reasonably able to recover parameter estimates using simulated data. Specifically, we generated true parameter values, simulated behavioral data based on the true parameters, and recovered parameter values with the same hierarchical Bayesian estimation procedure. Recovered and true estimates were correlated to visualize correspondence. As an additional computational check, we conducted posterior predictive checks in which simulated choice data were visually inspected to ensure high similarity with the true choice data. Accuracy was tested using a mixed effects linear model with the simulated dataset to confirm that results were similar to true accuracy results.

Statistical analysis of model parameters and self-report data

All statistical analyses were conducted in R Studio (Version 1.3.959). Self-report measures were submitted to analysis of variance (ANOVA) tests to characterize differences between groups. Less than 1% of the total self-report data was missing after calculating total scores; consequently, we did not conduct multiple imputation to estimate missing datapoints. Missing trial-level data due to non-responses from the gambling task were list-wise deleted, as non-reponses would not result in updates to value tracking of cues. Subject-level estimates from positive (η^+) and negative (η^-) learning rates were visually checked for shrinkage resulting from the hierarchical estimation procedure that could pull estimates towards the group mean (Baribault & Collins, 2023). Shapiro-Wilks tests were also conducted by group to check the normalcy of the distributions; only the BP group distribution was non-normal (p<.05). Consequently, we did not further transform the subject-level estimates. Covariates in all linear models were centered for interpretability.

To test group differences in positive (η^+) and negative (η^-) learning rates, we examined the posterior distributions of fixed group effects (γ) that estimate the difference between the

reference group (UG; γ_0) and the remaining groups (R, BP; γ_R, γ_{BP}). We used a Bayesian approach to determine the proportion of the highest density interval (HDI) of the posterior distributions for γ_R and γ_{BP} that fell within the region of practical equivalence (ROPE). The γ parameters effectively served as a weight for the group contrasts and therefore represented the difference between the reference group and the R and BP groups. Consequently, we would define the null hypothesis as: $H_0 = 0$, or no difference between UG and R/BP groups. ROPE was set to [-0.1, 0.1], given hyperpriors specifying draws from a normal distribution (Kruschke & Liddell, 2018). To conservatively determine "significance", we calculated the proportion of the 95% HDI that fell within ROPE and assessed the proportion of the posterior that fell above or below 0 (dependent on the sign of the median posterior value). If <5% of the HDI fell into ROPE and >95% of the posterior fell above or below 0, we would conclude that there was strong evidence to reject the null hypothesis.

To test group differences between R and BP, we calculated mean group posteriors by adding $\gamma_0 + \gamma_R$ and $\gamma_0 + \gamma_{BP}$ and then subtracting the BP from the R posterior values to calculate the difference in groups. "Significance" was determined with the same method described above, where the null hypothesis was rejected if <5% of the HDI fell into ROPE and >95% of the posterior fell above or below 0.

To address between-group learning bias, we subtracted the negative learning rate from the positive for each subject ($\eta^+ - \eta^-$), yielding a learning bias score, such that negative scores indicated a negative learning bias (i.e., weighting learning from negative outcomes more than positive outcomes) and positive scores indicated a positive learning bias (i.e., weighting learning from positive outcomes more than negative outcomes). We conducted a Kruskal-Wallis test to examine group differences. To control for age and trait anxiety, we then conducted a multiple

linear regression with learning bias as the dependent variable and group, trait anxiety, and age as independent variables. R was set as the reference group. Coefficients of interest were the main effect of group; post-hoc analyses of group were family-wise Bonferroni corrected (p<.025).

To examine associations between learning bias and self-report measures, we used spearman correlations. We combined the R and BP groups into one eating disordered group in order to examine relationships with eating disorder severity subscales dimensionally and to limit the number of comparisons made. Correlation analyses were conducted separately for the R/BP combined group and the UG group, and family-wise Bonferroni corrected for the number of comparisons made (p<.004). Any significant correlations, after correction, were included in a multiple linear regression analysis, with learning bias as the dependent variable, the self-report measure as the independent variable, and centered age and BMI included as covariates.

As an exploratory aim, we conducted a canonical correlation analysis (CCA) to test associations between eating pathology and winning model parameters using the *CCP* package. CCA is an efficient method for multivariate analysis involving two sets of variables and can be used for dimensionality reduction. Multivariate methods can also reveal patterns in behavior-symptom profiles that extend current understanding of relationships between model parameters and eating disorder pathology. In this case, we took a data-driven approach to understanding these relationships holistically, as a combination of behavioral, cognitive, and symptom-level measures. Because our eating pathology measure (EPSI) does not have a total or global score, we implemented a multivariate approach here to explore these relationships. Specifically, CCA was conducted between EPSI subscale scores (Set 1) and model parameters (Set 2). This approach yielded orthogonal variations of linear combinations that best explain the variance between and within the two sets of variables, called canonical dimensions or variates. The number of

canonical dimensions mirrors the number of variables in the smaller of the two sets, however we report only on the dimensions that obtained statistical significance using F approximations of various test statistics (e.g., Wilks lambda). We then report which model parameter(s) explained the most variance for each dimension.

RESULTS

Descriptive statistics for each group are shown in Table 2.1.

Table 2.1: Descriptive statistics by group. R – restricting; BP – binge-eating/purging; UG – undergraduates. STAI-T – State Trait Anxiety Inventory; EPSI – Eating Pathology Symptom Inventory; BIS – Behavioral Inhibition System; BAS – Behavioral Activation System; TCI – Temperament and Character Inventory.

Group	BP		R		UG		
Variable	**Mean/N**	**Std.Dev/%**	**Mean/N**	**Std.Dev/%**	**Mean/N**	**Std.Dev/%**	**Test**
Age	22.6	3.49	22.2	4.9	20.1	3.45	F=3.938**
Gender	25		28		40		X2=8.708*
... Female	20	80%	28	100%	30	75%	
... Male	5	20%	0	0%	9	22.5%	
... Non-binary	0	0%	0	0%	1	2.5%	
BMI	22	4.69	18.7	2.35			F=10.02***
STAI-T	60.6	11.1	60.2	10.9	42.4	10.7	F=31.414***
EPSI Body Dissatisfaction	20.2	6.36	20.5	7.25	11.6	6.57	F=19.217***
EPSI Binge Eating	12.2	8.14	4.66	4.01	9.6	6.27	F=10.34***
EPSI Cognitive Restraint	7.72	3.29	7.62	3.63	4.62	2.77	F=10.438***
EPSI Purging	5.84	6.02	1.86	3.75	0.5	1.22	F=15.443***
EPSI Restriction	11.4	7.29	13.3	6.76	5.12	4.13	F=17.898***
EPSI Excessive Excercise	7.92	6.91	8.04	7.46	7.72	6.03	F=0.018
EPSI Muscle Building	3.52	3.22	2.83	3.46	3.7	4.36	F=0.461
BIS	24.5	2.54	25	2.62	22.2	3.62	F=8.419***
BAS Drive	10.6	2.4	9.86	3.04	11.2	2.22	F=2.323
BAS Fun Seeking	10.9	2.29	9.86	2.75	12.8	1.95	F=14.31***
BAS Reward Responsiveness	16.2	3.06	15.4	2.62	17.8	1.89	F=8.434***
TCI Harm Avoidance	24.1	7.69	25.7	7.17	16.8	7.94	F=13.258***
Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01. N(BMI) = 50; N(EPSI Body Dissatisfaction) = 93.							

Raw Choice Data

Results from the model-agnostic, logistic regression analysis of the choice data revealed no between-group differences for weights associated with the computer number shown,

cumulative success, and cumulative failure, suggesting groups performed similarly overall in terms of factors influencing decisions to gamble/not gamble (Figure 2.2). However, linear models revealed within-subjects differences between weights associated with the computer number and the remaining weights, specific to group (Table 2.2). For the R group, the weight for cumulative success was significantly lower than the weight for the computer number (B = -2.34, p = .03), suggesting greater tendencies to gamble with lower numbers than gamble with decks linked to greater cumulative success (e.g., prioritizing rule-bound, "safe" gambles over adjusting gambling behavior to maximize success). For the BP group, the weight for cumulative failure was significantly lower than the weight for the computer number (B = -2.93, p = .006), suggesting greater tendencies to gamble with lower numbers than decline gambles with decks linked to greater cumulative failure (e.g., prioritizing rule-bound, "safe" gambles over adjusting gambling behavior to maximize avoidance of failure). Finally, the UG group showed significantly lower weights for both cumulative success (B = -2.81, p < .001) and cumulative failure (B = -2.12, p = .02) compared to the computer number, suggesting greater tendencies to gamble with lower numbers (e.g., prioritizing rule-bound, "safe" gambles over adjusting gambling behavior based on cumulative success or failure). Figure 2.3 illustrates the probability of gambling by group and deck over the course of the task, showing adequate learning of the task structure and the of the decks. Mixed effects linear regressions examining group differences in probability of gambling for each deck and task accuracy for each deck revealed no significant group differences (p>.05; Supplemental Figures 2.1-2.2, Supplemental Tables 2.1-2.2).

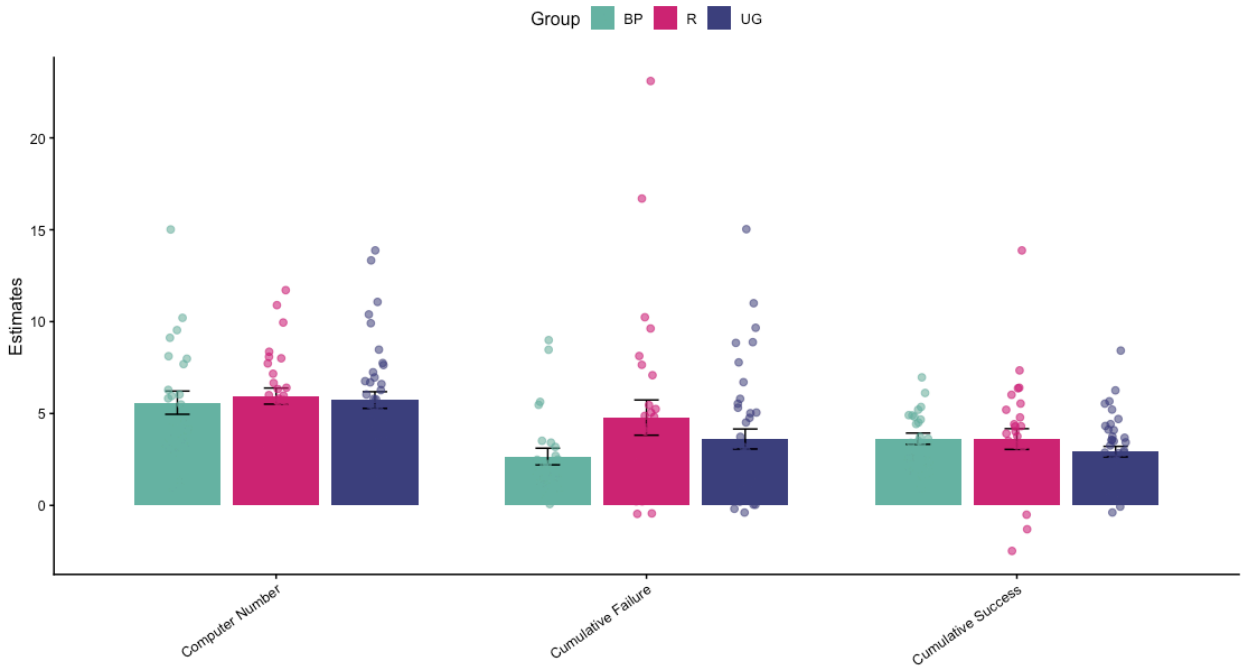


Figure 2.2: Logistic regression beta weights for the computer number, cumulative failure, and cumulative success. No significant between-group effects were observed. R – restricting, BP – binge-eating/purging, UG – undergraduates.

Table 2.2: Logistic regression results by group. R – restricting, BP – binge-eating/purging, UG – undergraduates.

	R				**BP**				**UG**			
Predictors	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*
Intercept	5.95	0.56	4.84 – 7.05	**<0.001**	5.59	0.60	4.40 – 6.78	**<0.001**	5.73	0.48	4.79 – 6.67	**<0.001**
Cumulative Failure	-1.17	0.79	-2.73 – 0.39	1.000	-2.93	0.85	-4.61 – -1.25	**0.006**	-2.12	0.68	-3.45 – -0.79	**0.017**
Cumulative Success	-2.34	0.79	-3.90 – -0.78	**0.031**	-1.96	0.85	-3.64 – -0.28	0.202	-2.81	0.68	-4.14 – -1.48	**<0.001**
Group[UG]	-0.22	0.74	-1.67 – 1.23	1.000	0.14	0.77	-1.37 – 1.66	1.000				
Group[BP]	-0.36	0.82	-1.98 – 1.26	1.000					-0.14	0.77	-1.66 – 1.37	1.000
Cumulative Failure x Group[UG]	-0.95	1.04	-3.00 – 1.10	1.000	0.81	1.09	-1.33 – 2.95	1.000				
Cumulative Success x Group[UG]	-0.47	1.04	-2.52 – 1.58	1.000	-0.85	1.09	-2.99 – 1.30	1.000				
Cumulative Failure x Group[BP]	-1.76	1.17	-4.05 – 0.53	1.000					-0.81	1.09	-2.95 – 1.33	1.000
Cumulative Success x Group[BP]	0.38	1.17	-1.92 – 2.67	1.000					0.85	1.09	-1.30 – 2.99	1.000
Group[R]					0.36	0.82	-1.26 – 1.98	1.000	0.22	0.74	-1.23 – 1.67	1.000
Cumulative Failure x Group[R]					1.76	1.17	-0.53 – 4.05	1.000	0.95	1.04	-1.10 – 3.00	1.000
Cumulative Success x Group[R]					-0.38	1.17	-2.67 – 1.92	1.000	0.47	1.04	-1.58 – 2.52	1.000
Observations	282				282				282			
R^2 / R^2 adjusted	0.137 / 0.112				0.137 / 0.112				0.137 / 0.112			

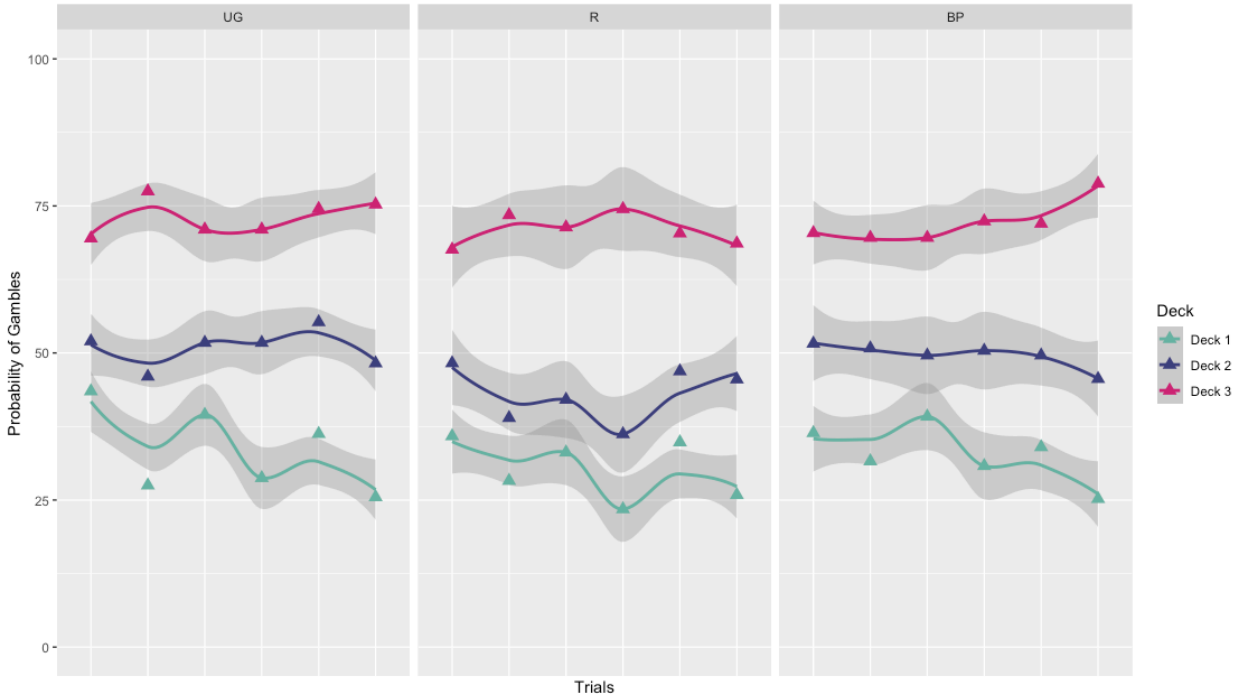
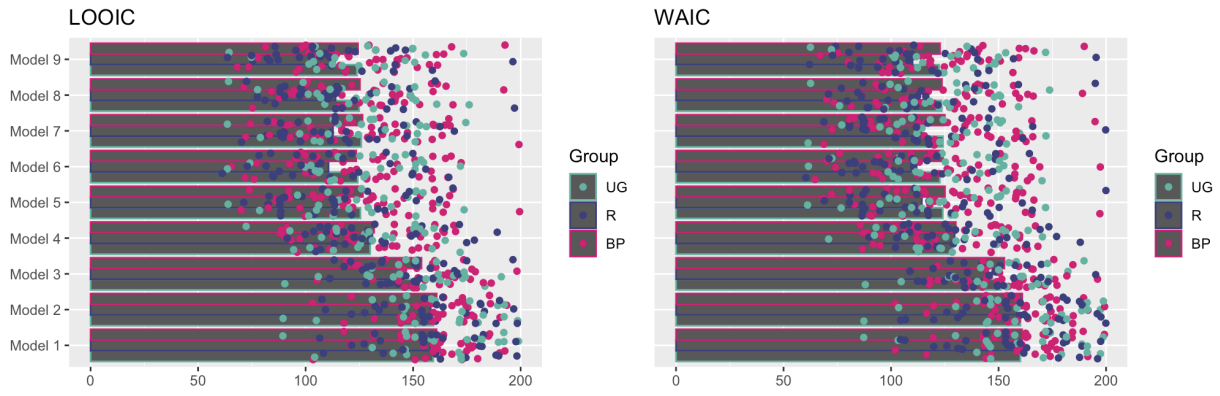


Figure 2.3: Visualization of participants' probability of gambling for each deck over the 180 trials of the task. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck ("high risk" deck); Deck 2 – even deck ("moderate risk" deck); Deck 3 – high deck ("low risk" deck).

Computational Modeling

Model comparison

Figure 2.4 shows the LOOIC and WAIC performance for each model. Algorithm 9 outperformed the other models when comparing ΔELPD for both LOOIC and WAIC approaches. Algorithm 10 was therefore constructed using algorithm 9 and all subsequent analyses were implemented with model parameters estimated from algorithm 10.



model	elpd_diff	se_diff
model9	0.00000	0.00000
model6	-59.01601	27.10162
model8	-62.35073	15.32268
model7	-129.01781	29.14138
model5	-129.70631	30.18680
model4	-542.98579	74.25567
model3	-1411.89119	97.84983
model1	-2086.11442	137.12588
model2	-2087.92673	136.96554

model	elpd_diff	se_diff
model9	0.00000	0.00000
model8	-65.14789	15.45513
model6	-71.23357	28.27541
model7	-131.08271	29.31607
model5	-131.47740	29.68594
model4	-545.04008	74.66418
model3	-1430.21353	98.26473
model1	-2107.88958	137.49637
model2	-2108.79341	137.60851

Figure 2.4: LOOIC (left) and WAIC (right) point estimates for each group. ΔELPD and standard errors are listed in the tables below each figure. Model 9 is the winning algorithm. R – restricting; BP – binge-eating/purging; UG – undergraduates. elpd_diff – ΔELPD estimates; se_diff – ΔELPD standard errors.

Computational checks

Estimation. Estimated parameters from algorithm 10 were assessed for adequacy of chain convergence visually with trace plots. Trace plots were acceptable. $\hat{R}$ values were ~1.0 (range[0.99,1.00]) for all freely estimated parameters and ESS values were all >400, indicating good convergence.

Parameter recovery. Parameter recovery results for algorithm 10 are shown in Supplemental Figure 2.3 for all estimated parameters. Correlations between true and recovered subject-level parameter estimates were all significantly and positively correlated (p<.05), suggesting adequate recovery. The recovered λ and η' parameters showed shrinkage; however,

given the adequacy of our posterior predictive checks, detailed below, we determined that the recovery remained acceptable overall.

Posterior predictive checks. Visual inspection of the simulated choice patterns over the course of the task showed adequate similarity to the true choice data (Supplemental Figures 2.4-2.5). Additionally, no group effects met significance when testing accuracy and probability of gambles, as was the case for the true data (Supplemental Table 2.3).

Analysis of model parameters and self-report data

Figure 2.5 shows transformed posterior distributions of the group means for positive and negative learning rates. Supplemental Figure 2.6 shows the posterior distributions for the remaining freely estimated parameters. Bayesian analysis of computational model parameters revealed strong evidence of attenuated positive (η^+) and negative (η^-) learning rates for the BP group compared to both UG and R groups (Figure 2.6). Posterior distributions for the contrast between BP and UG positive learning rates showed 1% of the HDI fell within ROPE, and with 100% of the probability of direction falling below 0. Posterior distributions for the contrast between BP and UG negative learning rates showed 2% of the HDI fell within ROPE, and with 99% of the probability of direction falling below 0. As for the difference between R and BP groups, posterior distributions for the positive learning rate showed 0% of the HDI fell within ROPE, and with 100% of the probability of direction falling above 0; posterior distributions for the negative learning rate showed 1% of the HDI fell within ROPE, and with 99% of the probability of direction falling above 0. Of note, the probability of direction is opposite for comparisons with UG and R, as the former is a contrast (γ) and the latter is a difference term.

All other computational model parameters did not show strong evidence of group differences (Supplemental Table 2.4).

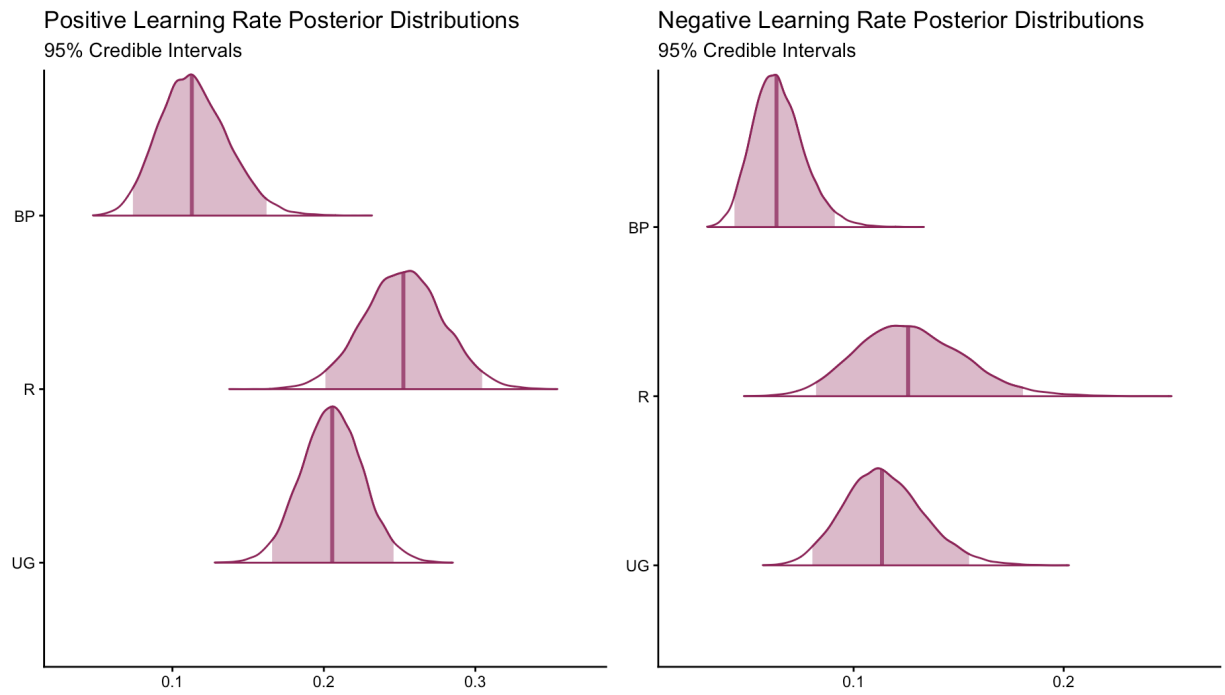
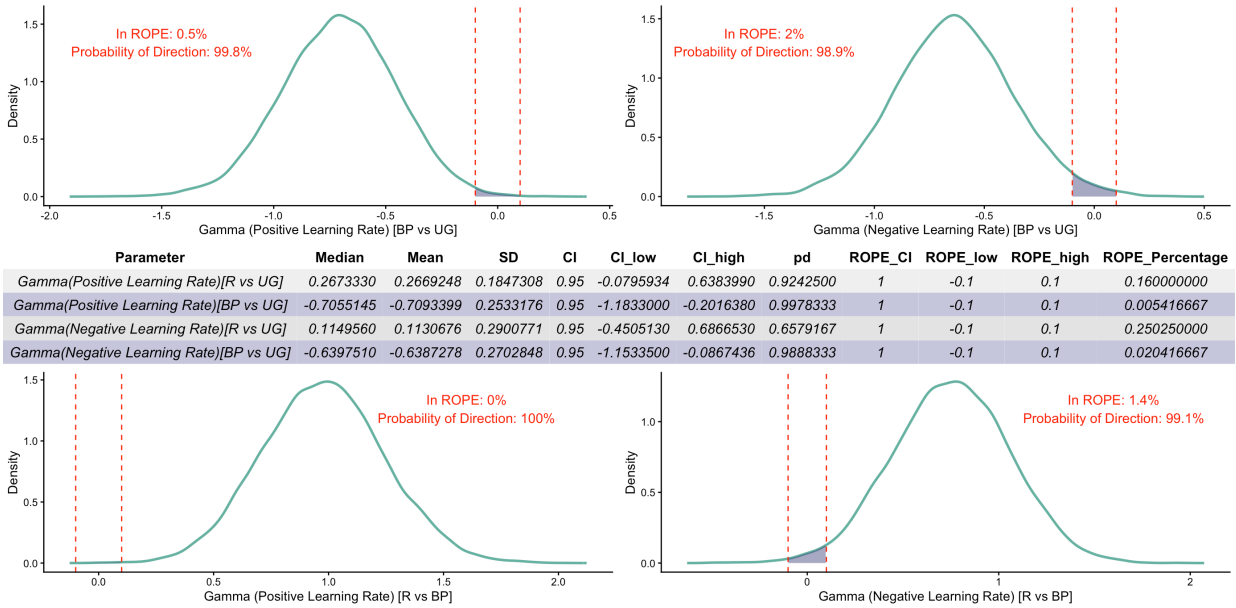


Figure 2.5: Transformed group-level posterior distributions for positive (left) and negative (right) learning rates. Credible intervals are 95% highest density intervals (HDI) and are represented by the shaded areas of the distributions. R – restricting; BP – binge-eating/purging; UG – undergraduates.



Parameter	Median	Mean	SD	CI	CI_low	CI_high	pd	ROPE_CI	ROPE_low	ROPE_high	ROPE_Percentage
Gamma(Positive Learning Rate)[R vs UG]	*0.2673330*	*0.2669248*	*0.1847308*	*0.95*	*-0.0795934*	*0.6383990*	*0.9242500*	*1*	*-0.1*	*0.1*	*0.160000000*
Gamma(Positive Learning Rate)[BP vs UG]	*-0.7055145*	*-0.7093399*	*0.2533176*	*0.95*	*-1.1833000*	*-0.2016380*	*0.9978333*	*1*	*-0.1*	*0.1*	*0.005416667*
Gamma(Negative Learning Rate)[R vs UG]	*0.1149560*	*0.1130676*	*0.2900771*	*0.95*	*-0.4505130*	*0.6866530*	*0.6579167*	*1*	*-0.1*	*0.1*	*0.250250000*
Gamma(Negative Learning Rate)[BP vs UG]	*-0.6397510*	*-0.6387278*	*0.2702848*	*0.95*	*-1.1533500*	*-0.0867436*	*0.9888333*	*1*	*-0.1*	*0.1*	*0.020416667*

Parameter	Median	Mean	SD	CI	CI_low	CI_high	pd	ROPE_CI	ROPE_low	ROPE_high	ROPE_Percentage
Gamma(Positive Learning Rate)[R vs BP]	*0.975247*	*0.9762647*	*0.2661036*	*0.95*	*0.4554236*	*1.495311*	*0.9998333*	*1*	*-0.1*	*0.1*	*0.00050000*
Gamma(Negative Learning Rate)[R vs BP]	*0.754868*	*0.7517954*	*0.3063097*	*0.95*	*0.1470950*	*1.334167*	*0.9914167*	*1*	*-0.1*	*0.1*	*0.01416667*

Figure 2.6: Posterior distributions for gamma parameters (top), indicating the difference between positive and negative learning rates for BP versus UG. Shaded regions represent percentage of the 95% highest density interval (HDI) in the region of practical equivalence (ROPE) and red dotted lines represent ROPE. Posterior distributions for difference terms (bottom) between positive and negative learning rates for R versus BP. Tables display posterior distribution metrics, including percentage of the HDI that falls within ROPE and the probability of direction (pd). CI refers to 95% HDI. R – restricting; BP – binge-eating/purging; UG – undergraduates.

Wilcoxon rank sum tests showed that learning bias, defined as the difference between positive and negative learning rates, differed significantly between groups after Bonferroni correction (Figure 2.7). The R group showed higher learning bias compared to both UG (W = 358, p = .02) and BP (W = 125, p < .001) groups, suggesting higher positive learning rates relative to negative learning rates. The BP group showed lower learning bias compared to both R and UG (W = 263, p = .003) groups, suggesting either greater balance between positive and negative learning rates (as learning bias approaches zero) or higher negative learning rates relative to positive. A post-hoc linear regression model with all participants included (Figure 2.8) showed that learning bias was associated with probability of gambling, such that the lower a

participant's learning bias, the greater the difference between propensity to gamble with the first deck ("low deck"/"high risk deck") and the third deck ("high deck"/"low risk deck"; B = -1.12, SE = 0.25, $p < .001$). This would suggest that higher negative learning rates, relative to positive learning rates, could result in behavior that is more intensely stratified by advantageous and disadvantageous decks. Importantly, this may also be associated with low learning rates, so that when deck values are learned, they are not quickly updated leading to more deck-specific choices that exploit past knowledge about the deck. This stratification becomes even more evident with lower overall propensity to gamble and lower learning bias. For example, Supplemental Figure 2.7 shows estimated marginal means for a linear model that predicts propensity to gamble with a three-way interaction term of deck, group, and learning bias. The R group displays slightly less propensity to gamble across decks and shows that the differences between decks become slightly (but not significantly) more steep at lower learning bias.

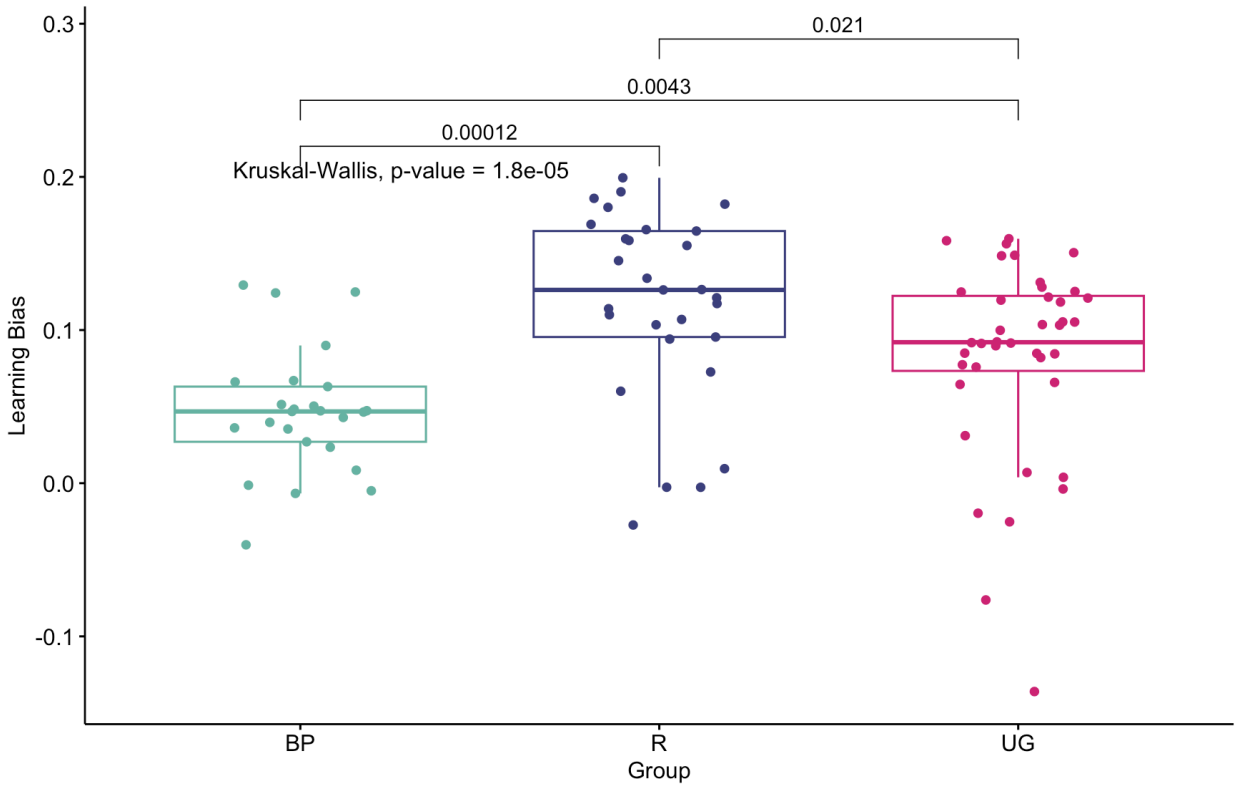


Figure 2.7: Boxplots of learning bias per group. Significant Wilcoxon rank sum tests and *p*-values are displayed. The R group shows significantly higher learning bias compared to the UG and BP groups. The BP group shows significantly lower learning bias compared to the UG group. R – restricting; BP – binge-eating/purging; UG – undergraduates.

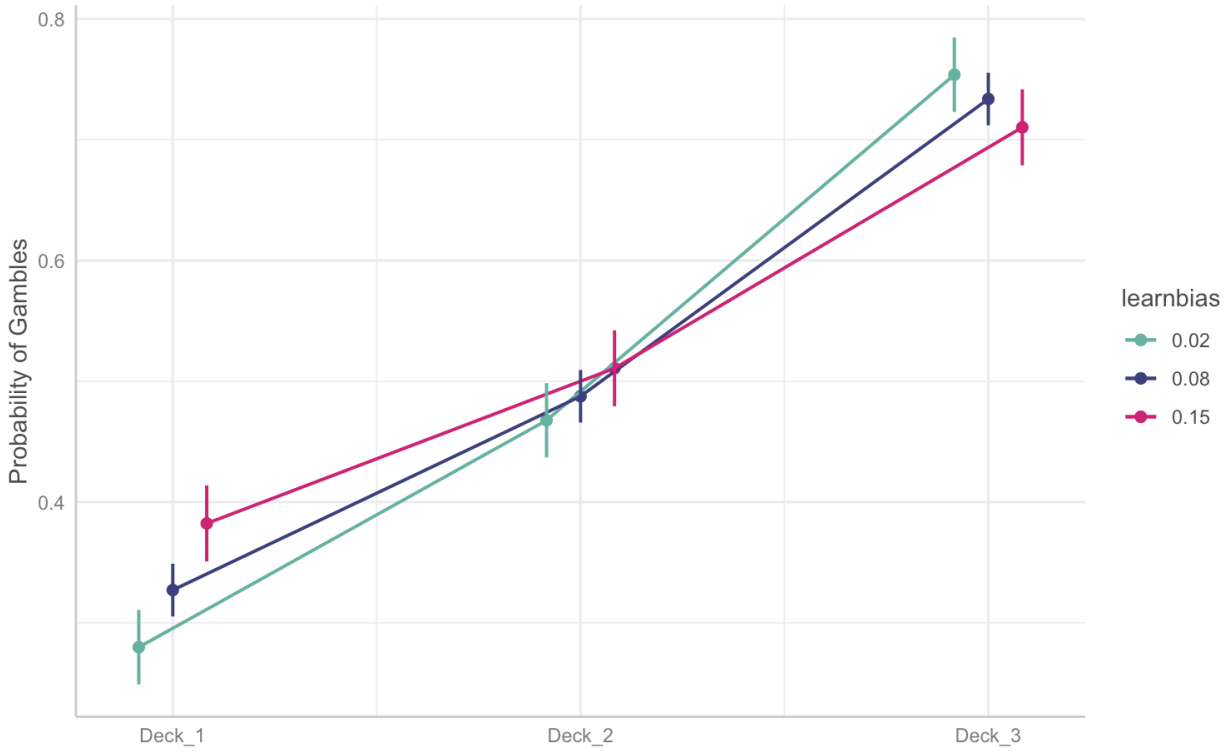


Figure 2.8: Line graph of estimated marginal means of probability of gambling for each deck by three levels of learning bias (.02, .08, .15). Deck 1 – "low deck"/"high risk deck"; Deck 2 – "even deck"/"moderate risk deck"; Deck 3 – "high deck"/"low risk deck".

Spearman correlations between learning bias and self-report measures for the combined R/BP group showed one significant association after Bonferroni correction (Figure 2.9). Specifically, learning bias was negatively correlated with the EPSI binge eating subscale (r = -0.43, p = .02)), suggesting greater positive learning bias is associated with reduced binge eating. This association held after controlling for age and BMI in a linear regression model (B = -0.004, SE = 0.001, p < .001; Supplemental Figure 2.8). No correlations for the UG group remained significant after correction.

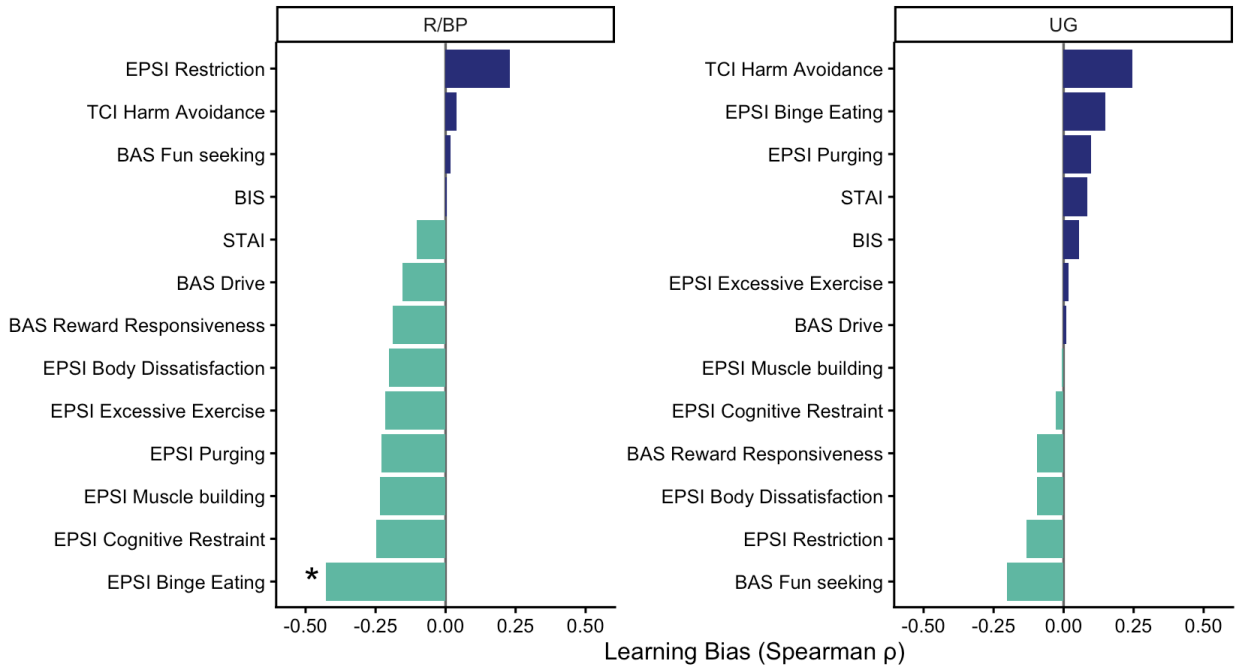


Figure 2.9: Association between learning bias and self-report measures for the combined R/BP group (left) and the UG group (right). R – restricting; BP – binge-eating/purging; UG – undergraduates. STAI-T – State Trait Anxiety Inventory; EPSI – Eating Pathology Symptom Inventory; BIS – Behavioral Inhibition System; BAS – Behavioral Activation System; TCI – Temperament and Character Inventory.

The exploratory canonical correlation analysis across all participants resulted in two significant canonical dimensions (Figure 2.10). Dimension 1 produced a correlation of .62 between the sets of computational model parameters and the eating disorder severity variables ($F_{49,\ 410.6}$, $p < .001$). Dimension 2 produced a correlation of .60 between the sets of variables ($F_{36,\ 358.5}$, $p = .002$). Supplemental Table 2.5 displays the standardized canonical coefficients for each dimension. Only the first two dimensions were significant; consequently, only the coefficients associated with these two dimensions will be reported here. The first dimension was most strongly influenced by body dissatisfaction (-.52) and restriction (-.42) from set 1, and the bias (.54) and associative learning update (.80) parameters from set 2, suggesting that augmented associative learning speed and a higher tendency to gamble may be associated with lower body

dissatisfaction and restrictive behavior. The second dimension was most strongly influenced by binge eating (-.82) and purging (-.55) from set 1, and positive learning rates (.64) from set 2. Consistent with our results showing a negative association between learning bias and binge eating severity, this dimension may reflect the relationship between learning from positive outcomes and reduced binge eating and purging behavior.

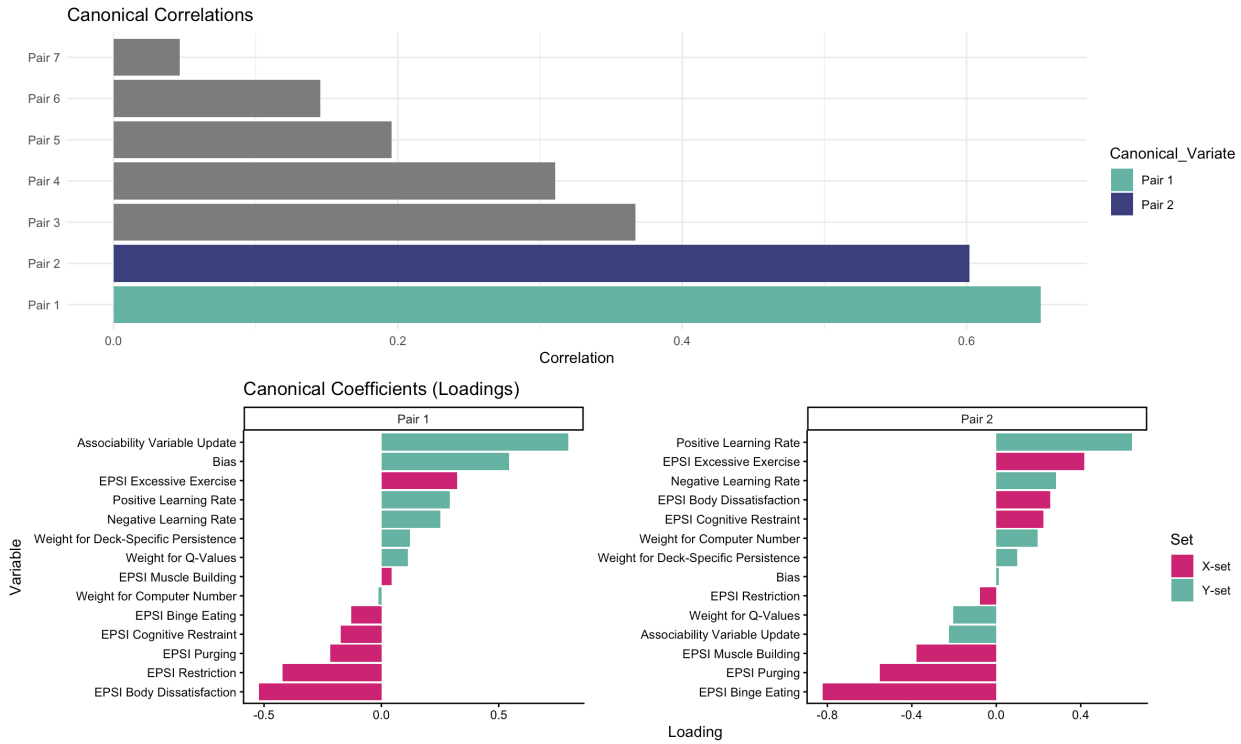


Figure 2.10: Canonical correlation analysis plots. Canonical dimension correlation coefficients (top); two pairs were significant. Canonical coefficient loadings for the first and second pairs (bottom) by set. X-set – eating disorder severity measures; Y-set – computational model parameters. EPSI – Eating Pathology Severity Inventory.

DISCUSSION

This study investigated learning bias in individuals with eating disorders who were subtyped based on restricting phenotype (anorexia nervosa restricting subtype; R) and binge-eating/purging phenotype (anorexia nervosa binge-eating/purging subtype and bulimia nervosa; BP), compared to a sample of undergraduates (UG) during a probabilistic gambling card task.

Learning bias was operationalized as the difference between positive and negative learning rates and generally quantified the degree of asymmetry in valenced learning, such that a positive learning bias represented faster updating for wins (relative to updating for losses), and negative learning bias represented faster updating for losses (relative to updating for wins). Notably, as learning bias approached zero, the learning rates become more "balanced" and less asymmetrical. Counter to our hypothesis, results from this study showed that the BP group was characterized by lower learning bias compared to both the R and UG groups, while the R group was characterized by higher learning bias compared to BP and UG groups. For the R and BP groups, learning bias was negatively associated with binge eating, linking valenced learning to eating disorder symptomatology. Exploratory canonical correlation analysis between computational model parameters and eating disorder severity self-report measures (Eating Pathology Symptom Inventory; EPSI) revealed two dimensions that linked faster associability variable updating and greater propensity to gamble to lower body dissatisfaction and restriction, and that linked higher positive learning rates to lower binge eating and purging. Overall, these results provide initial evidence of distinct latent mechanisms involved in reward learning, particularly under risk, that differentiate restricting and binge-eating/purging presentations of eating disorders. Attenuated learning rates and, consequently, learning bias in the BP group may reflect reduced sensitivity to prediction errors and reduced integration of the prediction error signal into cue valuation across valences.

Despite no observed group differences in raw choice behavior over the course of the task, computational modeling revealed dissociable speed of learning between the BP and UG groups and the R and BP groups. The BP group displayed attenuated positive and negative learning rates compared to R and UG groups, suggesting slower integration of the prediction error signal into

the computation of estimated value for each deck. This result is consistent with prior computational work in bulimia nervosa (BN), where women remitted from BN displayed reduced Bayesian prediction-error-dependent activation in the left dorsal caudate during an inhibitory control task while in a fed state and women currently ill with BN showed slower belief updating when formerly learned stimulus-response associations had to be inhibited (Berner, Fiore, et al., 2023; Berner, Harlé, et al., 2023). These studies, in combination with the results presented here, provide initial evidence of alterations in the pace of value adjustments following surprising or unexpected information about a stimulus, cue, or action and suggest that individuals with BP may be slower to adapt to changing environmental contingencies. This could be a putative mechanism involved in binge eating and purging cycles, in which reduced sensitivity to prediction errors (Frank et al., 2011), and consequently slower learning, result in difficulties updating previously learned responses to stimuli.

In this study, our BP group included both individuals diagnosed with BN and with anorexia nervosa binge-eating/purging subtype (AN-BP) in an effort to isolate the binge-eating/purging phenotype and examine cognitive processes unique to this eating disorder expression. Prior work has shown that AN-BP groups are not only dissociable from AN-R groups, but they have also shown greater similarity in task behavior, self-reported reward sensitivity, self-reported impulsivity, and cortisol reactivity to other binge-eating/purging eating disorder presentations, including BN and binge eating disorder (Dougherty et al., 2024; A. Harrison et al., 2010; Martinelli et al., 2024; P. Monteleone et al., 2014; Steward et al., 2017). Even so, there are relatively few studies that have examined differences in AN subtypes or that have taken a phenomenological approach to group categorization, as opposed to diagnostic. Here, we show that the R and BP groups display dissociable rates of learning; R was

characterized by faster positive and negative learning rates compared to the BP group, but did not show strong evidence of differences compared to the UG group. However, learning bias was increased in the R group relative to both the UG and the BP group, contrary to our hypothesis.

Higher learning bias in the current study represented greater weighting of positive learning rates relative to negative learning rates, indicating faster learning from wins compared to losses. Because the restricting subtype of AN has shown augmented punishment sensitivity and trait harm avoidance in prior studies (Glashouwer et al., 2014; A. Harrison et al., 2010; Jonker et al., 2020), we hypothesized that the R group would display more negative learning bias compared to the BP and UG groups. The opposite was observed, where the BP group showed lower learning bias and the R group showed higher. Although increased learning bias may be linked to reward sensitivity, valenced learning can be influenced by contextual factors, including reward rate, controllability, and volatility (Eckstein et al., 2022; Nussenbaum & Hartley, 2024). For example, a post hoc analysis of the interacting effect of learning bias and deck on probability of gambles in this study showed that, as learning bias increases, differences in probability of gambles tend to decrease or flatten; however, as learning bias decreases, differences between decks become more exaggerated. Although not a significant difference compared to the other groups, the R group showed a slight tendency towards reduced gambling behavior for every deck and opted to decline gambles at a higher rate, particularly for the low and even decks. Consequently, they would not be exposed to as many losing (or winning) outcomes, leading to less learning for these decks and reduced deck-specific behavior. This would also result in faster learning and more gambles for the high deck, as it was associated with the highest likelihood of wins. Therefore, learning bias results for the R group could potentially be interpreted as an "avoidant" learning strategy. Asymmetrical learning in favor of positive outcomes, within the

context of reduced gambling behavior overall, could result in reduced optimization of deck-specific gambling behavior, as less was learned about negative outcomes. Optimal responding to and valuation of stimuli in riskier states, in this case the low deck, could therefore be affected. Indeed, slower punishment learning has been reported in individuals currently ill with AN and remitted from AN during a probabilistic reward learning task (Wierenga et al., 2022). Importantly, Wierenga et al. (2022) reported that magnitude of negative prediction errors in the AN group was negatively associated with BMI at discharge, indicating that less optimized punishment learning may be linked to poorer outcomes, such as less weight restoration. Strengthening this possibility, recovery and weight-restoration in AN has been linked to *greater* punishment sensitivity (Bernardoni et al., 2021; Filoteo et al., 2014; Wierenga, Bischoff-Grethe, et al., 2025).

In contrast to the R group, the BP group showed lower learning bias (although not necessarily negative learning bias) which reflected less valence-specific weighting of positive and negative learning rates. Limited differentiation between positively and negatively-valenced feedback has been reported in a sample of women remitted from BN in prior work (Wagner et al., 2010) and may suggest alterations in the processing of rewarding and punishing outcomes. As has been described above, lower learning bias resulted in greater deck-specific distinction in gambling behavior, such that propensity to gamble was much lower for the low deck ("high risk") relative to the high deck ("low risk"). However, the BP group showed attenuated learning rates and reduced speed of learning for both wins and losses; despite the relative weight of learning from rewards and punishments possessing greater balance, this group still integrated prediction errors into deck valuation more slowly than the UG and R groups. In other words,

even though they learned from positive and negative outcomes at a more symmetrical rate, their rate of learning overall may have impacted the optimization of their choices.

Consistent with this, a negative association between learning bias and binge eating severity was observed in this study, specifically for the eating disorder groups. No significant relationship between learning bias and any self-report measure was observed in the UG group. This result suggests that attenuated bias towards learning from positive outcomes is linked to greater binge eating severity in the current sample. Higher sensitivity to punishment and harm avoidance have been commonly linked to eating disorders transdiagnostically (A. Harrison et al., 2010, 2011), however some theories of binge-eating/purging cycles, particularly in BN, reflect the hypothesis that harm avoidance (e.g., fear of gaining weight) is met with a hypo-sensitive reward system, predisposing an individual to engaging in these patterns of behavior (Frank, 2013). For example, augmented approach biases towards food may lead to binge eating (Brockmeyer et al., 2015), a behavior rife with reward-punishment conflict. Repeated binge eating behavior, despite the negative consequences that follow, are believed to reflect compulsivity that becomes entrenched and impervious to changes in environmental contingencies (Voon, 2015). In the BP group, lower learning bias, in the context of attenuated learning rates, may therefore reflect the combination of sluggish updating of values with the reduced weighting of positive feedback (relative to negative feedback) leading to less flexible behavior.

Relatedly, exploratory canonical correlation analysis revealed two significant canonical dimensions. The first was most explained by negative weighting of body dissatisfaction and restriction and positive weighting of the associability variable update and gambling bias. Consequently, greater propensity to gamble and faster tracking of prediction error magnitude for

each deck appear to be important for lower body dissatisfaction and restriction. Although not significant, the R group did gamble slightly less than the BP and UG groups, which may be associated with aversion to perceived risk (Adoue et al., 2015; Jenkinson et al., 2023; Neveu et al., 2016). Faster tracking of prediction error magnitude would modulate learning rate dynamically so that larger surprises result in faster learning than smaller surprises and could represent more efficient associative learning over time. It is possible that individuals with restrictive eating disorder presentations exhibit greater willingness to engage in effortful behavior, at the cost of efficient behavior. Learned industriousness has been posited as a putative mechanism that drives non-hedonic persistence in AN, for example (Haynos et al., 2023). However, it should be noted that there were no significant differences in group-level estimates of the associability variable update here. This interpretation is therefore speculative. The second canonical dimension was most explained by negative weighting of binge eating and purging and positive weighting of positive learning rate. Again, this finding may be supportive of the hypothesis that hypo-sensitivity to reward drives binge-eating/purging behaviors; although not strongly weighted, negative learning rate was also positively weighted for this dimension, which may also imply that faster learning and greater sensitivity to prediction errors across valences may be of particular importance for reducing binge-eating/purging behaviors.

Overall, results from this study support dissociability in latent cognitive mechanisms between eating disorder phenotypes during a gambling card game. There were no significant differences in task performance, and yet this work highlights the role of equifinality in cognitive processes that may be critical to address in future treatment development for eating disorders. For example, individuals in the BP group could benefit from more focus on integration of prediction errors into valuation estimates of stimuli or choices, whereas individuals in the R

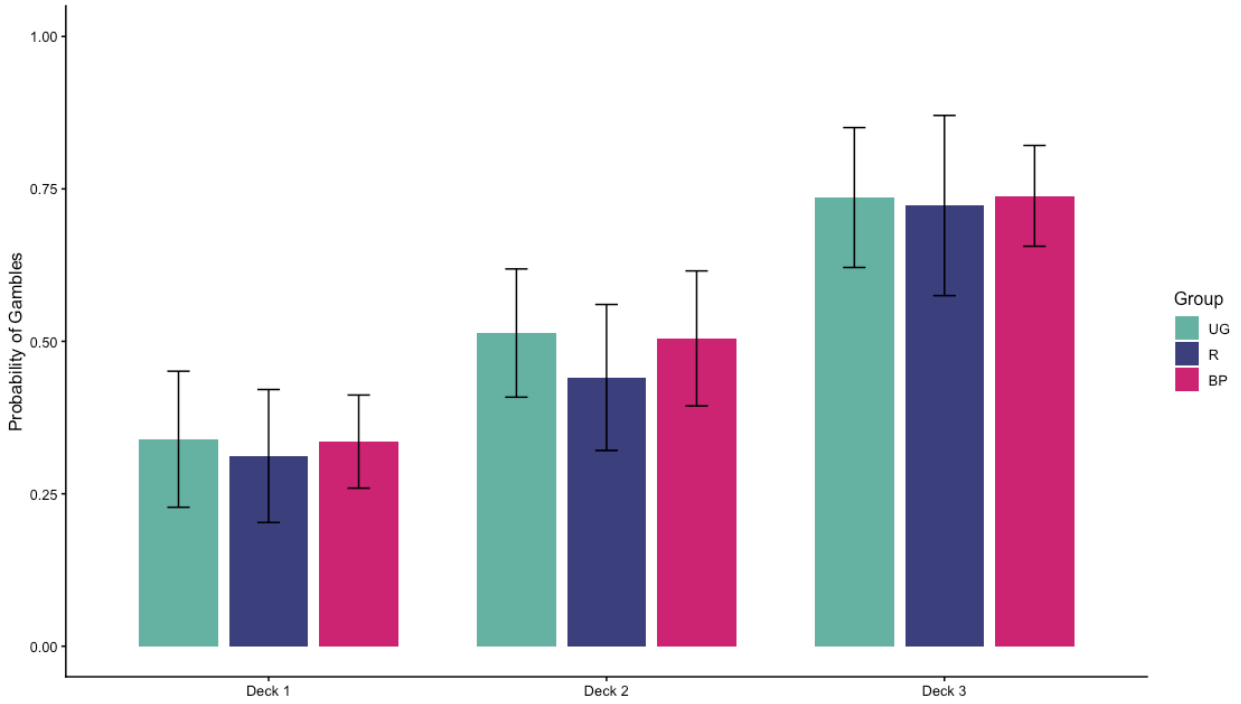
group could benefit from more balanced integration of negative information into decision-making processes.

This study had several limitations. The UG group was collected as a sample of convenience, meaning that the participants were not well characterized. Participants did not complete semi-structured clinical assessments, consequently possible psychiatric diagnoses in this group are unknown. Additionally, BMI data were not collected and could not be included as a covariate in analyses that included all groups. The heterogeneity of the UG group, however, may also be considered a strength, in that this can be considered a generalizable or representative random sample of undergraduates. Despite the heterogeneity of the sample, there remained substantive differences between the UG and the eating disorder groups, highlighting the specificity of these results to individuals diagnosed with eating disorders. The R and BP groups were also slightly older than the UG group. Age was included as a covariate in all analyses to control for differences. Yet another limitation of this work is the sample size of each group, particularly the eating disorder groups. Analyses using frequentist approaches were likely underpowered to detect group differences, so interpretations of these results should be done with caution. Additionally, combining individuals with BN and AN-BP may mask effects that are specific to these diagnoses. Future work should examine differences in learning bias between AN-R, AN-BP, and BN as separate groups to clarify whether AN-BP and BN groups are distinct. Finally, the task itself is a modified monetary version of a gambling card task that was originally used to investigate harm avoidance (Eldar et al., 2016) and, consequently, results cannot be interpreted in a manner that is overly similar to the original study. The original study operationalized positive and negative learning bias as representing active and passive harm avoidance, as the outcomes were solely punishment-related (i.e., shocks). Here, the outcome was

monetary gain or loss and learning bias was designed to represent valence-related weighting of learning rates. Finally, the task was coded to ensure relatively stable accuracy for every participant. While controlling for accuracy can be considered a strength of this work, we cannot extrapolate on accuracy per group and it is possible that potential differences on this performance metric were masked as a result.

ACKNOWLEDGEMENTS

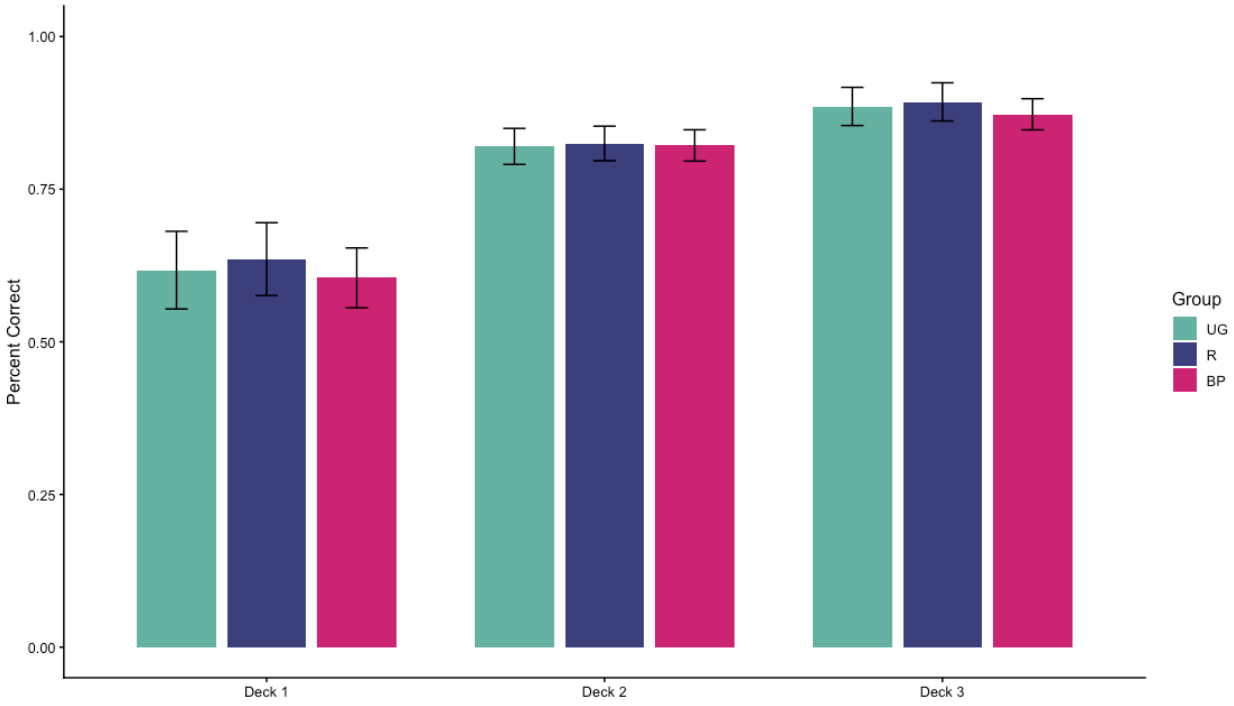
Chapter 2 contains unpublished material. I would like to acknowledge and thank: Audrey Nuñez, Amanda Bischoff-Grethe, PhD, and Christina Wierenga, PhD. I would also like to thank Nader Amir, PhD, Noriko Coburn, and the UCSD Eating Disorders Center for Treatment and Research staff and patients for their contributions to this project.



Supplemental Figure 2.1: Probability of gambling for each deck, by group. No significant between-group effects were observed. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.

Supplemental Table 2.1: Mixed effects linear regression predicting probability of gambles from deck, group, and their interaction. No significant interaction effects were observed. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.

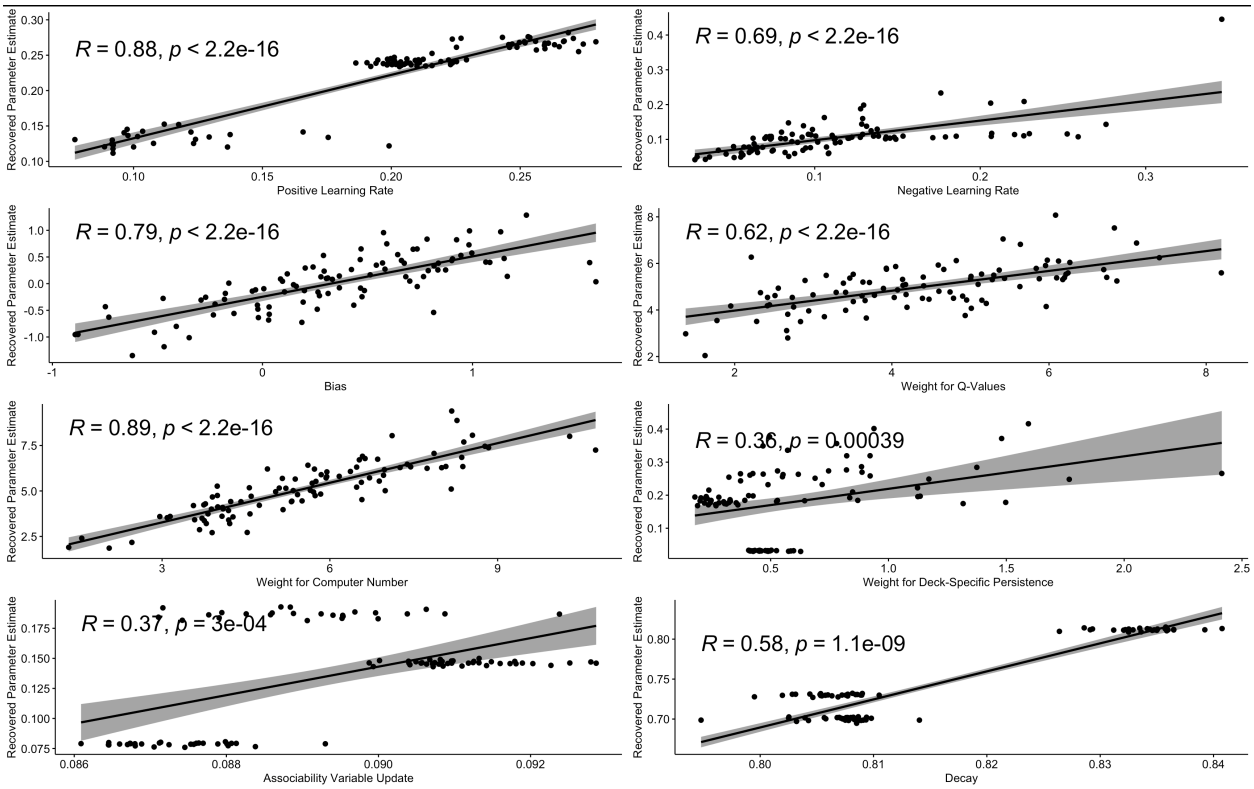
Predictors	*Estimates*	*CI*	*p*
Intercept	0.34	0.30 – 0.37	**<0.001**
Group[R]	-0.03	-0.08 – 0.03	0.313
Group[BP]	-0.00	-0.06 – 0.05	0.892
Deck[2]	0.17	0.13 – 0.22	**<0.001**
Deck[3]	0.40	0.35 – 0.45	**<0.001**
Group[R]xDeck[2]	-0.05	-0.12 – 0.03	0.239
Group[BP]xDeck[2]	-0.00	-0.08 – 0.07	0.902
Group[R]xDeck[3]	0.01	-0.06 – 0.09	0.709
Group[BP]xDeck[3]	0.01	-0.07 – 0.09	0.869
Random Effects			
σ^2	0.01		
$\tau_{00\ ID}$	0.00		
ICC	0.00		
$N_{\ ID}$	94		
Observations	282		
Marginal R^2 / Conditional R^2	0.693 / 0.693		



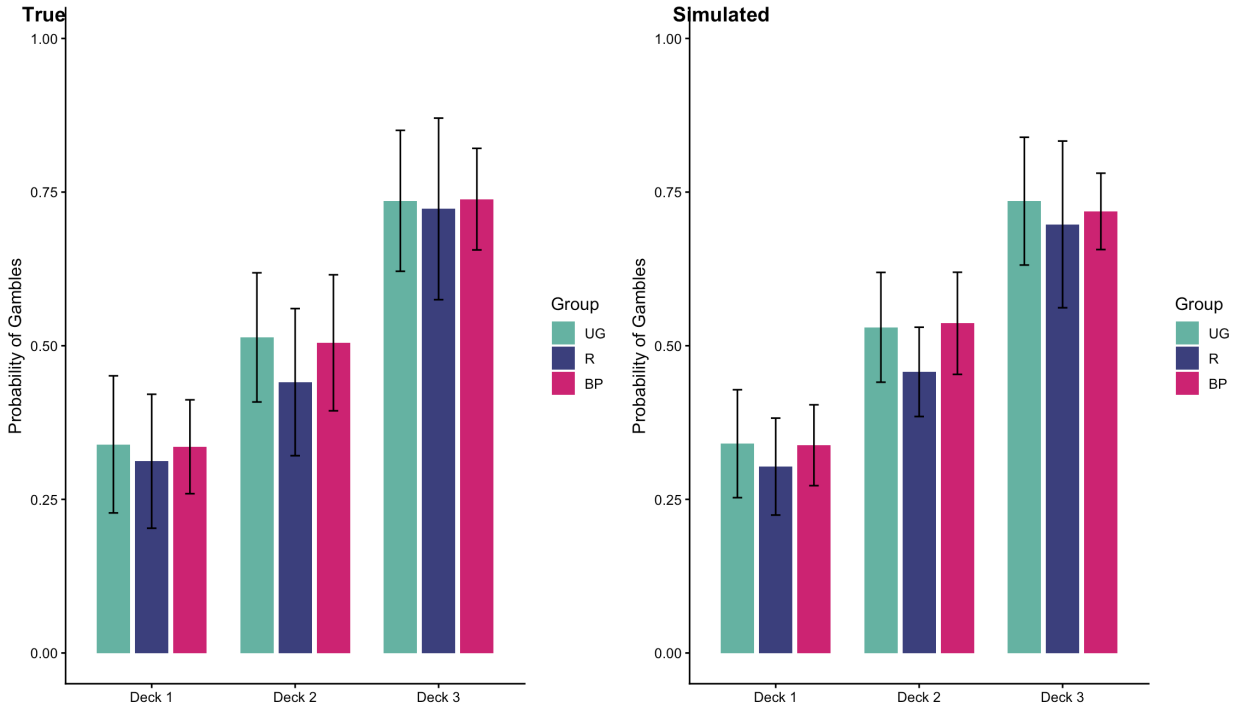
Supplemental Figure 2.2: Accuracy for each deck, by group. No significant between-group effects were observed. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.

Supplemental Table 2.2: Mixed effects linear regression predicting the percentage of correct gambles from deck, group, and their interaction. No significant interaction effects were observed. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.

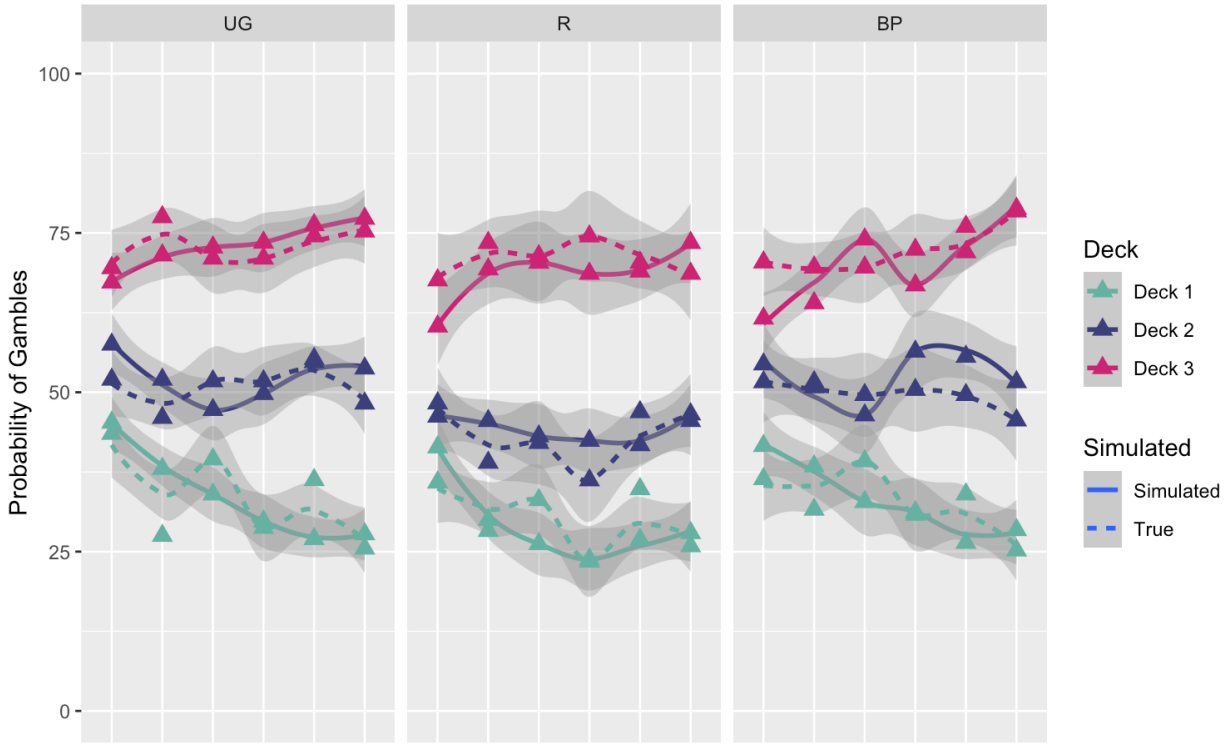
Predictors	*Estimates*	*CI*	*p*
Intercept	0.62	0.60 – 0.63	**<0.001**
Group[R]	0.02	-0.00 – 0.04	0.078
Group[BP]	-0.01	-0.03 – 0.01	0.233
Deck[2]	0.20	0.18 – 0.22	**<0.001**
Deck[3]	0.27	0.25 – 0.29	**<0.001**
Group[R]xDeck[2]	-0.01	-0.04 – 0.01	0.351
Group[BP]xDeck[2]	0.01	-0.01 – 0.04	0.337
Group[R]xDeck[3]	-0.01	-0.04 – 0.02	0.458
Group[BP]xDeck[3]	-0.00	-0.03 – 0.03	0.992
Random Effects			
σ^2	0.00		
$\tau_{00\ ID}$	0.00		
ICC	0.02		
$N_{\ ID}$	94		
Observations	282		
Marginal R^2 / Conditional R^2	0.883 / 0.885		



Supplemental Figure 2.3: Spearman correlations between recovered subject-level parameter estimates and true parameter estimates.



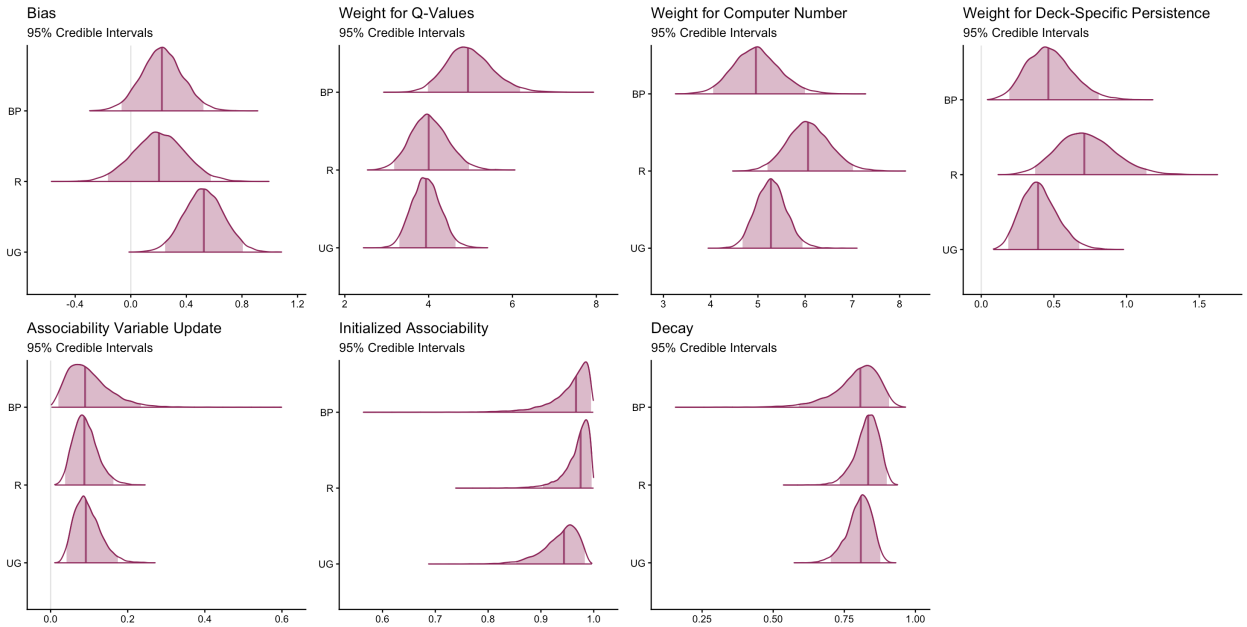
Supplemental Figure 2.4: True (left) and simulated (right) barplots of probability of gambles for each group and deck. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.



Supplemental Figure 2.5: Simulated and true probability of gambling for each group and deck over the course of the task. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.

Supplemental Table 2.3: Mixed effects linear regressions predicting simulated probability of gambles (left) and accuracy of gambles (right) from deck, group, and their interaction. No significant interaction effects were observed. R – restricting, BP – binge-eating/purging, UG – undergraduates. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck.

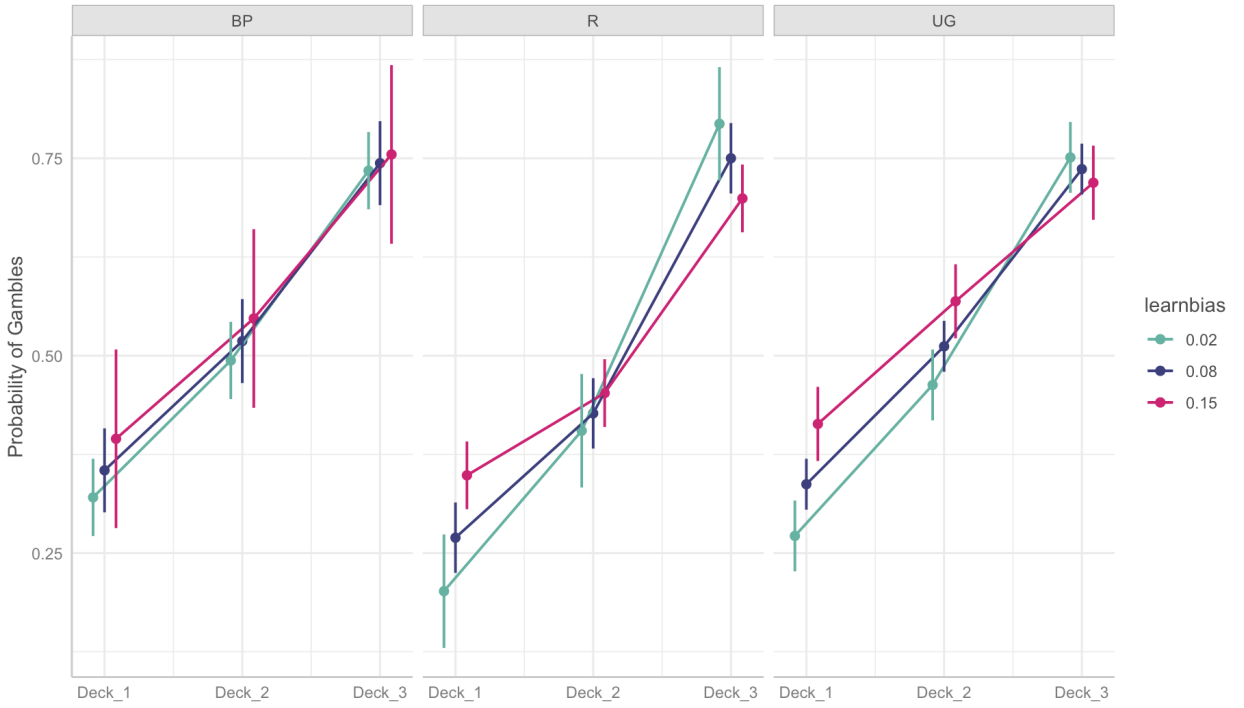
	Probability of Gambles			**Accuracy**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	0.34	0.31 – 0.37	**<0.001**	0.36	0.33 – 0.40	**<0.001**
Group[R]	-0.04	-0.08 – 0.01	0.095	0.03	-0.03 – 0.08	0.331
Group[BP]	-0.00	-0.05 – 0.04	0.914	-0.05	-0.11 – 0.00	0.069
Deck[2]	0.19	0.15 – 0.22	**<0.001**	0.23	0.18 – 0.28	**<0.001**
Deck[3]	0.39	0.36 – 0.43	**<0.001**	0.41	0.36 – 0.45	**<0.001**
Group[R]xDeck[2]	-0.04	-0.09 – 0.02	0.203	-0.04	-0.12 – 0.03	0.255
Group[BP]xDeck[2]	0.01	-0.05 – 0.07	0.756	0.00	-0.07 – 0.08	0.941
Group[R]xDeck[3]	-0.00	-0.06 – 0.05	0.984	-0.01	-0.08 – 0.06	0.794
Group[BP]xDeck[3]	-0.01	-0.07 – 0.04	0.625	0.05	-0.02 – 0.13	0.173
Random Effects						
σ^2	0.01			0.01		
τ_{00}	0.00_{ID}			0.00_{ID}		
ICC	0.22			0.06		
N	94_{ID}			94_{ID}		
Observations	282			282		
Marginal R^2 / Conditional R^2	0.762 / 0.813			0.703 / 0.720		



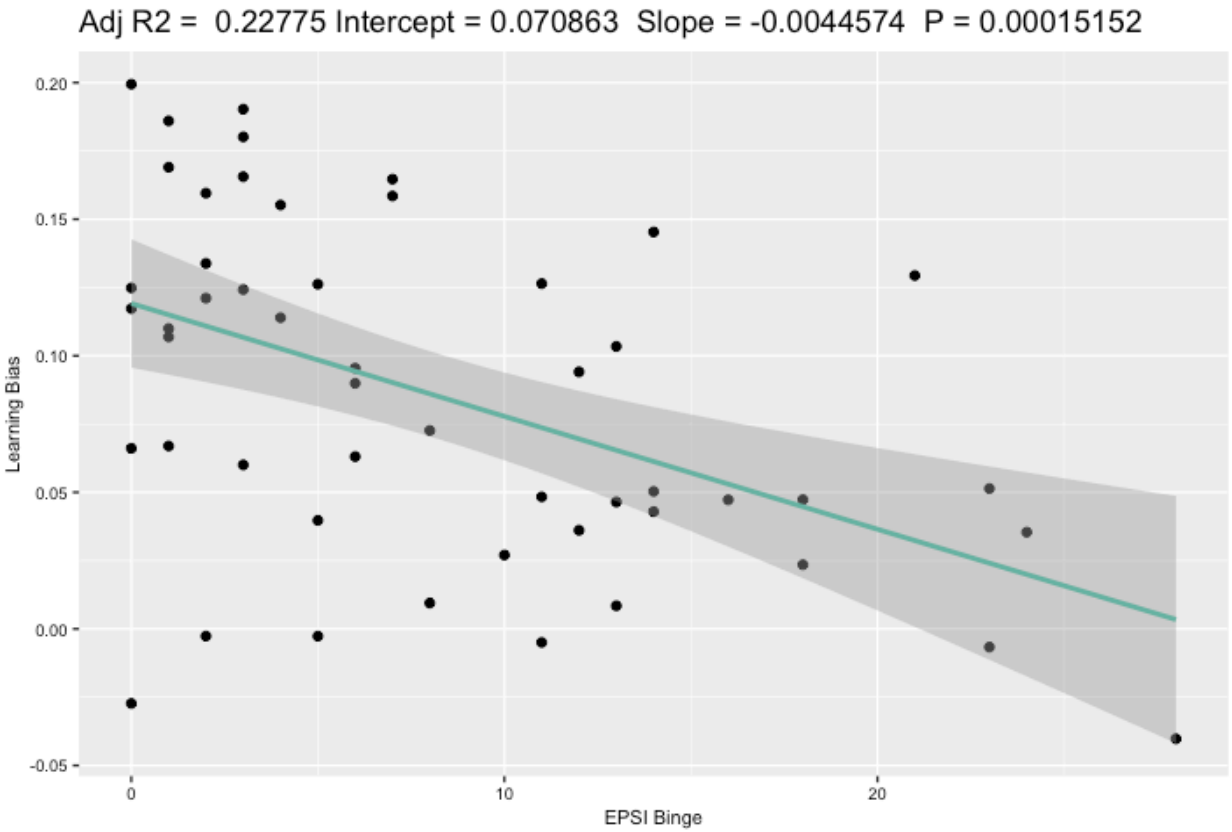
Supplemental Figure 2.6: Transformed group-level posterior distributions for bias (β), weight for the computer number (β'), weight for the Q-value (β''), weight for deck-specific persistence (β'''), associability variable update (η'), initialized associability ($\alpha_{t=0}^{d_t}$), and decay (λ). Credible intervals are 95% highest density intervals (HDI) and are represented by the shaded areas of the distributions. R – restricting; BP – binge-eating/purging; UG – undergraduates.

Supplemental Table 2.4: Posterior distribution metrics, including percentage of the HDI that falls within ROPE and the probability of direction (pd), for bias (β), weight for the computer number (β'), weight for the Q-value (β''), weight for deck-specific persistence (β'''), associability variable update (η'), initialized associability ($\alpha_{t=0}^{d_t}$), and decay (λ). CI refers to 95% HDI. R – restricting; BP – binge-eating/purging; UG – undergraduates.

Parameter	*Median*	*Mean*	*SD*	*CI*	*CI_low*	*CI_high*	*pd*	*ROPE_CI*	*ROPE_low*	*ROPE_high*	*ROPE_Percentage*
Gamma(Bias)[R vs UG]	-0.32	-0.32	0.23	0.95	-0.77	0.14	0.92	1	-0.10	0.10	0.13
Gamma(Bias)[BP vs UG]	-0.30	-0.30	0.20	0.95	-0.70	0.10	0.93	1	-0.10	0.10	0.13
Gamma(Weight for Q-Value)[R vs UG]	0.01	0.01	0.10	0.95	-0.19	0.21	0.55	1	-0.10	0.10	0.68
Gamma(Weight for Q-Value)[BP vs UG]	0.17	0.17	0.11	0.95	-0.04	0.38	0.95	1	-0.10	0.10	0.24
Gamma(Weight for Computer Number)[R vs UG]	0.12	0.12	0.08	0.95	-0.05	0.28	0.92	1	-0.10	0.10	0.41
Gamma(Weight for Computer Number)[BP vs UG]	-0.05	-0.05	0.09	0.95	-0.23	0.13	0.71	1	-0.10	0.10	0.66
Gamma(Weight for Deck-Specific Persistence)[R vs UG]	0.26	0.26	0.18	0.95	-0.09	0.62	0.92	1	-0.10	0.10	0.17
Gamma(Weight for Deck-Specific Persistence)[BP vs UG]	0.07	0.06	0.20	0.95	-0.32	0.45	0.64	1	-0.10	0.10	0.38
Gamma(Associability Variable Update)[R vs UG]	-0.06	-0.06	0.52	0.95	-1.12	0.93	0.54	1	-0.10	0.10	0.15
Gamma(Associability Variable Update)[BP vs UG]	-0.03	-0.06	0.69	0.95	-1.43	1.27	0.52	1	-0.10	0.10	0.12
Gamma(Decay)[R vs UG]	0.18	0.18	0.39	0.95	-0.56	0.97	0.69	1	-0.10	0.10	0.18
Gamma(Decay)[BP vs UG]	-0.01	-0.03	0.53	0.95	-1.07	1.00	0.51	1	-0.10	0.10	0.16
Gamma(Initialized Associability)[R vs UG]	0.86	0.89	0.80	0.95	-0.67	2.48	0.87	1	-0.10	0.10	0.06
Gamma(Initialized Associability)[BP vs UG]	0.56	0.57	0.86	0.95	-1.13	2.24	0.75	1	-0.10	0.10	0.08



Supplemental Figure 2.7: Line graph of estimated marginal means of probability of gambling for each deck by three levels of learning bias (.02, .08, .15) and by group. Deck 1 – low deck; Deck 2 – even deck; Deck 3 – high deck. R – restricting; BP – binge-eating/purging; UG – undergraduates.



Supplemental Figure 2.8: Association between learning bias and binge eating severity, controlling for age and BMI, in a combined group of R and BP participants. EPSI – Eating Pathology Severity Inventory.

Supplemental Table 2.5: Standardized canonical coefficients for each dimension and each set. Set 1 includes EPSI subscale; Set 2 includes computational model parameters, including positive learning rate (η_+), negative learning rate(η_-), bias (β), weight for the computer number (β'), weight for the Q-value (β''), weight for deck-specific persistence (β'''), and associability variable update (η'). EPSI – Eating Pathology Severity Inventory.

	Set 1 Variables						
	EPSI Body Dissatisfaction	EPSI Binge Eating	EPSI Cognitive Restraint	EPSI Purging	EPSI Restriction	EPSI Excessive Exercise	EPSI Muscle Building
Dimension_1	−0.52	−0.13	−0.17	−0.22	−0.42	0.32	0.04
Dimension_2	0.26	−0.82	0.22	−0.55	−0.08	0.42	−0.38

	Set 2 Variables						
	Positive Learning Rate	Negative Learning Rate	Bias	Weight for Q-Value	Weight for Computer Number	Weight for Deck-Specific Persistence	Associative Learning Update
Dimension_1	0.29	0.25	0.54	0.11	−0.01	0.12	0.80
Dimension_2	0.64	0.28	0.01	−0.20	0.20	0.10	−0.22

Chapter 3 AVERSIVE INTEROCEPTIVE LEARNING IN RESTRICTING AND BINGE-EATING/PURGING EATING DISORDER PHENOTYPES

ABSTRACT

Alterations in interoceptive processes, including perception of one's internal milieu and integration of internal sensory information, have been observed in individuals with eating disorders and may contribute to disordered eating behavior. Sensory experiences relevant to some eating disorders, including body image disturbance and disrupted hunger/satiety cues, may therefore be the consequence of improper integration of bottom-up viscerosensory inputs and top-down cognitive sensory expectations. Individuals with anorexia nervosa (AN) and bulimia nervosa (BN) have shown difficulty accurately perceiving internal sensory experiences and have shown altered insula activity, particularly during anticipation of sensory stimulation. The current study tested whether women diagnosed with a predominantly restricting eating disorder, specifically AN restricting subtype, (R; $n_R = 31$) and women diagnosed with a predominantly binge-eating/purging eating disorder, specifically AN binge-eating/purging subtype and BN (BP; $n_{BP} = 36$), displayed alterations in interoceptive learning compared to a healthy control group (HC; $n_{HC} = 25$). Participants completed a reversal learning task modified to deliver probabilistic inspiratory resistive loads via a mouthpiece and a resting-state magnetic resonance imaging scan (rsfMRI). Behavior was analyzed with a series of temporal difference learning models, while connectivity between regions-of-interest (ROI), including insula, amygdala, nucleus accumbens, and prefrontal cortex, were tested for group differences. Results showed that the R group was characterized by faster learning rates compared to the BP group and faster counterfactual learning rates compared to the BP and HC groups, while the BP group showed greater exploitative choice behavior compared to the R group. No ROI-to-ROI group differences were observed. However, there was a significant interaction effect of learning rate on group

differences between right insula (IC) and left anterior prefrontal cortex (aPFC) connectivity, such that the R group showed a more negative association between IC-aPFC connectivity and learning rate. This study suggests dissociable latent information processing from aversive sensory stimuli between R and BP groups, where the R group displayed greater sensitivity to prediction errors and faster updating of sensory expectations, but noisier decision-making and less exploitative action selection. Additionally, for the R group, faster updating of sensory expectations was associated with weaker IC-aPFC functional connectivity, and may suggest that disrupted neural synchrony in regions involved in higher-order cognitive processes is linked to faster learning rates in individuals with restrictive eating behaviors.

INTRODUCTION

Interoception, or the perception and integration of the internal physiological state of the body (Craig, 2002), is thought to be a critical phenomenological component of eating disorders (Jacquemot & Park, 2020; Klabunde et al., 2017; Martin et al., 2019). Specifically, impairments in interoception are thought to contribute to body image disturbance, inappropriate responses to hunger/satiety cues, and sensory distortions (e.g., feelings of fullness) that both propel and maintain eating disorder pathology (Badoud & Tsakiris, 2017; Berner et al., 2018; Kaye et al., 2009). Several studies have found maladaptive interoception to be central in eating disorder symptom-level networks (T. A. Brown et al., 2020; Levinson & Williams, 2020), suggesting that impairments in sensory perception and/or processing are a contributing factor in eating disorder maintenance and are even present in individuals at risk of developing these disorders (Fairburn et al., 2003; Jenkinson et al., 2018).

Notably, some disordered eating behaviors (e.g., restriction, purging) are believed to function as a means of avoiding or escaping anticipated threats (e.g., weight gain) and associated

aversive internal sensations (e.g., feeling bloated), maintained by maladaptive interoceptive processes (Berg et al., 2013; Haynos et al., 2017; Khalsa et al., 2022; Schaumberg et al., 2021). Meanwhile others have posited that altered interoceptive processing relevant to gastrointestinal functioning contributes to a range of disordered eating behaviors, including disrupted hunger/satiety cues and less or more sensitivity to stomach distension that may lead to restriction or binge eating behavior (Khalsa et al., 2022). Despite recent theoretical models proposing that aversive conditioning to physiological sensations may contribute to restrictive eating (Zucker & Bulik, 2020), there is little empirical evidence demonstrating how individuals who engage in restriction and/or binge-eating/purging behaviors learn to associate stimuli with aversive bodily sensations, which may be a key component in the pathogenesis of eating disorders more broadly. Perturbations in negative valence systems (e.g., responses to expected threat) may be central to understanding how aversive learning and subsequent anxiety-related responses are dynamically integrated into the architecture of eating disorder symptomatology. However, the bulk of emerging work in reinforcement learning (RL) investigates RL in the context of explicit reward/punishment cognitive tasks, thus we extended this framework to probe whether aversive interoceptive learning (i.e., updating beliefs about an aversive physiological stimulus), and associated neural circuitry, are altered in eating disorder presentations characterized predominantly by restrictive and binge-eating/purging symptoms.

More recent theories of interoceptive learning, informed by the field of computational neuroscience, highlight the insula as a crucial hub for the regulation of interoceptive signals and, importantly, for tracking the difference between anticipated and experienced body states (i.e., PEs; Craig, 2009). This is believed to stem from interoceptive predictive coding, a data processing strategy in which the brain attempts to minimize PEs by updating generative models

of the world (i.e., learning; Seth & Critchley, 2013). Specifically, descending interoceptive predictions are met with ascending viscerosensory inputs; the difference between these two signals (i.e., PEs) is believed to be integrated in the anterior insula and is involved in the production of subjective feeling states. Large PEs that are not effectively updated, resulting in disrupted learning, may instantiate pathological emotions and behaviors (e.g., avoidance of non-threatening stimuli (Paulus & Stein, 2006). Thus, psychopathology is purported to manifest from negative expectations or beliefs about physiological states, increased hypervigilance towards internal signs of threat, and impaired sensory learning over time (Hakamata et al., 2010; Paulus & Stein, 2010; Van den Bergh et al., 2017). For example, alterations in top-down cognitive and bottom-up somatic signaling are hypothesized to maintain eating disorders in cyclical feedback loops that are driven by misperceptions of bodily states.

Critically, insular cytoarchitecture, supported by distinct parallel connections with the striatum and prefrontal cortex (PFC), is believed to mediate visceral learning and higher-order representations of interoceptive feelings (Barrett & Simmons, 2015; Fermin et al., 2022). Neuroanatomical networks linking the insula with the striatum, PFC, and amygdala are believed to reflect interoceptive reinforcement learning processes (Adhikari et al., 2015; Allen, 2020; Gehrlach et al., 2019), with atypical activity in this loop suggesting impaired predictive coding. Specifically, connections between the insula and PFC are linked to cognitive integration of PEs, while connections between the insula and the ventral striatum and amygdala are linked to sensory predictions (Fermin et al., 2022). No study has yet examined functional connectivity in this interoceptive reinforcement learning network in eating disorders, and investigations into this circuitry may yield important insights into the neural basis of interoceptive learning relevant to restricting and binge-eating/purging presentations of eating disorder pathology.

Preliminary evidence of altered interoceptive predictive coding has been documented in individuals with eating disorders, with one study showing that, under low-arousal conditions, participants with anorexia nervosa (AN) had difficulty discriminating between experienced visceral signals and anticipated sensations (Khalsa et al., 2015), and another showing failure to adapt sensory precision estimates during a heartbeat tapping task in a mixed sample of individuals with eating disorders (Smith et al., 2020). Yet another study found that, after infusions of isoproterenol, participants with AN experienced greater panicogenic effects during meal anticipation compared to controls, and reported localized heartbeat sensations in the absence of cardiorespiratory stimulation (Khalsa et al., 2018). Neuroimaging studies in women remitted from AN (Berner et al., 2018) and women remitted from bulimia nervosa (BN; Berner et al., 2019; Wierenga et al., 2020) showed altered insula activation during an anticipatory phase of sensory stimuli tasks, compared to control women. Such findings may indicate disrupted processing of interoceptive cues and associated PEs in AN and BN (Jacquemot & Park, 2020; Oberndorfer et al., 2013; Wagner et al., 2008).

This study utilized a reversal learning paradigm with subsequent inspiratory breathing loads to investigate computational markers of aversive interoceptive learning in women diagnosed with a predominantly restricting eating disorder, specifically AN restricting subtype (R), and women diagnosed with a predominantly binge-eating/purging eating disorder, specifically AN binge-eating/purging subtype and BN (BP), and healthy controls (HC) who were demographically matched with R and BP women. Inspiratory breathing restriction is a powerful tool to perturb internal bodily states, as it engenders strong sensory inputs that are slow to habituate (Davenport & Vovk, 2009), allows for effective measurement of interoceptive PEs and their relationship to aversive learning, and has been shown to elicit strong affective responses in

individuals with AN and BN (Lapidus et al., 2020). This form of aversive stimulus, paired with a more standard reversal learning paradigm, allows for examination of interoceptive associative learning processes in a dynamic environment, which aligns with exemplar reinforcement learning behavioral paradigms (Wierenga, Brown, et al., 2025). A prior study successfully utilized this breathing load task in a sample of participants with low and moderate trait anxiety, where participants higher in trait anxiety demonstrated greater anterior insula activation when predicting an upcoming breathing load (O. K. Harrison et al., 2021). We hypothesized that the eating disorder groups would show attenuated speed of learning compared to HC and that learning rates estimated from the computational model would be negatively associated with eating disorder severity. Given the well-established role of the insula in interoceptive processing and decision-making PEs, a second aim of this study was to examine group differences in resting state functional connectivity (rsFC) between insula and brain regions commonly associated with aversive learning and interoceptive predictive coding (e.g., striatum, amygdala, PFC). We hypothesized that eating disorder groups would show attenuation of connectivity between insula and PFC and augmented connectivity of insula and amygdala compared to HC. Finally, an exploratory aim examined the relationship between behavioral and neural signatures of interoceptive learning in AN.

METHODS

Participants

A sample of acutely ill and partially remitted adult women with anorexia nervosa (AN) and bulimia nervosa (BN) were recruited from the community and the UCSD Eating Disorder Center for Treatment and Research Partial Hospitalization Program (n_{AN} = 46, n_{BN} = 29); additionally, a demographically-matched HC group was recruited from the community (n_{HC} =

26). Participants were diagnostically assessed based on DSM-5 eating disorder and comorbid psychiatric disorders using the Structured Clinical Interview for the DSM-5 (First et al., 2015). Participants assigned to an eating disorder group were required to: meet full criteria for current DSM-5 AN or BN diagnoses and be medically stable based on 1) vital signs and lab values within normal limits, 2) no complications due to coexisting medical problems, 3) no active suicidality or inability to commit to safety, and 4) weight at least 75% of ideal body weight (IBW). Control participants were required to: 1) have no eating pathology; 2) have no current Axis I disorder or medical illness; 3) not be taking psychotropic medications; and 4) have maintained 90-120% ideal body weight since menarche. All participants were required to be: 1) aged 18-39 and 2) female. Exclusion criteria included: 1) MRI contraindications; 2) psychotic or other mental illness requiring hospitalization; 3) current nicotine, alcohol, or drug dependence; 4) physical condition known to influence weight or eating (e.g., pregnancy, diabetes); 5) neurological disorder or history of head injury with >30 minute loss of consciousness; 6) neurodevelopmental disorder. To address our question about differences in eating disorder phenotype, we created a restricting group composed of only individuals diagnosed with AN restricting subtype (R; $n_R = 33$), and a binge-eating/purging group composed of both AN binge-eating/purging subtype and BN (BP; $n_{BP} = 42$). Nine participants were removed from analyses due to poor performance on the breathing learning task (BLT; for more details, see the Methods section), leaving 31 in the R group, 36 in the BP group, and 25 in the HC group.

The study was approved by the University of California, San Diego Institutional Review Board and all participants provide written informed consent.

Measures

Anthropometrics. BMI was calculated (kg/m^2) for each participant using height and weight information.

Eating disorder severity. The Eating Pathology Symptoms Inventory (EPSI; (Forbush et al., 2013) is a 45-item questionnaire used to assess eating disorder symptoms and is made up of eight subscales: Body Dissatisfaction, Binge Eating, Cognitive Restraint, Purging, Restricting, Excessive Exercise, Negative Attitudes toward Obesity, and Muscle Building. Items were rated on a 0–4 Likert scale, with higher values indicating higher frequency of eating disorder symptoms. The EPSI has demonstrated good construct and external validity within heterogeneous samples of clinical and non-clinical subjects (Coniglio et al., 2018).

Interoceptive awareness. The Multidimensional Assessment of Interoceptive Awareness (MAIA; Mehling et al., 2012) is a 32-item self-report measure of eight IA dimensions: 1) Noticing ("I notice where in my body I am comfortable"), Not-Distracting ("I distract myself from sensations of discomfort"), Not-Worrying ("I can notice an unpleasant sensation without worrying"), Attention Regulation ("I am able to consciously focus on my body as a whole"), Emotional Awareness ("When something is wrong in my life, I can feel it in my body"), Self-Regulation ("When I am caught in my thoughts, I can calm my mind by focusing on my body/breathing"), Body Listening ("I listen to my body to inform me about what to do"), and Trusting ("I trust my body sensations"). Items are rated on a 0 (never) to 5 (always) scale, with higher scores indicating better IA. The MAIA subscales have demonstrated acceptable psychometric properties in eating disorders (T. A. Brown et al., 2017).

Effortful Control. The Effortful Control Scale from the short form of the Adult Temperament Questionnaire (ATQ; Evans & Rothbart, 2007) is a 19-item self-report measure that assesses self-regulation. Participants self-reported ability to exercise control over their behavior on a 1-7 Likert scale. The ATQ Effortful Control Scale consists of three subscales: attention control, or the capacity to focus and shift attention; inhibitory control, or the capacity to suppress inappropriate approach behavior; and activation control, or the capacity to act despite strong motivation to avoid.

Procedure

Participants first completed a phone screen to assess eligibility. Those who qualified were scheduled for a screening visit to complete clinical interviews to determine diagnoses and establish final eligibility. Participants then attended a study visit that included completion of self-report measures, the Breathing Learning Task (BLT), and a resting-state functional magnetic resonance imaging scan (rsfMRI).

Breathing Learning Task. To measure dynamic updating of interoceptive beliefs, participants completed an aversive breathing load reversal learning task (BLT; O. K. Harrison et al., 2021) outside of the MRI scanner (Figure 2), which was coded in MATLAB and presented with PsychToolbox. Two visual cues were paired with an 80% or 20% chance of a subsequent inspiratory resistive load and were counter-balanced, such that each cue was first matched with an 80% chance of breathing resistance for half of the participants. Contingencies swapped four times over the course of 80 trials. Following presentation of one of the cues, participants predicted (via button press) whether they would experience a breathing resistance. Following a brief pause (2.5s), they were shown a circle representing the stimulus period (5s), where

participants either experienced breathing resistance (40% of maximal resistance) or no resistance. Participants then indicated (via button press) whether they felt there was a breathing resistance or not. Rest periods (7-9s) were pseudo-randomized between trials. Following the last trial, participants answered several visual analogue scale (VAS) questions, including: "How anxious are you?", from "Not at all anxious" to "Extremely anxious", and "How difficult was it to breathe?", from "Not at all difficult" to "Extremely difficult". Some participants (n = 84) answered two additional questions: "How confident are you that you were able to accurately predict upcoming breathing restrictions?", from "Not at all confident" to "Extremely confident", and "How confident are you that you were able to correctly identify when there was a breathing restriction?", from "Not at all confident" to "Extremely confident". To deliver the breathing resistance, participants wore a silicone rubber face mask that covered the bottom half of their face and breathed through a mouthpiece with a non-rebreathing valve (2600 series, Hans Rudolph). Resistance loads consisted of sintered bronze disks placed in series in a Plexiglas tube, with stoppered ports between disks. During the task, a 40cm H_2O/l/s breathing restriction was employed. Participants were also fitted with a pneumatic belt (GDX-RB, Vernier Graphical Analysis) placed around the abdomen to collect respiratory data, including breaths per minute (BPM).

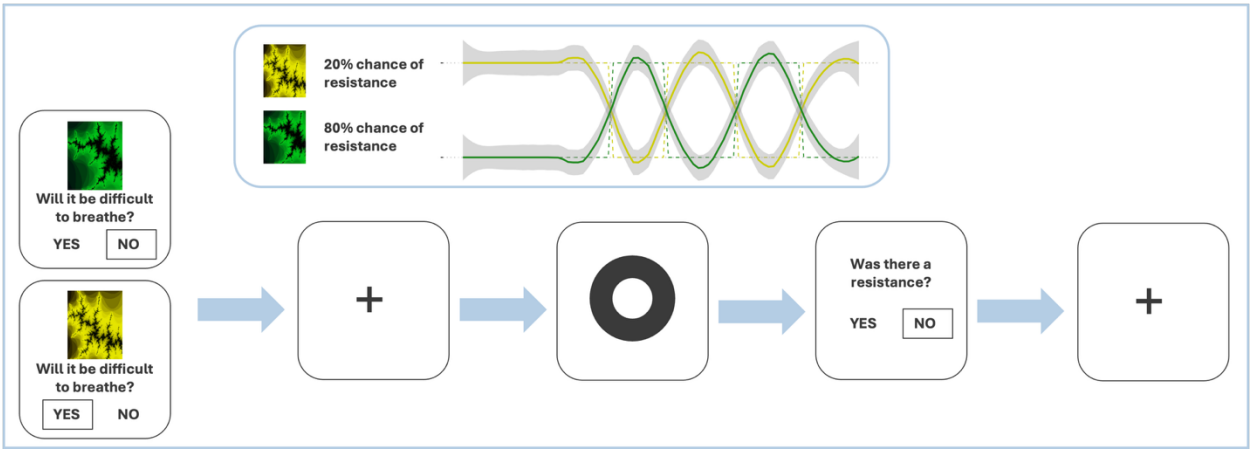


Figure 3.1: Breathing learning task (BLT) modified from Harrison et al. (2021).

Resting-state fMRI. Participants completed the imaging visit during the early follicular phase (days 1 to 10) of their menstrual cycle, if they had regular menses. The scan session took place at approximately 8:30 am following an overnight fast. Following MRI, subjects had a snack and completed self-report assessments.

Data were acquired using a 3.0 T Siemens Prisma scanner with a 32-channel head coil and followed the HCP/Lifespan standardized protocols (Glasser et al., 2013, 2016). Participants were instructed to look at fixation crosshair for the duration of the scan. The scan included two T2* weighted echo-planar imaging (EPI; TR= 720 ms, TE=37 ms, flip angle=52°, FOV=180 (RL) x 208 (AP) mm, 72 2mm thick axial slices, scan time=8:00) scans to measure BOLD resting state functional activity with phase maps to correct field inhomogeneities.

Analysis

Here we describe analysis methodology for both the behavioral and rsfMRI data. We employed computational modeling for assessment of task performance and estimation of latent learning processes, consistent with prior work investigating interoceptive learning (O. K.

Harrison et al., 2021). For the neuroimaging data, we used an ROI-to-ROI approach on our *a priori* neural regions.

Raw choice data

Behavioral data were assessed for overall accuracy of detection of breathing loads and of prediction of breathing loads. Participants who displayed <80% accuracy of detection were omitted from the sample, as inaccurate detection can result in failure to learn underlying task structure. Accuracy of prediction across reversals was compared between groups with a mixed-effects linear regression, where accuracy was predicted by a block-by-group interaction term, controlling for age, BMI, post-task VAS anxiety ratings, and post-task VAS difficulty ratings. HC was set as the reference group. The independent variables described were defined as fixed effects and a random intercept was included. Coefficients of interest were main effect of group, the main effect of block, and the block-by-group interaction effect. Post-hoc analyses of group were family-wise Bonferroni corrected (p<.025). Win-stay (i.e., probability of repeating a prior choice following a positive outcome) and lose-shift (i.e., probability of choosing an alternative choice following a negative outcome) averages for each group were calculated and compared with mixed-effects linear regression using the same method.

Computational modeling

We then compared ten algorithms to identify the one that best explained participants' choices (Crawley et al., 2020; O. K. Harrison et al., 2021). The first two were considered "null" models and did not have update functions relevant to reinforcement learning. The null models were used as a validity check to ensure that participants were indeed learning from outcomes and not relying on a heuristic choice rule or random choice. Eight algorithms were based on common

temporal difference learning models. In all temporal difference learning algorithms, the probability of choosing an action (i.e., "yes" or "no") was modeled by applying a softmax function to a term that represented all available evidence (i.e., action policy).

Algorithm 1 ("biased random decision")

Action policy: $p_t(a_1) = b \; AND \; p(a_2) = 1 - b$

Where, $p_t(a_k)$ is the probability of choosing action *k* at time *t* and b is a bias parameter that controls the probability of picking action 1 ($p(a_1)$).

Algorithm 2 ("noisy win-stay lose-shift")

Action policy: $p_t^k = \begin{cases} 1 - \frac{\epsilon}{2} \; if \; (c_{t-1} = k \; AND \; r_{t-1} = 1) \\ \quad OR (c_{t-1} \neq k \; AND \; r_{t-1} = 0) \\ \frac{\epsilon}{2} \; if \; (c_{t-1} \neq k \; AND \; r_{t-1} = 1) \\ \quad OR (c_{t-1} = k \; AND \; r_{t-1} = 0) \end{cases}$

Where, p_t^k is the probability of making choice *k* at time *t, c* is the choice made, *r* is the outcome, and ϵ is a noise parameter.

Algorithm 3 ("temporal difference learning")

Update function: $V_{t+1}^c = V_t^c + \alpha \delta_t$

Action policy: $p_t^k = \frac{\exp(\beta V_t^c)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$

Where, $\delta = r_t - V_t^c$ or outcome minus expected value for the shown cue *c* at time *t* (V_t^c), also known as the prediction error (PE). α is defined as the learning rate, or the speed of updating based on the generated PE. The probability of selecting a particular choice (p_t^k) is given by the softmax function, which defines β as an inverse temperature parameter, controlling the stochasticity of choices.

Algorithm 4 ("temporal difference learning – counterfactual update")

Update function: $V_{t+1}^c = V_t^c + \alpha\delta_t^c$

$$V_{t+1}^u = V_t^u + \alpha\delta_t^u$$

Action policy: $p_t^k = \frac{\exp(\beta V_t^c)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$

Where, $\delta_t^c = r_t - V_t^c$ is the PE for the shown cue *c* at time *t,* and $\delta_t^u = r_t - V_t^u$ represents the counterfactual PE for the unseen cue *u* at time *t*. α is defined as the learning rate, or the speed of updating based on the generated PE. The action policy remains identical to algorithm 3.

Algorithm 5 ("temporal difference learning – counterfactual update, dual learning rates")

Update function: $V_{t+1}^c = V_t^c + \alpha\delta_t^c$

$$V_{t+1}^u = V_t^u + \alpha'\delta_t^u$$

Action policy: $p_t^k = \frac{\exp(\beta V_t^c)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$

Where, the update functions are identical to algorithm 4, except that α' is defined as the learning rate for the counterfactual update. The action policy remains identical to algorithm 3.

Algorithm 6 ("temporal difference learning – separate learning rates for +/- PE")

Update function: $V_{t+1}^c = \begin{cases} if\ \delta_t \geq 0: V_t^c + \alpha_+ \delta_t \\ if\ \delta_t < 0: V_t^c + \alpha_- \delta_t \end{cases}$

Action policy: $p_t^k = \frac{\exp(\beta V_t^c)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$

Where, $\delta = r_t - V_t^k$ or outcome minus expected value for cue *c* at time *t* (V_t^c). α is defined as the learning rate for either positive (α_+) or negative (α_-) PEs. The action policy remains identical to algorithm 3.

Algorithm 7 ("temporal difference learning – separate learning rates for +/- PE, counterfactual update")

Update function: $V_{t+1}^c = \begin{cases} if\ \delta_t^c \geq 0: V_t^c + \alpha_+ \delta_t^c \\ if\ \delta_t^c < 0: V_t^c + \alpha_- \delta_t^c \end{cases}$

$$V_{t+1}^u = \begin{cases} if\ \delta_t^u \geq 0: V_t^u + \alpha_+ \delta_t^u \\ if\ \delta_t^u < 0: V_t^u + \alpha_- \delta_t^u \end{cases}$$

Action policy: $p_t^k = \frac{\exp(\beta V_t^c)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$

Where, the update functions are a combination of algorithm 4 and algorithm 6. The action policy remains identical to algorithm 3.

Algorithm 8 ("temporal difference learning – separate learning rates for +/- PE, counterfactual update, dual learning rates")

$$\text{Update function: } V_{t+1}^{c} = \begin{cases} if\ \delta_t^c \geq 0: V_t^c + \alpha_+ \delta_t^c \\ if\ \delta_t^c < 0: V_t^c + \alpha_- \delta_t^c \end{cases}$$

$$V_{t+1}^{u} = \begin{cases} if\ \delta_t^u \geq 0: V_t^u + \alpha'_+ \delta_t^u \\ if\ \delta_t^u < 0: V_t^u + \alpha'_- \delta_t^u \end{cases}$$

$$\text{Action policy: } p_t^k = \frac{\exp(\beta V_t^c)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$$

Where, the update functions are a combination of algorithm 5 and algorithm 7. The action policy remains identical to algorithm 3.

Algorithm 9 ("temporal difference learning – adaptive learning rate")

$$\text{Update function: } V_{t+1}^{k} = V_t^k + \kappa \alpha \delta_t$$

$$\alpha_{t+1} = \eta |\delta_t| + (1 - \eta)\alpha_t$$

$$\text{Action policy: } p_t^k = \frac{\exp(\beta V_t^k)}{\sum_{i=0}^{k} \exp(\beta V_t^i)}$$

Where, $\delta = r_t - V_t^k$ or outcome minus expected value for the shown cue at time *t* (V_t^k). α is defined as an adaptive learning rate that is updated based on α_t, an associability learning

rate η, and the absolute value of the PE ($|\delta_t|$). Initial learning rate was set to α_0. Additionally, κ is a scaling factor for updating, such that small PEs result in small learning rates and vice versa. The action policy remains identical to algorithm 3.

Algorithm 10 ("temporal difference learning – separate inverse temperature for +/- PE")

Update function: $V_{t+1}^k = V_t^k + \alpha\delta_t$

$$\text{Action policy: } p_t^k = \begin{cases} if\ \delta_t \geq 0\text{: } \frac{\exp(\beta_+ V_t^k)}{\sum_{i=0}^{k} \exp(\beta_+ V_t^i)} \\ if\ \delta_t < 0\text{: } \frac{\exp(\beta_- V_t^k)}{\sum_{i=0}^{k} \exp(\beta_- V_t^i)} \end{cases}$$

Where, $\delta = r_t - V_t^k$ or outcome minus expected value for the shown cue at time *t* (V_t^k). α is defined as the learning rate. The action policy remains identical to algorithm 3, except that β is labeled as the inverse temperature for either positive (β_+) or negative (β_-) PEs.

Model fitting

We used a hierarchical Bayesian estimation procedure in the Stan modeling language (Stan Development Team, 2023) to fit freely estimated parameters to trial-by-trial choices. Model parameters were estimated using Markov Chain Monte Carlo sampling, implemented with the No-U-Turn variant of Hamiltonian Monte Carlo, which efficiently samples from posterior distributions in models with correlated parameters and a hierarchical structure (Gelman et al., 2015). Due to the hierarchical nature of the data, freely estimated parameters were taken as

random subject-level effects that were drawn from larger group-level distributions, specified by diagnostic group.

Priors for group-level means were defined as normal distributions and priors for group-level standard deviations were defined as half-normal distributions (i.e., hyperparameters with associated hyperpriors; Gelman, 2006). Given the modest sample sizes within each group, weakly informative priors were used to limit the influence of prior assumptions on the resulting posterior estimates.

To mitigate estimation difficulties arising from the hierarchical structure of the data and to improve sampling efficiency, free parameters, p, were reparameterized at the subject level, j. Parameters were first drawn from a unit normal distribution (μ^{UT}, UT = untransformed) and subsequently transformed to the appropriate scale using a group-specific location parameter, μ (indexed by group, g), and independently estimated variance terms, σ. Subject-level parameters were therefore defined as:

$$\beta_{p,j} = \mu_{p,g} + \sigma_{p,j}\mu_{p,j}^{UT}$$

Learning rates (α), associability learning rates (η'), scaling factors (κ), and initialized associability variables (α_0) were bounded between 0 and 1 (e.g., learning rates $\in (0,1)$. We used an inverse logit transformation to convert unconstrained values into this range. Hyper-priors and priors are detailed below for the learning rate, as has been done with similar work (Homan et al., 2019):

$$\mu_\alpha \sim Normal(0, 1)$$

$$\mu_\alpha^{UT} \sim Normal(0, 1)$$

$$\sigma_\alpha \sim half - Normal(0, 0.2)$$

$$\eta_{untranformed} \sim \mu_\alpha + \sigma_\eta \mu_\alpha^{UT}$$

$$\eta = Logit^{-1}(\eta_{untransformed})$$

Inverse temperature parameters were bounded between 0 and an upper limit $\in (0,20)$. To do this, we adapted the declaration rule for [0,1] bounded parameters by multiplying the upper limit (*UL*) with the inverse probit-transformed parameter, as has been done in prior work (Ahn et al., 2017). Hyper-priors and priors are detailed below for the computer number weight:

$$\mu_\beta \sim Normal(0, 1)$$

$$\mu_\beta^{UT} \sim Normal(0, 1)$$

$$\sigma_\beta \sim half - Normal(0, 0.2)$$

$$\beta_{untransformed} \sim \mu_\beta + \sigma_\beta \mu_\beta^{UT}$$

$$\beta = Probit^{-1}\left(\beta_{untransformed}\right) * UL$$

Three-thousand samples were drawn after 1,000 burn-in samples for each of four chains, with no thinning, leaving a total of 12,000 posterior samples. Convergence was assessed using the Gelman–Rubin diagnostic ($\hat{R}$), which should remain close to 1.0 (Gelman & Rubin, 1992) and by visual inspection of trace plots for each parameter. Sampling efficiency was evaluated using the effective sample size (ESS), with recommended values exceeding 100x the number of chains specified (Baribault & Collins, 2023; Vehtari et al., 2021).

Model comparison

Models were compared using leave-one-out information criterion (LOOIC) and the widely applicable information criterion (WAIC), computed using the *loo* package (Vehtari et al.,

2017). Pointwise log-likelihoods were calculated within Stan and passed to *loo* to obtain LOOIC and WAIC estimates, with lower values indicating better predictive performance. Uncertainty in model comparisons was quantified using standard errors and differences in expected log pointwise predictive density (ΔELPD).

Parameter estimates from the best-fitting model were used for all subsequent analyses. To assess recoverability, true parameter values were sampled, behavioral data were simulated from these values, and parameters were then re-estimated using the same hierarchical Bayesian procedure. Recovery was evaluated by correlating true and recovered parameters. As an additional validation step, posterior predictive checks were performed by simulating choice data from the fitted model and visually comparing these simulations with the observed behavioral data. Accuracy, win-stay, and lose-shift behavior was tested using a mixed effects linear model with the simulated dataset to confirm that results were similar to true behavioral results.

Statistical analysis of model parameters and self-report data

All statistical analyses were conducted in R Studio (Version 1.3.959). Self-report measures were submitted to analysis of variance (ANOVA) tests to characterize differences between groups. Approximately 1% of the total self-report data was missing after calculating total scores; consequently, we did not conduct multiple imputation to estimate missing datapoints. Missing trial-level data due to non-responses from the BLT were list-wise deleted, as non-reponses would not result in updates to value tracking of cues. Subject-level estimates from the freely estimated computational model parameters were visually checked for shrinkage that would pull estimates to the group mean and reduce variability in the data (Baribault & Collins, 2023). Although Shapiro-Wilks tests were significant, indicating non-normalcy of the distributions, visual inspection of the data did not show substantial shrinkage, therefore we did

not further transform the subject-level estimates. Covariates in all linear models were centered for interpretability.

To test group differences in freely estimated parameters from the winning model, we examined the posterior distributions of the location parameter ($\mu_{p,g}$) for each parameter, p, and each group, g. The location parameter represented the untransformed group mean; consequently, group differences were defined by taking the difference between them (e.g., $\Delta_p = \mu_{p,R} - \mu_{p,HC}$). We used a Bayesian approach to determine the proportion of the highest density interval (HDI) of the posterior distributions for Δ_p that fell within the region of practical equivalence (ROPE). The difference terms effectively served as a weight for the group contrasts and therefore represent the difference between groups. Consequently, we defined the null hypothesis as: $H_0 = 0$, or no difference between UG and R/BP groups. ROPE was set to [-0.1, 0.1], given that location parameters were untransformed, with hyperpriors specifying draws from a normal distribution (Kruschke & Liddell, 2018). To conservatively determine "significance", we calculated the proportion of the 95% HDI that fell within ROPE and assessed the proportion of the posterior that fell above or below 0 (dependent on the sign of the median posterior value). If <5% of the HDI fell into ROPE and >95% of the posterior fell above or below 0, we would conclude that there was strong evidence to reject the null hypothesis.

To test potential associations between learning rates and eating disorder symptomatology, EPSI subscale scores were included as a dependent variable in a series of linear regression models, with learning rate parameter estimates (α) from the best fitting model, group (R, BP, HC), age, BMI, post-task VAS anxiety ratings, and post-task VAS difficulty ratings included as independent variables. HC was set as the reference group. Family-wise bonferroni correction was used to correct for multiple comparisons in the case of post hoc analyses.

Post hoc analyses were conducted to examine associations between post-task VAS certainty ratings, including certainty of detection of breathing resistance (detection certainty) and certainty of prediction of breathing resistance (prediction certainty), and the difference between learning rates and counterfactual learning rates ($\Delta_{\alpha-\alpha'}$) from the winning model, based on observed group differences in computational model parameters. A linear model was specificed, with difference values as the dependent variable and post-task VAS detection/predicton certainty ratings as the independent variable, controlling for age, BMI, post-task VAS anxiety ratings, and post-task VAS difficulty ratings. The R group was set as the reference group. Family-wise bonferroni correction was used to correct for multiple comparisons.

Statistical analysis of rsfMRI data

Functional and structural MRI data were preprocessed with a pipeline from the CONN functional connectivity toolbox (Nieto-Castanon & Whitfield-Gabrieli, 2022; RRID:SCR_009550) release 22.a, using SPM (Penny et al., 2011; RRID:SCR_007037) release 12.7771. Preprocessing pipeline consisted of realignment with correction of susceptibility distortion interactions, outlier detection, direct segmentation and MNI-space normalization, and smoothing. Functional data were realigned using SPM *realign & unwarp* procedure (Andersson et al., 2001). All scans were co-registered to the reference image using a six-parameter rigid-body transformation (Friston et al., 1995) estimated with a least-squares approach, and resampled using b-spline interpolation to correct for motion and magnetic susceptibility artifacts. Outlier scans were detected using artifact detection tools (ART). Scans were flagged as outliers if framewise displacement exceeded 0.9 mm or if global BOLD signal intensity deviated by more than 5 standard deviations from the mean. A reference BOLD image for each participant was computed by averaging all non-outlier volumes. Functional and anatomical data were

normalized into standard MNI space using SPM unified segmentation and normalization algorithm (Ashburner, 2007; Ashburner & Friston, 2005). Images were segmented into gray matter, white matter, and cerebrospinal fluid (CSF) compartments and resampled to 2-mm isotropic resolution following the normalization procedure (Calhoun et al., 2017). Functional images were smoothed with spatial convolution using a 4-mm full-width half-maximum (FWHM) Gaussian kernel.

Denoising was performed using a standard denoising pipeline (Nieto-Castanon, 2020). Nuisance regressors included white matter timeseries (Compcor noise components), CSF timeseries (CompCor noise components), movement regressors and their first-order derivatives, motion parameters and their first order derivatives, outlier scans identified using ART, session effects and their first-order derivatives, and linear trends within each run. CompCor noise components (Behzadi et al., 2007; Chai et al., 2012) within white matter and CSF were derived from eroded segmentation masks by extracting principal components orthogonal to the mean BOLD signal, motion parameters, and identified outlier volumes. Following confound regression, temporal band-pass filtering (0.008–0.09 Hz) was applied to the residual BOLD time series. The effective degrees of freedom after denoising ranged from 102.5 to 111.8 across participants (mean = 110.9).

To investigate differences in rsfMRI, we conducted an ROI-to-ROI analysis on an *a priori* network and an exploratory seed-to-voxel analysis. Specifically, we selected regions that have been implicated in an interoceptive learning network, including insula, prefrontal cortex (PFC), ventral striatum, and amygdala. Connectivity analyses were conducted in CONN (Whitfield-Gabrieli & Nieto-Castanon, 2012; RRID:SCR_009550) release 22.a (Nieto-Castanon & Whitfield-Gabrieli, 2022) and SPM (Penny et al., 2011; RRID:SCR_007037) release 12.7771.

Seventeen anatomical ROIs were defined using the Harvard-Oxford Cortical and Subcortical Atlases (Desikan et al., 2006; Frazier et al., 2005; Goldstein et al., 2007; Makris et al., 2006), including bilateral insula, accumbens, amygdala, inferior frontal gyrus, middle frontal gyrus, superior frontal gyrus, and frontal pole. Timeseries from each ROI was extracted and submitted to ROI-to-ROI, as well as seed-to-voxel, analyses, estimating both ROI-to-ROI connectivity matrices and seed-based connectivity maps. Functional connectivity strength was represented by Fisher-transformed bivariate correlation coefficients from a weighted general linear model (GLM). In order to test group effects, GLMs were estimated with functional connectivity measures and specified as dependent variables for ROI-to-ROI analyses. Voxel-wise effects were evaluated using multivariate parametric statistics with random effects across subjects and covariance estimated across repeated measurements. Statistical inference was performed at the cluster level (i.e., contiguous groups of voxels) using Gaussian random field theory (Nieto-Castanon, 2020). A cluster-defining threshold of $p < 0.001$ (uncorrected, voxel level) was applied, and clusters were considered significant at $p_{\mathrm{FDR}} < 0.05$ (familywise error-corrected at the cluster level). ROI-to-ROI and seed-to-voxel GLMs included a group-by-learning rate interaction term to test group differences in associations between intrinsic connectivity and speed of learning during the BLT.

RESULTS

Descriptive statistics for each group are shown in Table 3.1. Nine participants were removed from the sample ($n_R = 2$, $n_{BP} = 6$, $n_{HC} = 1$) due to <50% accuracy of predictions and/or <80% accuracy in detection of breathing resistance. Excluded participants were older and had lower intelligence quotient (IQ) scores compared to included participants (Supplemental Table 3.1).

Table 3.1: Descriptive statistics by group. R – restricting; BP – binge-eating/purging; HC – healthy controls. EPSI – Eating Pathology Symptom Inventory; MAIA – Multidimensional Assessment of Interoceptive Awareness; ATQ – Adult Temperament Questionnaire (Effortful Control subscale); WASI – Weschler Abbreviated Scale of Intelligence.

Group	BP		HC		R		
Variable	Mean/N	Std.Dev/%	Mean/N	Std.Dev/%	Mean/N	Std.Dev/%	Test
Education	15.2	2.05	14.3	1.97	14.4	2.08	F=2.036
Race	36		25		31		X2=12.525*
... Asian	4	11.1%	7	28%	8	25.8%	
... Black	0	0%	0	0%	2	6.5%	
... More than one race	3	8.3%	5	20%	1	3.2%	
... White	29	80.6%	13	52%	20	64.5%	
Ethnicity	36		25		31		X2=1.388
... Hispanic or Latino	5	13.9%	3	12%	7	22.6%	
... Not Hispanic or Latino	31	86.1%	22	88%	24	77.4%	
Age	24.4	5.23	22.7	4.3	23.5	4.97	F=0.905
BMI	23.2	3.8	22.7	1.89	19.2	2.24	F=18.145***
MAIA Noticing	3.02	0.815	3.06	0.905	3.06	1.09	F=0.02
MAIA Not Distracting	1.78	1.01	2.53	1.22	1.58	0.903	F=6.379***
MAIA Not Worrying	2.22	1.08	3.29	0.817	2.55	0.885	F=9.366***
MAIA Attention Regulation	2.29	0.938	2.83	0.788	2.21	0.79	F=4.235**
MAIA Emotional Awareness	3.19	0.868	2.9	1.11	2.86	1.1	F=0.978
MAIA Self Regulation	2.4	1.18	3.02	1.09	2.03	0.924	F=5.546***
MAIA Body Listening	2.18	1.16	2.41	1.28	1.8	1.22	F=1.762
MAIA Body Trust	1.69	1.2	4.05	0.916	1.89	1.33	F=34.191***
EPSI Body Dissatisfaction	20.6	4.71	5.16	3.73	17.4	7.28	F=61.524***
EPSI Binge Eating	16.9	8.82	3.76	3.69	5.03	4.89	F=40.6***
EPSI Cognitive Restraint	8.56	2.68	2.64	1.91	7.26	3.19	F=37.446***
EPSI Purging	5.08	5.04	0	0	1.68	3.24	F=15.426***
EPSI Restricting	9.47	4.77	2.12	3.19	11	5.44	F=28.321***
EPSI Excessive Exercise	9.44	6.43	3.56	3.97	6.74	5.63	F=8.221***
EPSI Obesity Attitude	8.11	5.82	2.24	2.83	6.1	6.18	F=9.034***
EPSI Muscle Building	4.97	4.63	1.72	2.73	2.9	3.46	F=5.773***
ATQ Activation Control	4.05	0.965	5.04	0.766	4.35	1.31	F=6.67***
ATQ Attentional Control	3.48	1	4.23	1	3.77	1.08	F=3.962**
ATQ Inhibitory Control	4.02	0.913	4.7	0.954	4.71	0.98	F=5.67***
WASI IQ Score	114	11.9	117	8.98	117	10.5	F=0.715

Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01. N(MAIA Noticing) = 90; N(MAIA Not Distracting) = 90; N(MAIA Not Distracting) = 90; N(MAIA Not Worrying) = 90; N(MAIA Attention Regulation) = 90; N(MAIA Emotional Awareness) = 90; N(MAIA Self Regulation) = 84; N(MAIA Body Listening) = 84; N(WASI IQ Score) = 84.

Raw Choice Data

Figure 3.2 illustrates choice behavior by group over the course of the task, showing adequate learning of the reversal structure. Results from the mixed effects linear regression models testing group differences in prediction and breathing resistance detection accuracies, win-stay, and lose-shift metrics revealed no significant main effects of group or interaction effects after correction (Figure 3.3, Table 3.2). Wilcoxon rank sum tests examining group differences in reaction times for prediction of breathing resistance and for detection of breathing resistance revealed no significant effects (Supplemental Figure 3.1). Additionally, there were no significant group differences in VAS difficulty ratings; there was a trend-level difference in VAS anxiety ratings between the BP and HC groups after correction ($W = 282.5$, $p = .06$), showing that the BP group rated their anxiety post-task as higher than the HC group (Supplemental Figure 3.2A – B). There were no group differences in post-task VAS certainty ratings for both predictions and detection of breathing resistance (Supplemental Figure 3.2C – D).

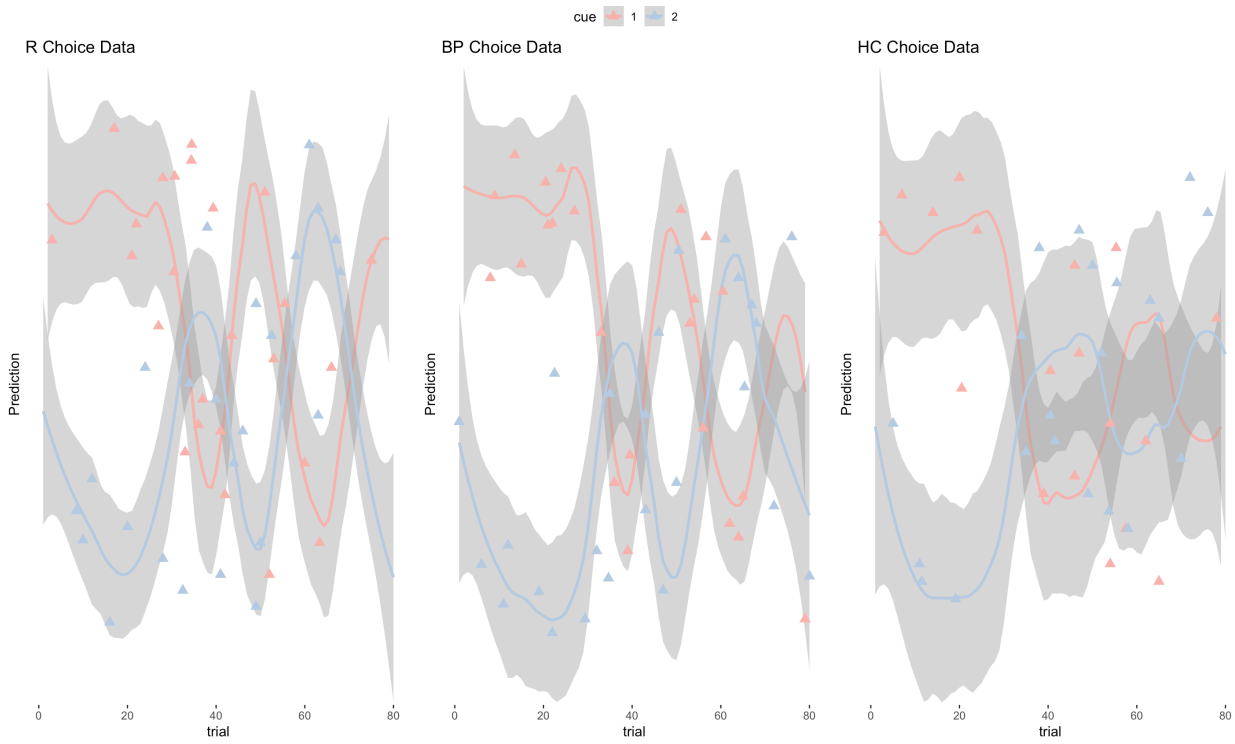


Figure 3.2: Visualization of participant trial-by-trial choices over the course of the task, by group and by cue. R – restricting, BP – binge-eating/purging, HC – healthy control.

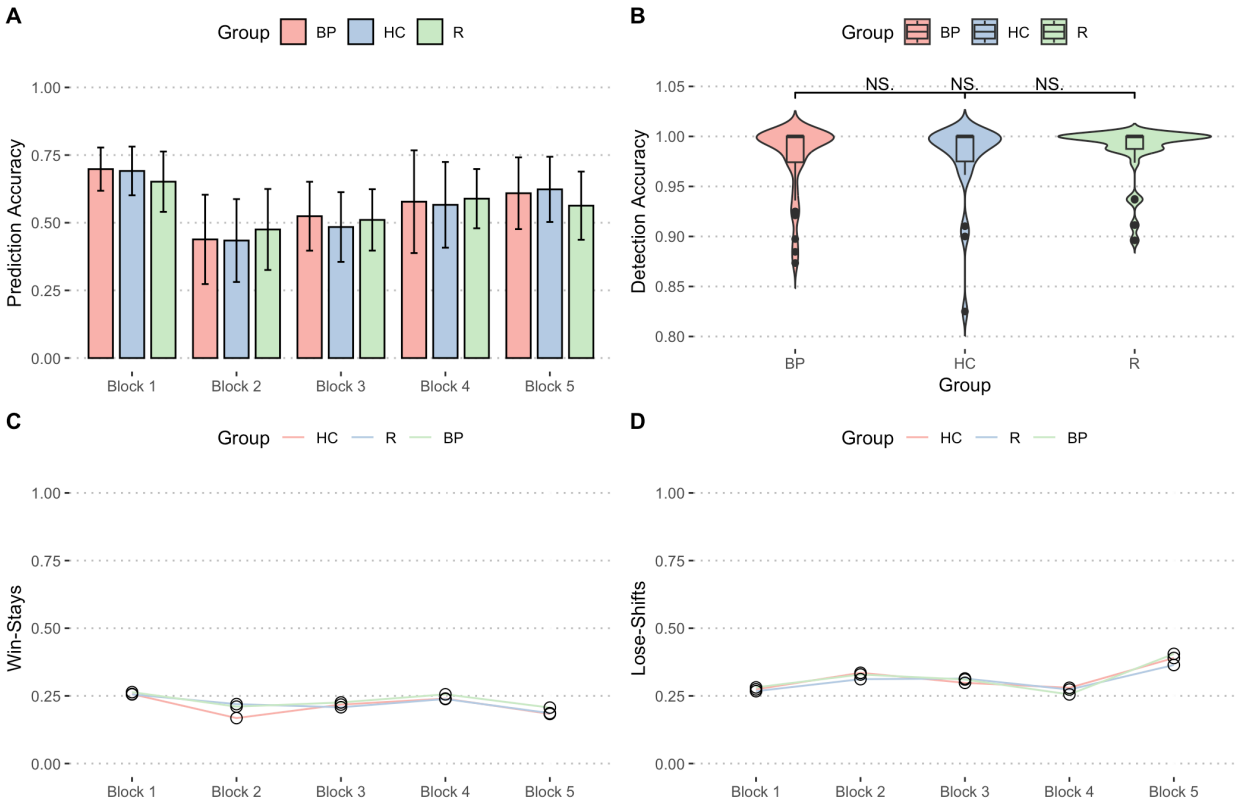


Figure 3.3: Raw choice plots from the Breathing Learning Task (BLT). (A) Accuracy of trial-by-trial predictions of whether a presented cue will be associated with breathing resistance, by group and by block. (B) Accuracy of trial-by-trial detection of breathing resistance, by group. (C) Win-stay probabilities, by group and by block. (D) Lose-shift probabilities, by block and by group. R – restricting, BP – binge-eating/purging, HC – healthy control.

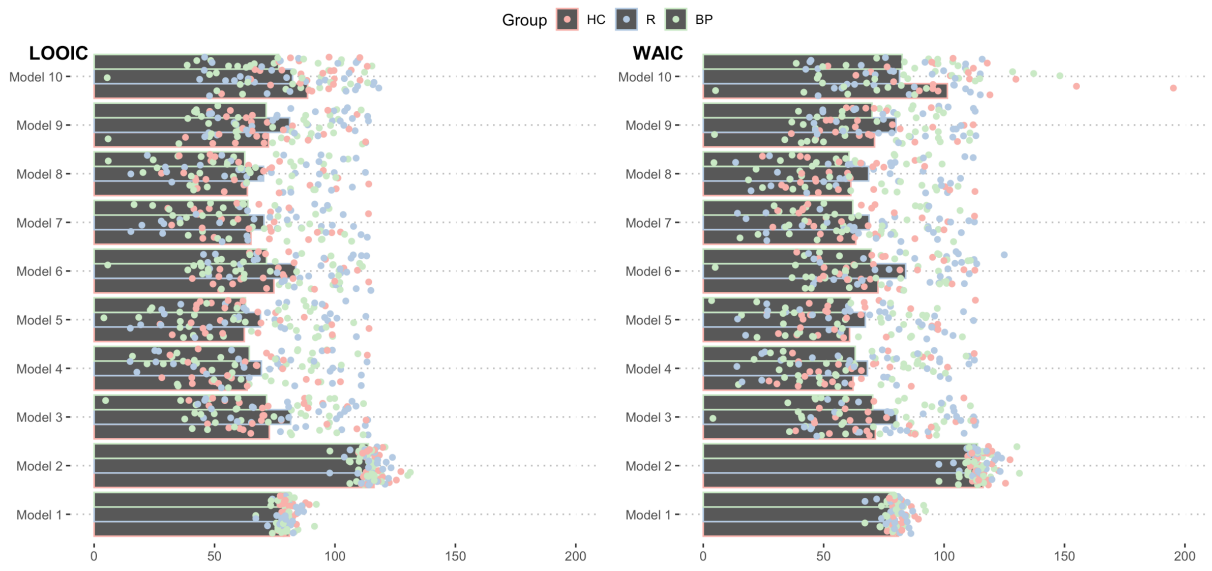
Table 3.2: Mixed effects linear regression results for prediction accuracy, win-stay, and lose-shift behavioral data. R – restricting, BP – binge-eating/purging, HC – healthy control.

	Accuracy				Win-Stay				Lose-Shift			
Predictors	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*
Intercept	0.82	0.18	0.47 – 1.16	**<0.001**	0.43	0.19	0.05 – 0.81	0.517	-0.39	0.19	-0.75 – -0.02	0.724
Block[2]	-1.66	0.25	-2.15 – -1.18	**<0.001**	-1.12	0.26	-1.63 – -0.61	**<0.001**	0.59	0.24	0.12 – 1.06	0.262
Block[3]	-1.34	0.25	-1.82 – -0.85	**<0.001**	-0.49	0.26	-1.00 – 0.02	1.000	0.23	0.24	-0.24 – 0.70	1.000
Block[4]	-0.81	0.25	-1.29 – -0.32	**0.022**	-0.21	0.26	-0.72 – 0.30	1.000	0.06	0.24	-0.41 – 0.53	1.000
Block[5]	-0.44	0.25	-0.92 – 0.05	1.000	-0.93	0.26	-1.43 – -0.42	**0.007**	1.12	0.24	0.65 – 1.59	**<0.001**
Group[R]	-0.27	0.24	-0.75 – 0.21	1.000	-0.07	0.27	-0.60 – 0.47	1.000	-0.08	0.26	-0.59 – 0.43	1.000
Group[BP]	0.06	0.23	-0.40 – 0.51	1.000	0.10	0.25	-0.40 – 0.60	1.000	0.10	0.24	-0.37 – 0.58	1.000
BMI	0.02	0.05	-0.07 – 0.12	1.000	-0.04	0.06	-0.16 – 0.09	1.000	-0.03	0.07	-0.16 – 0.10	1.000
Age	-0.02	0.04	-0.10 – 0.06	1.000	-0.00	0.05	-0.11 – 0.11	1.000	-0.16	0.05	-0.27 – -0.05	0.080
Anxiety	-0.06	0.05	-0.17 – 0.04	1.000	-0.00	0.07	-0.14 – 0.13	1.000	0.03	0.07	-0.11 – 0.17	1.000
Difficulty	0.07	0.05	-0.03 – 0.17	1.000	0.01	0.07	-0.12 – 0.15	1.000	-0.00	0.07	-0.14 – 0.13	1.000
Block[2]xGroup[R]	0.52	0.33	-0.14 – 1.17	1.000	0.69	0.35	0.00 – 1.37	0.947	-0.18	0.32	-0.81 – 0.46	1.000
Block[3]xGroup[R]	0.44	0.33	-0.21 – 1.10	1.000	-0.08	0.35	-0.76 – 0.61	1.000	0.21	0.32	-0.43 – 0.84	1.000
Block[4]xGroup[R]	0.43	0.33	-0.23 – 1.09	1.000	0.00	0.35	-0.68 – 0.69	1.000	0.01	0.32	-0.62 – 0.65	1.000
Block[5]xGroup[R]	-0.12	0.33	-0.78 – 0.53	1.000	0.06	0.35	-0.62 – 0.75	1.000	-0.23	0.32	-0.86 – 0.41	1.000
Block[2]xGroup[BP]	-0.02	0.32	-0.66 – 0.61	1.000	0.48	0.34	-0.19 – 1.14	1.000	-0.08	0.31	-0.70 – 0.53	1.000
Block[3]xGroup[BP]	0.20	0.32	-0.43 – 0.84	1.000	-0.02	0.34	-0.68 – 0.64	1.000	0.06	0.31	-0.55 – 0.68	1.000
Block[4]xGroup[BP]	0.01	0.32	-0.63 – 0.64	1.000	0.11	0.34	-0.55 – 0.77	1.000	-0.29	0.31	-0.91 – 0.32	1.000
Block[5]xGroup[BP]	-0.13	0.32	-0.77 – 0.50	1.000	0.21	0.34	-0.45 – 0.87	1.000	0.11	0.31	-0.51 – 0.72	1.000
Random Effects												
σ^2	0.76				0.83				0.71			
τ_{00}	0.00_{ID}				0.09_{ID}				0.12_{ID}			
ICC					0.10				0.15			
N	90_{ID}				90_{ID}				90_{ID}			
Observations	450				450				450			
Marginal R^2 / Conditional R^2	0.270 / NA				0.117 / 0.207				0.199 / 0.318			

Computational Modeling

Model comparison

Figure 3.4 shows the LOOIC and WAIC values for each tested model. Model 5 outperformed all other models when examining ΔELPD for both LOOIC and WAIC. All subsequent analyses using computational model parameter estimates were performed based on parameter estimates from Model 5.



model	elpd_diff	se_diff
model5	0.00000	0.00000
model8	-30.34358	29.67660
model4	-49.90994	31.17263
model7	-76.48166	39.99539
model9	-468.48599	48.28971
model3	-481.89307	51.32746
model6	-537.19061	61.61169
model1	-694.60023	129.65641
model10	-782.93058	72.54174
model2	-2296.71434	132.13021

model	elpd_diff	se_diff
model5	0.00000	0.00000
model8	-17.30132	31.14525
model4	-67.59298	30.93475
model7	-74.30364	39.55912
model9	-489.12679	48.22017
model3	-501.04745	51.25767
model6	-563.09707	68.78077
model1	-773.72887	129.97740
model10	-1403.94681	285.51048
model2	-2375.84763	132.43553

Figure 3.4: LOOIC (left) and WAIC (right) point estimates for each group. ΔELPD and standard errors are listed in the tables below each figure. Model 5 is the winning algorithm. R – restricting; BP – binge-eating/purging; HC – healthy control. elpd_diff – ΔELPD estimates; se_diff – ΔELPD standard errors.

Computational checks

Estimation. Estimated parameters from Model 5 were assessed for adequacy of chain convergence visually with trace plots. Trace plots were acceptable. $\hat{R}$ values were ~1.0 (range[0.99,1.00]) for all freely estimated parameters and ESS values were all >400, indicating good convergence.

Parameter recovery. Parameter recovery results for Model 5 are shown in Supplemental Figure 3.3 for all estimated parameters. Correlations between true and recovered subject-level

parameter estimates were all significantly and positively correlated (p<.05), suggesting adequate recovery.

Posterior predictive checks. Visual inspection of the simulated choice patterns over the course of the task showed adequate similarity to the true choice data (Supplemental Figures 3.4-3.5). Additionally, no group effects met significance when testing accuracy, win-stay, or lose-shift behavior, as was the case for the true data (Supplemental Table 3.3).

Analysis of model parameters and self-report data

Transformed posterior distributions of group means for parameter estimates, including learning rate, counterfactual learning rate, and inverse temperature, are displayed in Figure 3.5. Bayesian analysis of model parameters showed strong evidence of higher learning rates and counterfactual learning rates in the R group compared to the BP group. Posterior distributions for the difference between R and BP learning rates showed 2% of the HDI fell within ROPE, and with 97% of the probability of direction falling above 0. Similarly, posterior distributions for the difference between R and BP counterfactual learning rates showed <1% of the HDI fell within ROPE, and with 99% of the probability of direction falling above 0. Additionally, the R group showed strong evidence of higher counterfactual learning rates compared to the HC group, with 2% of the HDI of the difference posterior distribution falling within ROPE and 98% of the probability of direction falling above 0. Finally, the BP group demonstrated greater tendencies to exploit cue valuation in their action selection process compared to the R group. This was revealed through higher inverse temperature estimates in the BP group, with 3% of the HDI falling into ROPE and 97% of the probability of direction falling below 0. Figure 3.6 displays

visualizations and metrics of difference posterior distributions for each freely estimated parameter.

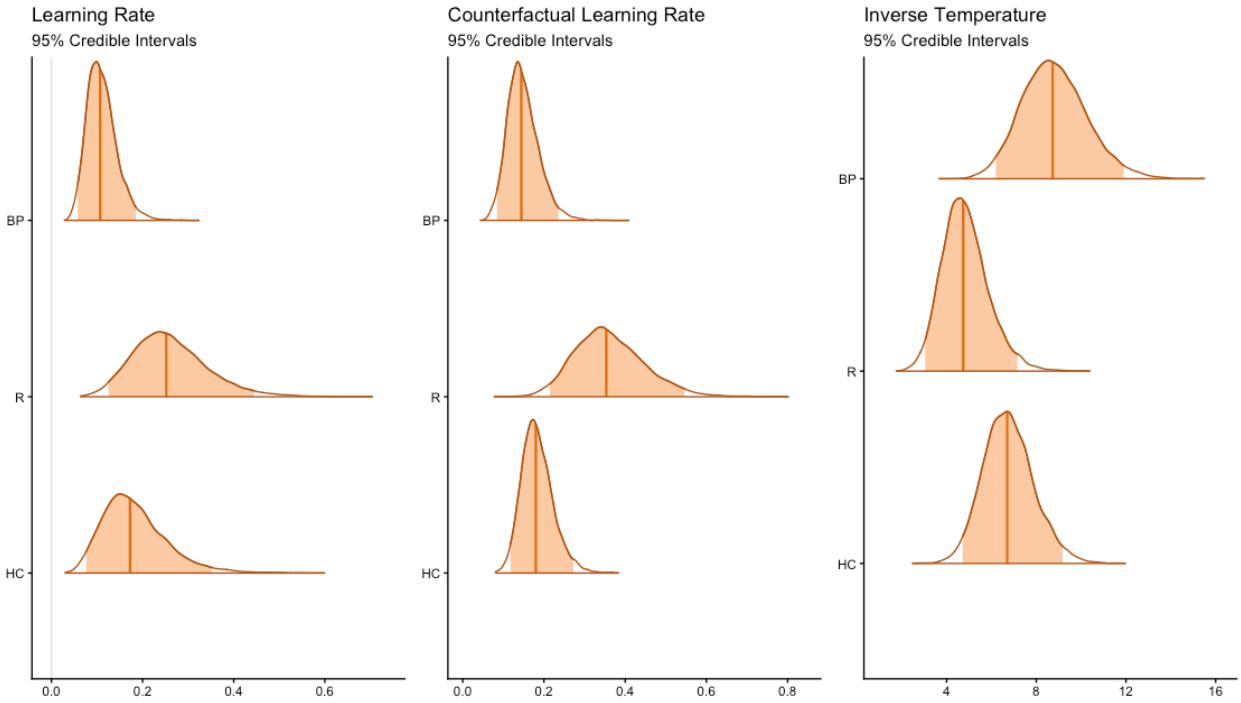
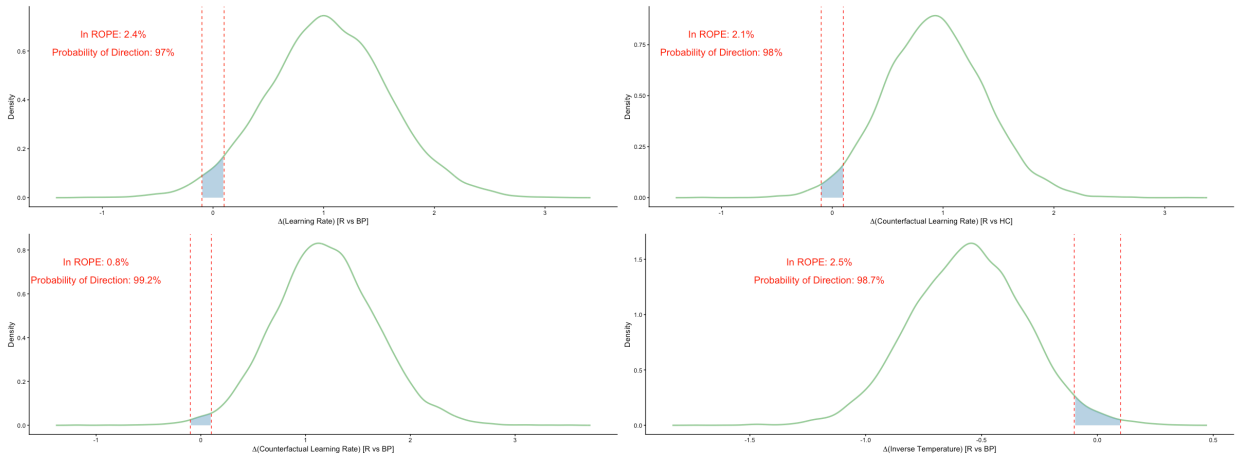


Figure 3.5: Transformed group-level posterior distributions for positive (left) and negative (right) learning rates. Credible intervals are 95% highest density intervals (HDI) and are represented by the shaded areas of the distributions. R – restricting; BP – binge-eating/purging; HC – healthy control.



Parameter	Median	Mean	SD	CI	CI_low	CI_high	pd	ROPE_CI	ROPE_low	ROPE_high	ROPE_Percentage
Delta(Alpha)[R vs HC]	0.4762535	0.4786431	0.6496339	0.95	-0.77853800	1.7591800	0.7723333	1	-0.1	0.1	0.09541667
Delta(Alpha)[BP vs HC]	-0.5615275	-0.5601312	0.5781353	0.95	-1.65735000	0.6104510	0.8355833	1	-0.1	0.1	0.08508333
Delta(Alpha)[R vs BP]	1.0337600	1.0387743	0.5513438	0.95	-0.03132840	2.1433300	0.9704167	1	-0.1	0.1	0.02433333
Delta(Counterfactual Alpha)[R vs HC]	0.9202030	0.9257218	0.4560162	0.95	-0.00206605	1.7947900	0.9805000	1	-0.1	0.1	0.02141667
Delta(Counterfactual Alpha)[BP vs HC]	-0.2605045	-0.2643416	0.3992631	0.95	-1.05078000	0.5070400	0.7450833	1	-0.1	0.1	0.16691667
Delta(Counterfactual Alpha)[R vs BP]	1.1783950	1.1900634	0.4885914	0.95	0.24413700	2.1620300	0.9924167	1	-0.1	0.1	0.00825000
Delta(Beta)[R vs HC]	-0.2881430	-0.2878821	0.2249005	0.95	-0.71399700	0.1622100	0.9039167	1	-0.1	0.1	0.15633333
Delta(Beta)[BP vs HC]	0.2663745	0.2709572	0.2438259	0.95	-0.20578300	0.7574990	0.8715000	1	-0.1	0.1	0.17808333
Delta(Beta)[R vs BP]	-0.5572935	-0.5588393	0.2474337	0.95	-1.06295000	-0.0974503	0.9865833	1	-0.1	0.1	0.02516667

Figure 3.6: Posterior distributions for difference parameters (top), showing differences between groups in group means for learning rates, counterfactual learning rates, and inverse temperature parameters. Shaded regions represent percentage of the 95% highest density interval (HDI) in the region of practical equivalence (ROPE) and red dotted lines represent ROPE. Tables display posterior distribution metrics, including percentage of the HDI that falls within ROPE and the probability of direction (pd). CI refers to 95% HDI. R – restricting; BP – binge-eating/purging; HC – healthy control.

Linear models investigating associations between eating disorder self-report measures (EPSI) and subject-level computational model parameters were non-significant following correction. Significant exploratory spearman correlations between model parameters and all self-report measures, by group, are shown in Supplemental Figures 3.6 – 3.8. A post-hoc analysis investigating relationships between post-task VAS certainty ratings and the difference between learning rates and counterfactual learning rates ($\Delta_{\alpha-\alpha'}$) revealed a significant omnibus test ($F_{9,70}$ = 2.36, p = .02), main effect of post-task VAS detection certainty for the R group (b = -0.17, se = 0.05, p = .002), and significant group-by-detection certainty interaction effects for the BP group (b = 0.23, se = 0.07, p = .004) and the HC group (b = 0.18, se = 0.07, p = .03) after correction. This indicates that as difference scores become more negative, meaning that counterfactual learning rates become higher than learning rates, certainty in one's ability to

detect breathing resistance increased for the R group (Supplemental Figure 3.9). This relationship was in the opposite direction for the BP and HC groups. No significant main or interaction effects were observed when conducting the model with post-task VAS prediction certainty ratings.

Analysis of rsfMRI data

Group-level ROI-to-ROI analyses revealed no significant group differences. Moreover, interaction effects between group and subject-level learning rates were non-significant. Exploratory seed-to-voxel GLM analyses, with insula selected as the seed region, revealed a significant interaction between group and learning rates. Specifically, a significant interaction effect ($F_{2,83} = 14.7$; $p_{FDR\text{-}corrected} < .01$) was observed when associating learning rates with connectivity between left insula (IC) and right anterior prefrontal cortex (aPFC; cluster size = 51 voxels, $p_{FDR\text{-}corrected} = .01$), shown in Figure 3.7. Post-hoc analyses revealed that the R group showed more negative associations between subject-level learning rates and IC-aPFC functional connectivity compared to the HC group ($t_{83} = 5.1$, $p_{FDR\text{-}corrected} < .01$). Specifically, as learning rates increased in the R group, IC-aPFC functional connectivity decreased, whereas the opposite was observed in the HC group.

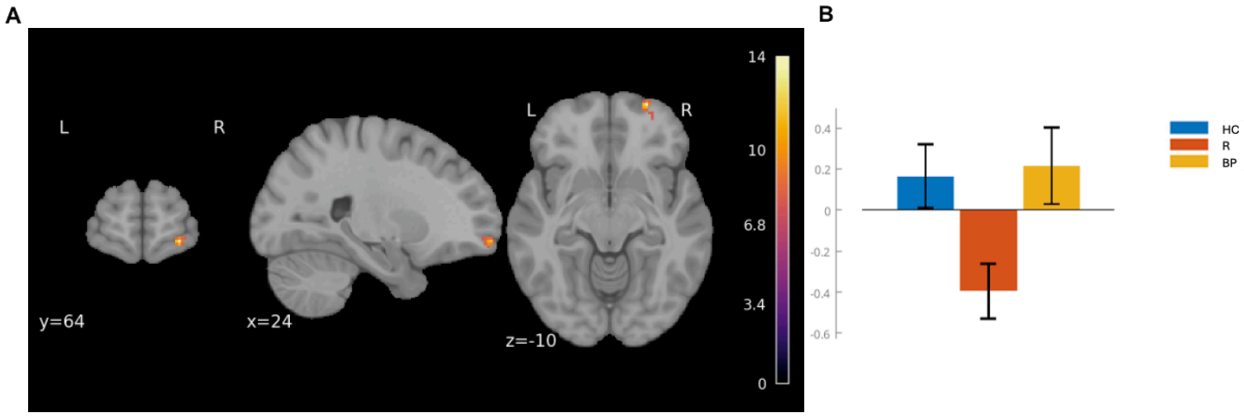


Figure 3.7: Results from exploratory seed-to-voxel GLMs examining group differences in associations between connectivity and learning rates. (A) A significant group-by-learning rate interaction term was present for connectivity between left insula and right anterior prefrontal cortex (center coordinates: x = 26, y = 64, z = -12). (B) Plotted weights of the interaction effect. R – restricting; BP – binge-eating/purging; HC – healthy control.

DISCUSSION

This study examined interoceptive learning and resting-state functional connectivity in women with eating disorders, who were characterized by restricting (AN restricting subtype; R) and binge-eating/purging phenotypes (AN binge-eating/purging subtype and BN; BP), compared to a sample of healthy control women (HC). Interoceptive learning was operationalized by behavioral performance during a reversal learning task outfitted with probabilistic breathing resistance outcomes (i.e., breathing learning task; BLT), and associated with functional connectivity in key interoceptive neural regions during a resting-state functional magnetic resonance imaging (rsfMRI) scan, including bilateral insula, amygdala, ventral striatum, and prefrontal cortex. Overall, the R group showed faster learning rates and counterfactual learning rates, as well as lower inverse temperature parameter estimates, compared to the BP group, which was not consistent with our hypotheses. Compared to the HC group, the R group showed augmented counterfactual learning rates, but no other differences were observed in the R and BP groups when compared to HC. And while no significant group differences in functional

connectivity between regions-of-interest (ROI) emerged, exploratory analyses revealed group differences in associations between learning rates and functional connectivity between left insula (IC) and right anterior prefrontal cortex (aPFC). Most notably, these results highlight distinct latent processes involved in integration of internal sensory information into higher-order cognitive functions between primarily restricting and binge-eating/purging eating disorder phenotypes.

Although raw choice behavior during the BLT did not differ between groups, relatively distinct patterns of behavior emerged when examining latent cognitive processes via estimated computational model parameters. The R group appeared to be characterized by faster updating of cue valuation (i.e., learning rates), faster updating of unseen cue valuation (i.e., counterfactual learning rates), and noisier or more exploratory action selection (i.e., inverse temperature parameter). The BP group showed the opposite, compared to the R group: slower updating of cue valuation, slower updating of unseen cue valuation, and more exploitative action selection. Faster updating of cue valuation in the R group, and thus greater sensitivity to prediction errors (or the difference between outcomes and expected outcomes), is consistent with prior work showing heightened prediction error signaling in the caudate and insula within a sample of adolescents with AN (DeGuzman et al., 2017). However, evidence of faster learning rates in individuals with AN have largely been described in the context of learning from negative feedback more specifically (Bernardoni et al., 2018, 2021; Filoteo et al., 2014; Wierenga, Bischoff-Grethe, et al., 2025), although this is not necessarily consistent across the literature (Shott et al., 2012; Wierenga et al., 2022). Sensitivity to valence in latent learning processes has also been described in individuals diagnosed with mood and anxiety disorders and may reflect a transdiagnostic feature of negative affective bias (Pike & Robinson, 2022). Although the winning

computational model in this study specified a valence-agnostic learning rate, the task itself was designed as an aversive interoceptive learning task, with "positive" outcomes defined as the absence of breathing resistance, or rather, the absence of punishment. Within this framework, it is possible that learning rates possessed an inherently negative valence, as participants were asked to learn about cues that predict punishment and avoidance of punishment. This would be consistent with prior literature in eating disorders (C. S. Brown et al., 2024; Radzikowska et al., 2025), although it is speculative and in need of further exploration.

Interestingly, counterfactual learning rates in the R group were also faster compared to both BP and HC groups. In this case, counterfactual learning refers to how an individual learns from both direct experience via cues that are presented to them on each trial, as well as hypothetical or fictive outcomes associated with the unseen cue (Boorman et al., 2011; Fischer & Ullsperger, 2013; Hayden et al., 2009; Lohrenz et al., 2007). Consequently, a counterfactual learning algorithm is somewhat linked to model-based reinforcement learning, in that it assumes use of mental representations of the environment (i.e., knowledge of the 80/20 probability structure) and mental simulations of possible outcomes for an unseen cue (J. P. O'Doherty et al., 2015; Wan Lee et al., 2014); for example, if the presented cue on trial *t* results in a breathing resistance, the unseen cue would hypothetically result in no breathing resistance. Counterfactual learning has been suggested as a more computationally expensive or demanding cognitive process compared to standard temporal difference learning and has been linked to neural regions associated with cognitive control and punishment learning, including dorsolateral prefrontal cortex, insula, and right ventral striatum (Boorman et al., 2011; Koechlin, 2014; Palminteri et al., 2012; Santo-Angles et al., 2021). Despite the greater complexity in cognitive processing inherent in counterfactual learning, it has been shown to align with developmental trajectories from

adolescence to adulthood (Palminteri et al., 2016), similar to transitions from model-free to model-based learning (Decker et al., 2016).

Augmented counterfactual learning rates in the R group therefore represent faster rates of updating for fictive valuations compared to the remaining groups and may suggest higher perceptions of volatility in the environment, as well as greater efforts to detect change in the environment (i.e., reversals; Zhang et al., 2015). Faster counterfactual learning also implies that the R group made strong inferences about unseen outcomes and were more sensitive to fictive prediction errors; while this could be suggestive of a model-based learning strategy, it also invariably influences cue valuation to be biased based on an expected or mentally simulated prediction error. In the context of interoceptive learning, this type of cognitive processing may result in top-down expectations about internal sensory experiences and their cues influencing how an individual learns about their internal milieu. For example, when confronted with a decision to eat or not to eat, expectations about bloating being associated with food consumption is influenced not just by the information gained from the experience itself, but also by the hypothetical expectation of whether the unchosen action would result in bloating or not. This hypothesis is strengthened by our post hoc finding that, for the R group, the difference between learning rates and counterfactual learning rates was negatively associated with confidence in the ability to accurately detect breathing resistances. This relationship was opposite to that of the HC and BP groups and suggests that underweighting of learning from direct experience (i.e., learning rates), relative to fictive experience (i.e., counterfactual learning rates), is paradoxically linked to greater confidence in interoceptive detection. It is therefore possible that individuals in the R group more heavily relied on mental simulations or expectations of sensory outcomes to feel less uncertain about their bodily state. Instead of relying solely on experienced sensory input,

individuals with restricting phenotypes may also rely on expectations tied to outcomes that they did not directly experience (similar to the concept of hyper-precise Bayesian priors), in line with interoceptive Bayesian inference theories (Petzschner et al., 2021).

Adding some amount of complexity to the behavioral picture, inverse temperature parameters for the R group were attenuated when compared to the BP group, a finding that is somewhat consistent with a prior study comparing women with AN restricting subtype to control women (Wierenga et al., 2022). However, because posterior distributions of the group means were all relatively high, the impact of this group difference should not be overstated, as inverse temperature parameters past a certain threshold will begin to result in action selection that is more deterministic, favoring the higher value in an absolute sense (Sutton & Barto, 1998). This is particularly true for cue values that are far apart; the closer they are, and thus, the more ambiguity in optimal choice, the greater possibility for stochasticity in choice behavior, depending on the parameterization of the inverse temperatures. Nevertheless, this result suggests greater stochasticity in choice behavior for the R group relative to the BP group. Attenuated inverse temperature parameters could be the result of several factors, including greater exploratory behavior linked to uncertainty, general noisiness in choice selection, or inattentiveness (Li et al., 2024; Wilson et al., 2014). It is possible that greater exploratory behavior in the R group could be related to efforts to resolve uncertainty, particularly if they perceived the task environment to be volatile or dynamic. Such an interpretation could be plausible given the high learning and counterfactual learning rates seen in the R group, as learning rates tend to increase when environments become less stable (Behrens et al., 2007; Browning et al., 2015; McGuire et al., 2014). This finding also indicates that although participants show higher sensitivity to prediction errors, they do not necessarily overly exploit

their cue valuation to make choices. Meanwhile, the BP group showed the opposite, where they displayed less sensitivity to trial-by-trial prediction errors, but more deterministic, exploitative choice selection, relying more heavily on historical cue valuations. This is in contrast to more recent theories about the explore-exploit tradeoff and its relevance to eating disorders (Hagan et al., 2024), particularly in combination with evidence of non-optimal use of exploration related to excessive switching behavior in individuals with binge eating disorder (Reiter et al., 2017). However, given the high values for group-level inverse temperature estimates and differences in task design reported here, interpretations should be made with caution.

Counter to our hypotheses, we did not find any group differences in ROI-to-ROI connectivity in our *a priori* interoceptive network. However, our exploratory seed-to-voxel analyses, specifying insula as our seed region, showed group differences in associations between IC-aPFC functional connectivity and learning rate. Specifically, as learning rate increased, IC-aPFC functional connectivity decreased in the R group, indicating that higher learning rates in this sample are linked to reduced neural synchrony in these regions. Lesion studies examining aPFC function have linked this region to higher-order cognitive processes, including cognitive control (Huettel, 2006), meta-cognitive awareness (Stuss & Levine, 2002; Wheeler et al., 1997), and meta-awareness of emotional states (Ochsner et al., 2004; Ochsner & Gross, 2005). Both lateral prefrontal cortex regions and anterior insula are hubs within the salience network, noted to be involved in detecting salient stimuli and initiating attentional control processes (Dosenbach et al., 2007; Menon & Uddin, 2010). Interestingly, it has been hypothesized that streams of information processing within the salience network involve integration of bottom-up, sensory information through the anterior insula to lateral prefrontal regions to elicit top-down attentional control, similar to theories of neuroanatomical networks involved in interoceptive learning

(Fermin et al., 2022). Prior work has also shown that functional connectivity between anterior insula and anterior prefrontal cortex is involved in attentional and cognitive control (Deen et al., 2011; Simmons et al., 2013; Touroutoglou et al., 2012). Within the context of an aversive interoceptive learning task, it is surprising that higher learning rates would be associated with attenuated connectivity in these regions; one possibility is that higher learning rates in individuals with attenuated connectivity in regions linked to integration of relevant sensory stimuli and allotment of control/attention may be understood as reflecting an imbalance or miscommunication between bottom-up and top-down systems. If communication with cognitive control regions is weakened or disrupted in individuals with restricting eating disorder phenotypes, weighting learning processes so that they are heavily influenced by immediate sensory experience or feedback could potentially contribute to certain forms of eating disorder symptomatology that become entrenched (e.g., dietary restriction, body image disturbance). This is speculative, however, and should be replicated in future work.

The current study supports dissociable latent interoceptive learning processes between the R and BP groups, which may be an important step towards understanding core mechanisms involved in how individuals with eating disorders experience their internal bodily states. In this context, individuals specifically presenting with a restricting phenotype appeared to learn about sensory experiences both from direct experience, but also from expected or assumed outcomes. Within the framework of interoceptive predictive coding, over-reliance on top-down processes that assume a particular model of the world may lead to chronic negative beliefs about physiological states and may help explain prior findings highlighting difficulties discriminating between anticipated and experienced sensory stimuli (Khalsa et al., 2018). Clinical interventions may be improved by focusing on learning from direct experience to reduce bias introduced by

hypothetical outcomes and/or changing existing mental models. Alternatively, individuals presenting with binge-eating/purging phenotypes may see improved outcomes by focusing on faster integration of immediate sensory feedback.

To our knowledge, this is the first study to examine latent mechanisms involved in aversive interoceptive learning in individuals with eating disorders. There were, however, several limitations to this work. Although the sample was well characterized, sample sizes for each group were small, especially for the HC group, which may have affected power in some of our analyses. Additionally, combining individuals with BN and AN-BP may have masked effects specific to these diagnoses; future work should investigate interoceptive learning by diagnosis, as the underweight status required for an AN-BP diagnosis might be associated with distinct interoceptive processing or experience. Finally, a major limitation of this study is the lack of task-based fMRI to understand neural functioning during the BLT. We examined associations between computational model parameters and rsfMRI only, which can lead to spurious conclusions that are more difficult to interpret. A next step would be to acquire task-based fMRI with the BLT in order to make conclusions about neural activity involved in interoceptive learning processes.

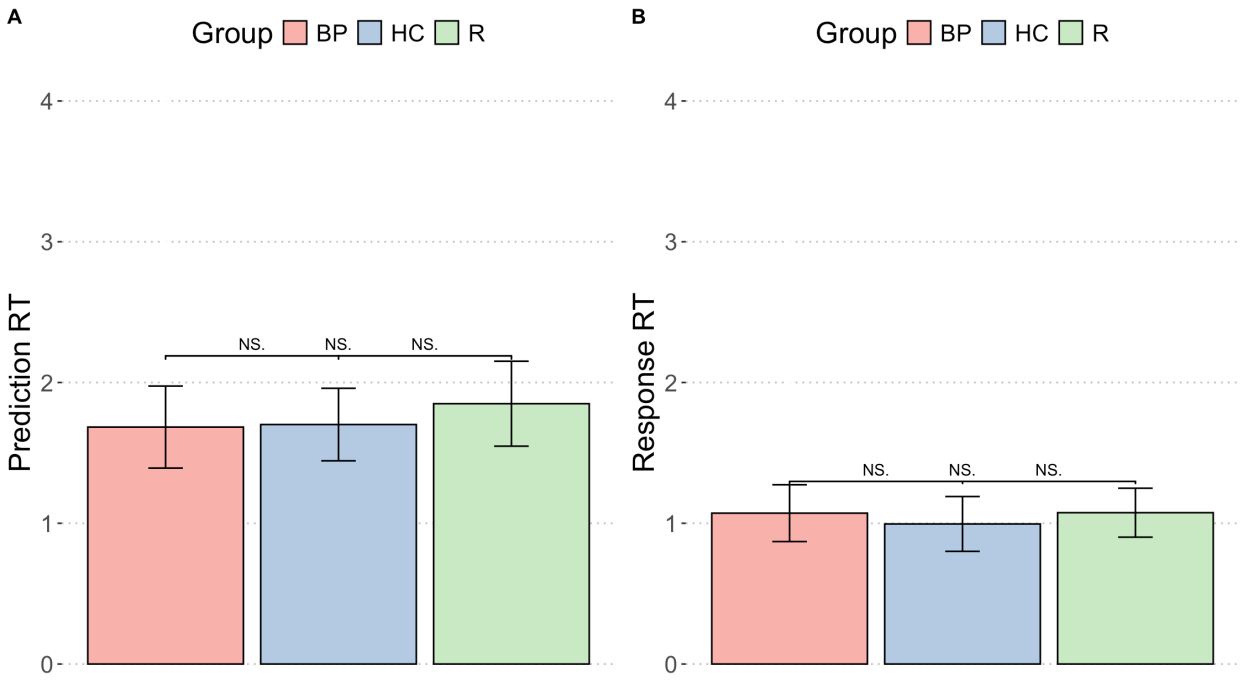
ACKNOWLEDGEMENTS

Chapter 3 contains unpublished material. I would like to acknowledge and thank: Katherine Gandee, Megan Martinho, Amanda Bischoff-Grethe, PhD, and Christina Wierenga, PhD. I would also like to thank Angeline Krueger, the UCSD Eating Disorders Center for Treatment and Research staff and patients, Xiaosi Gu, PhD, and Laura Berner, PhD for their

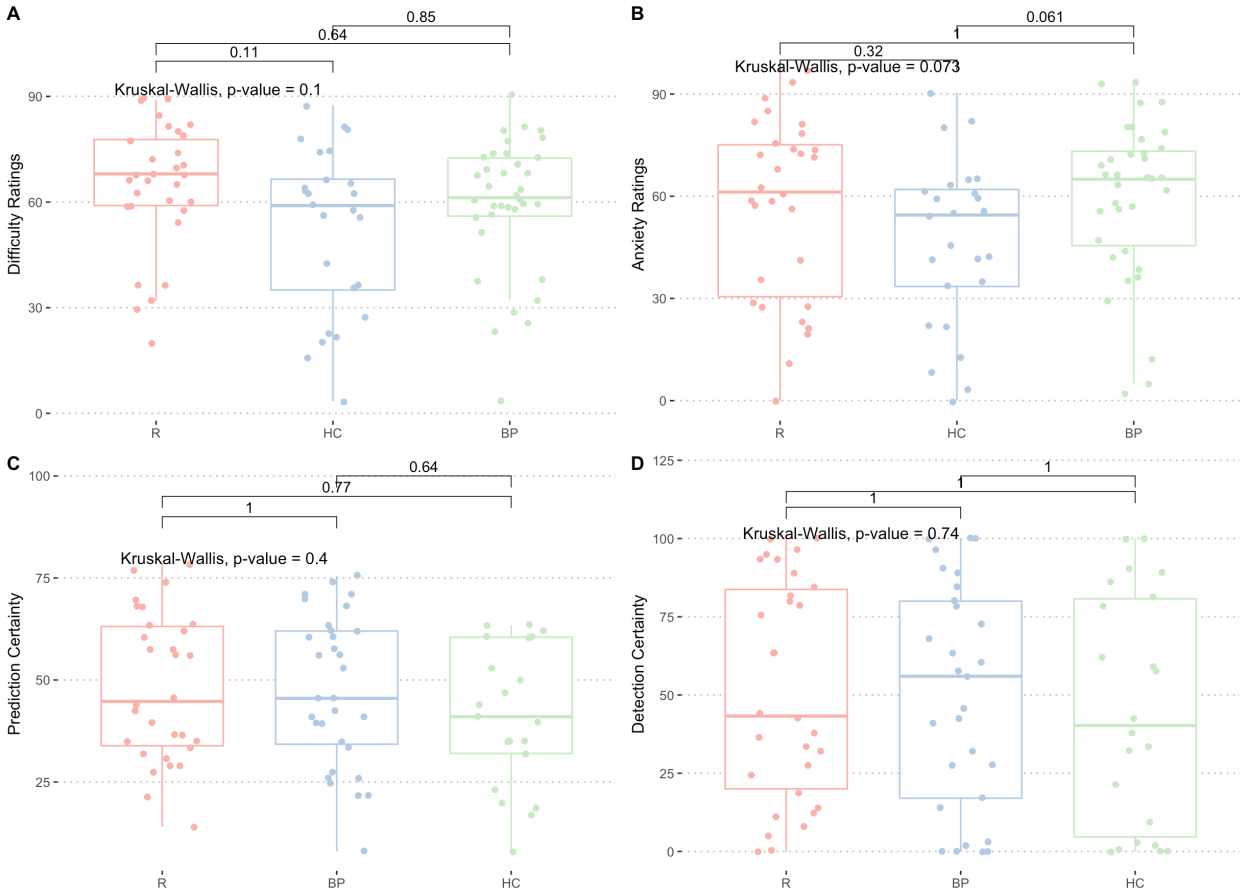
contributions to this project. This project was supported by National Institutes of Health Grant No. F31MH133362 (awarded to CSB) and Grant No. R01MH125880 (awarded to CEW).

Supplemental Table 3.1: Descriptive statistics by group for participants excluded from analyses compared to participants included in analyses. R – restricting; BP – binge-eating/purging; HC – healthy controls. EPSI – Eating Pathology Symptom Inventory; MAIA – Multidimensional Assessment of Interoceptive Awareness; ATQ – Adult Temperament Questionnaire (Effortful Control subscale); WASI – Weschler Abbreviated Scale of Intelligence.

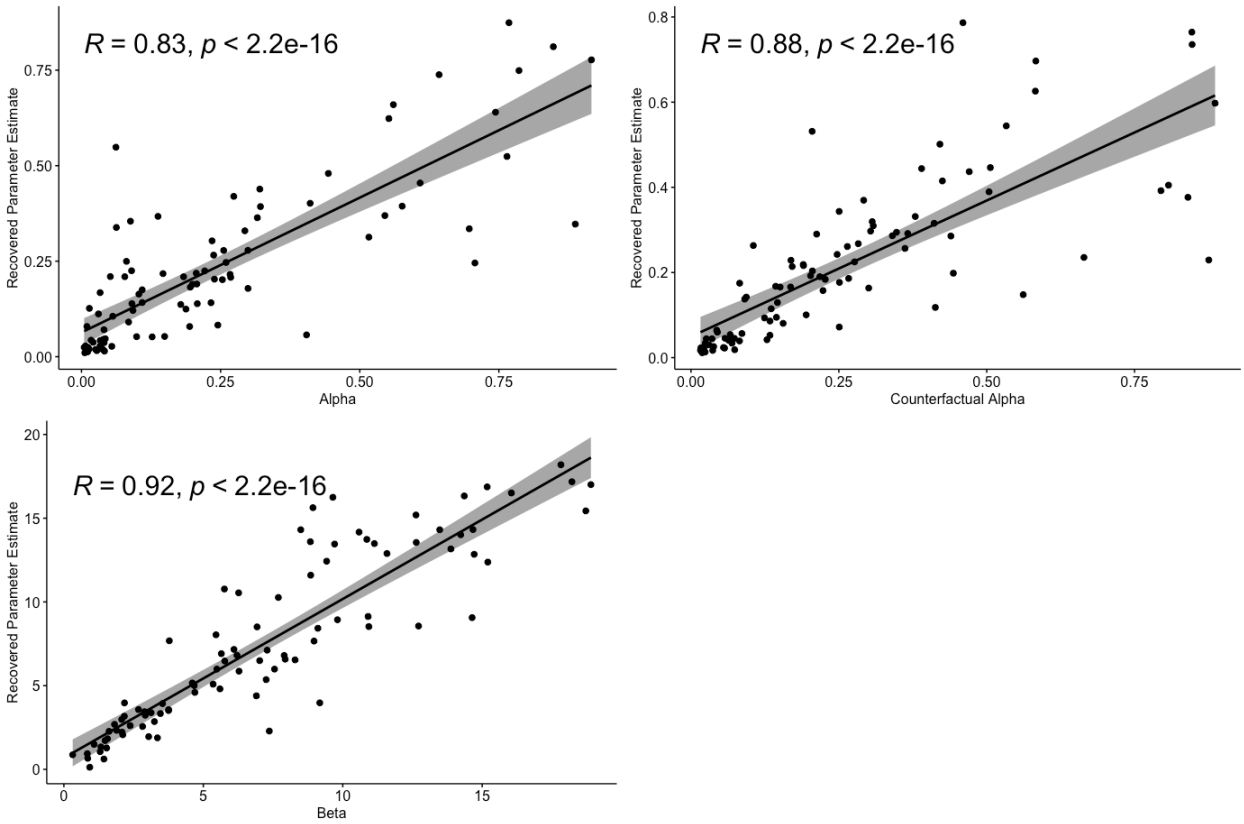
Group	Excluded			Included			
Variable	N	Mean	SD	N	Mean	SD	Test
Education	9	15	3	92	15	2.1	F=0.451
Race	9			92			X2=2
... Asian	1	11%		19	21%		
... Black	0	0%		2	2%		
... More than one race	0	0%		9	10%		
... White	8	89%		62	67%		
Ethnicity	9			92			X2=0.675
... Hispanic or Latino	0	0%		15	16%		
... Not Hispanic or Latino	9	100%		77	84%		
Age	9	28	7.8	92	24	4.9	F=5.384**
BMI	9	20	3.3	92	22	3.4	F=1.613
MAIA Noticing	9	2.9	0.61	90	3	0.93	F=0.105
MAIA Not Distracting	9	2	0.52	90	1.9	1.1	F=0.03
MAIA Not Worrying	9	2.6	0.72	90	2.6	1	F=0.046
MAIA Attention Regulation	9	2.1	0.67	90	2.4	0.88	F=0.895
MAIA Emotional Awareness	9	3.1	0.47	90	3	1	F=0.073
MAIA Self Regulation	3	2.1	0.58	84	2.4	1.1	F=0.278
MAIA Body Listening	3	2	1	84	2.1	1.2	F=0.019
MAIA Body Trust	9	2	1.2	92	2.4	1.5	F=0.564
EPSI Body Dissatisfaction	9	19	6.6	92	15	8.4	F=1.24
EPSI Binge Eating	9	7.3	5.9	92	9.3	8.9	F=0.44
EPSI Cognitive Restraint	9	8.8	3.5	92	6.5	3.6	F=3.242*
EPSI Purging	9	4.2	6.9	92	2.6	4.2	F=1.131
EPSI Restricting	9	11	5.1	92	8	5.9	F=1.86
EPSI Excessive Exercise	9	10	6.1	92	6.9	6	F=2.786*
EPSI Obesity Attitude	9	4.2	6.3	92	5.8	5.8	F=0.631
EPSI Muscle Building	9	3.1	3.3	92	3.4	4	F=0.041
ATQ Activation Control	9	4.4	1.1	92	4.4	1.1	F=0.004
ATQ Attentional Control	9	3.8	0.84	92	3.8	1.1	F=0.002
ATQ Inhibitory Control	9	4.4	0.64	92	4.4	0.99	F=0
WASI IQ Score	8	106	7.7	84	116	11	F=5.62**
Group	9			92			X2=2.647
... HC	1	11%		25	27%		
... BP	6	67%		36	39%		
... R	2	22%		31	34%		
Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01							



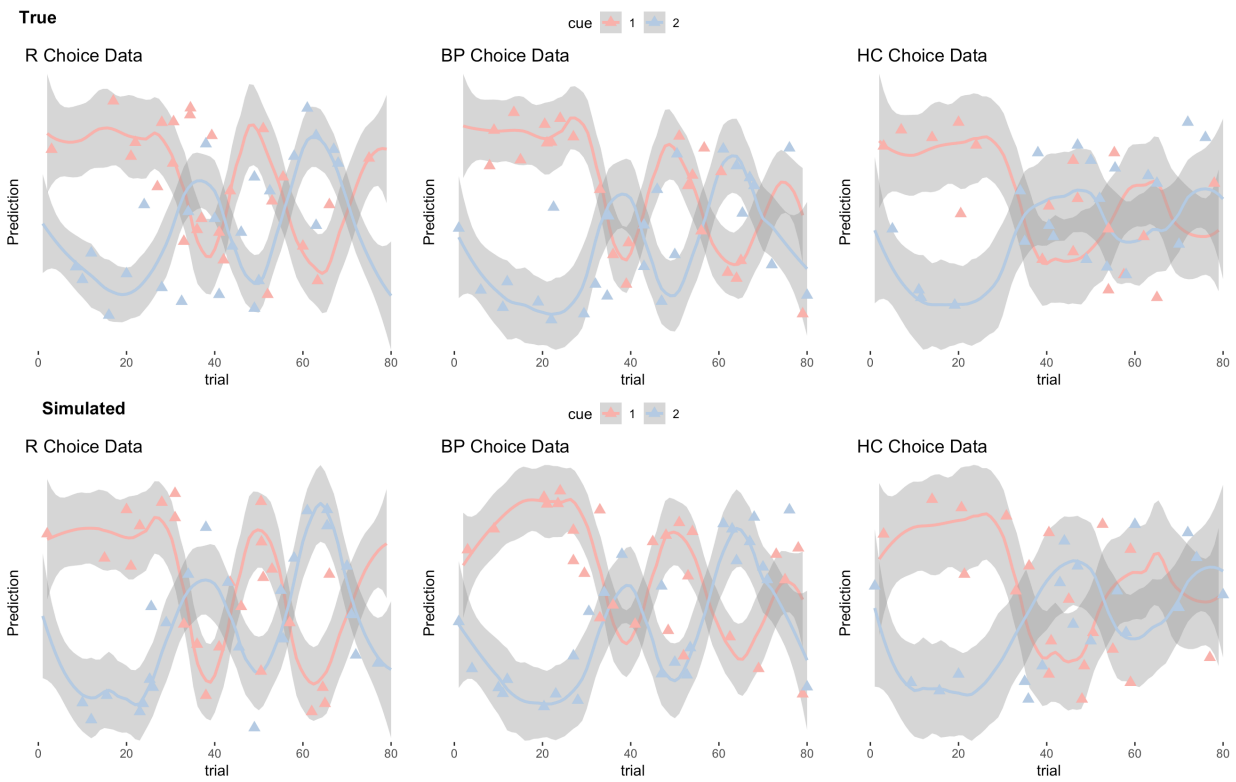
Supplemental Figure 3.1: Reaction times (RT) per group during the Breathing Learning Task (BLT). No group differences were observed. (A) RTs for predicting whether a presented cue would result in a breathing resistance. (B) RTs for detection responses (i.e., whether a breathing resistance occurred). N.S. – not significant. R – restricting; BP – binge-eating/purging; HC – healthy controls.



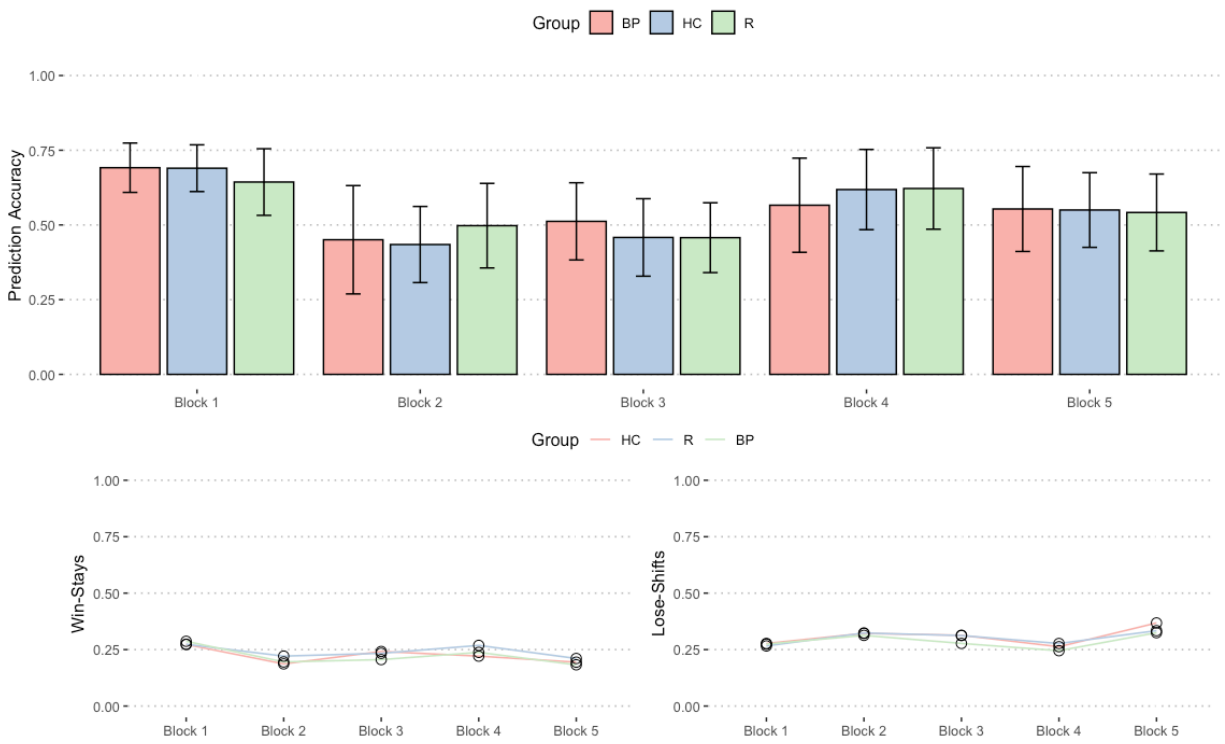
Supplemental Figure 3.2: Boxplots of post-task visual analogue scale (VAS) ratings for perceived difficulty (A), anxiety (B), prediction certainty (C), and detection certainty (D), by group. No significant group differences were observed. R – restricting; BP – binge-eating/purging; HC – healthy controls.



Supplemental Figure 3.3: Spearman correlations between recovered subject-level parameter estimates and true parameter estimates. Alpha – learning rate; Counterfactual alpha – counterfactual learning rate; Beta – inverse temperature.



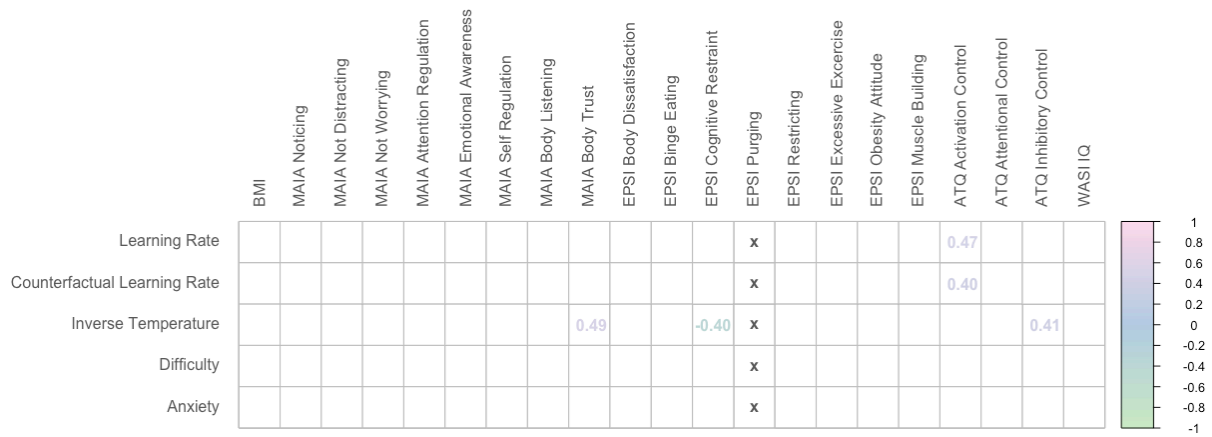
Supplemental Figure 3.4: Visualization of participant trial-by-trial choices over the course of the task, by group and by cue. (Top) True choice data. (Bottom) Simulated choice data. R – restricting, BP – binge-eating/purging, HC – healthy control.



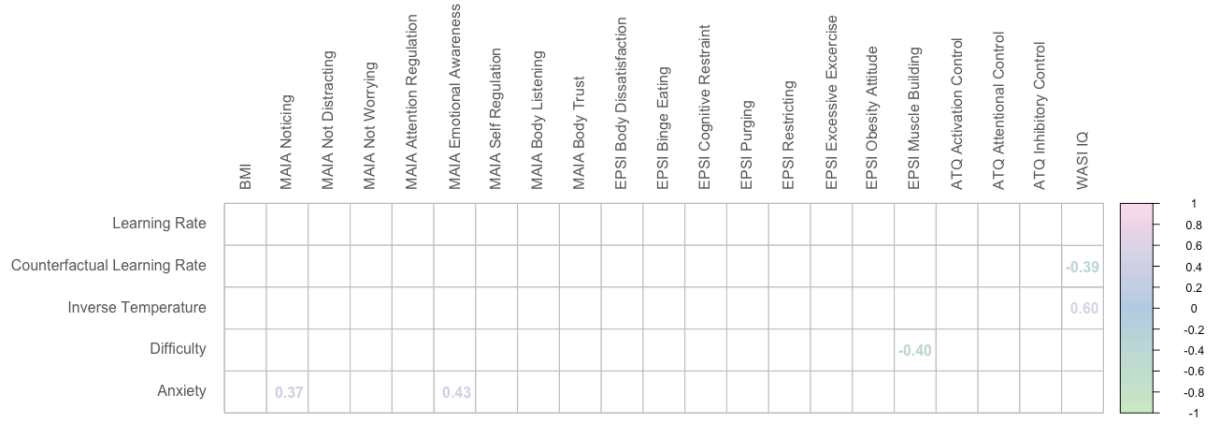
Supplemental Figure 3.5: Choice plots from simulated data. (Top) Accuracy of trial-by-trial predictions of whether a presented cue will be associated with breathing resistance, by group and by block. (Bottom) Win-stay and lose-shift probabilities, by group and by block. R – restricting, BP – binge-eating/purging, HC – healthy control.

Supplemental Table 3.2: Mixed effects linear regression results for prediction accuracy, win-stay, and lose-shift simulated behavioral data. R – restricting, BP – binge-eating/purging, HC – healthy control.

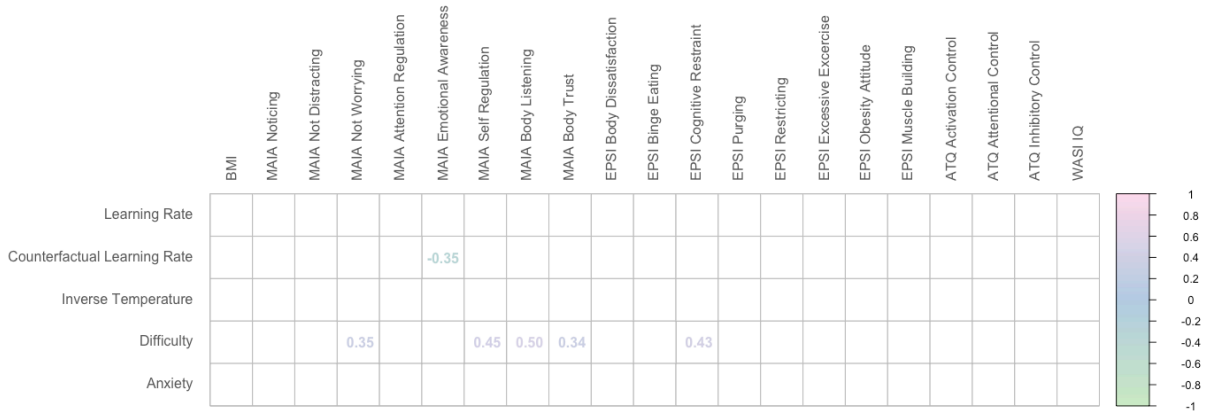
	Accuracy				**Win-Stay**				**Lose-Shift**			
Predictors	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*
Intercept	0.87	0.17	0.53 – 1.21	**<0.001**	0.56	0.19	0.19 – 0.94	0.063	-0.20	0.20	-0.59 – 0.19	1.000
Block[2]	-1.66	0.23	-2.12 – -1.21	**<0.001**	-1.06	0.27	-1.58 – -0.53	**0.002**	0.47	0.28	-0.08 – 1.01	1.000
Block[3]	-1.51	0.23	-1.97 – -1.06	**<0.001**	-0.38	0.27	-0.91 – 0.14	1.000	0.38	0.28	-0.17 – 0.92	1.000
Block[4]	-0.47	0.23	-0.92 – -0.01	0.851	-0.63	0.27	-1.16 – -0.11	0.339	-0.15	0.28	-0.69 – 0.39	1.000
Block[5]	-0.91	0.23	-1.37 – -0.46	**0.002**	-0.96	0.27	-1.49 – -0.44	**0.007**	0.95	0.28	0.40 – 1.49	**0.013**
Group[BP]	0.08	0.23	-0.37 – 0.53	1.000	0.21	0.25	-0.28 – 0.70	1.000	-0.06	0.26	-0.57 – 0.46	1.000
Group[R]	-0.29	0.24	-0.77 – 0.19	1.000	-0.08	0.27	-0.62 – 0.45	1.000	-0.16	0.28	-0.71 – 0.39	1.000
BMI	0.00	0.06	-0.11 – 0.12	1.000	-0.12	0.06	-0.23 – -0.01	0.500	-0.01	0.06	-0.13 – 0.10	1.000
Age	-0.06	0.05	-0.16 – 0.03	1.000	-0.00	0.05	-0.09 – 0.09	1.000	-0.05	0.05	-0.15 – 0.05	1.000
Anxiety	-0.01	0.06	-0.13 – 0.11	1.000	-0.00	0.06	-0.12 – 0.11	1.000	0.05	0.06	-0.07 – 0.17	1.000
Difficulty	-0.03	0.06	-0.15 – 0.09	1.000	-0.03	0.06	-0.15 – 0.08	1.000	0.01	0.06	-0.11 – 0.13	1.000
Block[2]xGroup[BP]	0.03	0.30	-0.57 – 0.62	1.000	-0.06	0.35	-0.75 – 0.63	1.000	-0.07	0.36	-0.78 – 0.65	1.000
Block[3]xGroup[BP]	0.31	0.30	-0.29 – 0.91	1.000	-0.65	0.35	-1.34 – 0.04	1.000	-0.36	0.36	-1.07 – 0.36	1.000
Block[4]xGroup[BP]	-0.41	0.30	-1.00 – 0.19	1.000	0.03	0.35	-0.66 – 0.72	1.000	-0.14	0.36	-0.85 – 0.58	1.000
Block[5]xGroup[BP]	0.02	0.30	-0.57 – 0.62	1.000	-0.34	0.35	-1.03 – 0.34	1.000	-0.38	0.36	-1.10 – 0.33	1.000
Block[2]xGroup[R]	0.73	0.31	0.11 – 1.35	0.387	0.45	0.37	-0.27 – 1.18	1.000	0.16	0.38	-0.59 – 0.92	1.000
Block[3]xGroup[R]	0.33	0.31	-0.29 – 0.95	1.000	-0.15	0.37	-0.88 – 0.57	1.000	0.07	0.38	-0.68 – 0.82	1.000
Block[4]xGroup[R]	0.34	0.31	-0.28 – 0.95	1.000	0.57	0.37	-0.16 – 1.30	1.000	0.26	0.38	-0.49 – 1.02	1.000
Block[5]xGroup[R]	0.25	0.31	-0.37 – 0.86	1.000	0.19	0.37	-0.54 – 0.92	1.000	-0.30	0.38	-1.05 – 0.46	1.000
Random Effects												
σ^2	0.67				0.85				0.91			
τ_{00}	0.07_{ID}				0.01_{ID}				0.02_{ID}			
ICC	0.10				0.01				0.02			
N	90_{ID}				83_{ID}				83_{ID}			
Observations	450				415				415			
Marginal R^2 / Conditional R^2	0.285 / 0.355				0.177 / 0.186				0.112 / 0.127			



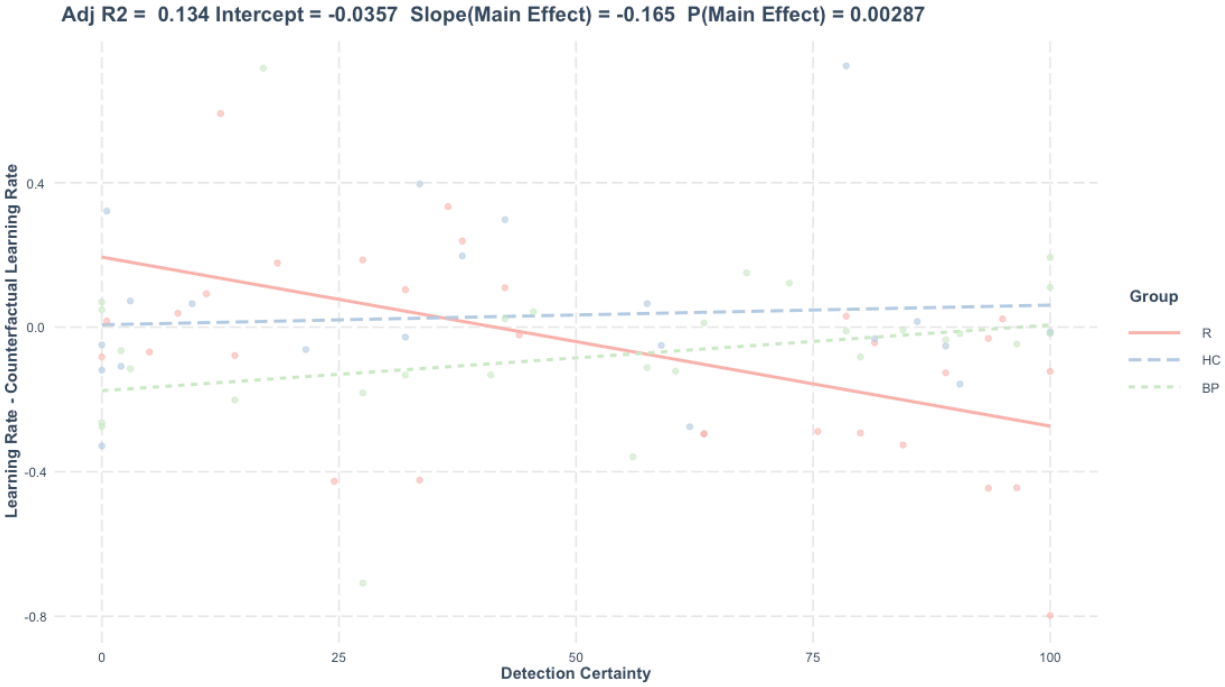
Supplemental Figure 3.6: Correlation plot for the healthy control (HC) group. Shown values (Spearman's ρ) remained significant following Bonferroni correction. Columns marked with "x" were blocked out, as there was no variance in the data. EPSI – Eating Pathology Symptom Inventory; MAIA – Multidimensional Assessment of Interoceptive Awareness; ATQ – Adult Temperament Questionnaire (Effortful Control subscale); WASI – Weschler Abbreviated Scale of Intelligence.



Supplemental Figure 3.7: Correlation plot for the restricting (R) group. Shown values (Spearman's ρ) remained significant following Bonferroni correction. EPSI – Eating Pathology Symptom Inventory; MAIA – Multidimensional Assessment of Interoceptive Awareness; ATQ – Adult Temperament Questionnaire (Effortful Control subscale); WASI – Weschler Abbreviated Scale of Intelligence.



Supplemental Figure 3.8: Correlation plot for the binge-eating/purging (BP) group. Shown values (Spearman's ρ) remained significant following Bonferroni correction. EPSI – Eating Pathology Symptom Inventory; MAIA – Multidimensional Assessment of Interoceptive Awareness; ATQ – Adult Temperament Questionnaire (Effortful Control subscale); WASI – Weschler Abbreviated Scale of Intelligence.



Supplemental Figure 3.9: Plot depicting the interaction effect between group and post-task visual analogue scale (VAS) detection certainty ratings in predicting the difference between learning rate and counterfactual learning rates. R – restricting, BP – binge-eating/purging, HC – healthy control.

Chapter 4 INTEGRATED SUMMARY

Despite higher risk of mortality, psychological and psychosocial impairment, medical and psychiatric comorbitiy, and significant public health burden (Arcelus et al., 2011; Mitchell & Crow, 2006; Solmi et al., 2024), treatment efficacy for individuals suffering from eating disorders is generally low, resulting in either failure to remediate or eventual relapse for a substantial proportion of patients (Eddy et al., 2017). First-line psychological interventions for these populations (e.g., Cognitive Behavioral Therapy – Enhanced) are often based in cognitive-behavioral therapies that assume intact learning and decision-making processes in patients and rely heavily on these, and other, executive functions to shape patients' thoughts, emotions, and behaviors; however, there is evidence to suggest that learning and decision making in individuals with eating disorders is altered (C. S. Brown et al., 2024; Christian & Levinson, 2022; Schaefer & Steinglass, 2021), possibly impacting the effectiveness of current gold-standard psychotherapeutic approaches (Reilly et al., 2019). Consequently, there is an established need for improved understanding of neurocognitive mechanisms that contribute to eating disorder symptoms in these populations, with computational psychiatry frameworks offering a potential pathway towards more granular knowledge of latent mechanisms underlying observed neurocognitive processes that interact with and maintain pathology (Haynos et al., 2022; Radzikowska et al., 2025; Wierenga et al., 2022). This dissertation represents a first step towards application of computational psychiatry principles in an effort to elicit greater understanding of underlying cognitive processes involved in eating disorder phenomenology, across eating disorder diagnoses, contexts, and valences, using a reinforcement learning lens.

Collectively, the three studies included in this dissertation suggest two core considerations: the first is the utility in model-based, computational approaches to analyzing

behavioral data in eating disorder populations, and the second is the orthogonality of relevant latent learning mechanisms between restricting and binge-eating/purging eating disorder phenotypic expressions. Considering the former, in all three studies, model-agnostic analyses of observed choice behavior during the two-step, probabilistic gambling, and the breathing learning tasks revealed almost no significant group differences between eating disorder groups and control groups. Using computational (i.e., model-based) reinforcement learning approaches revealed group differences in latent learning processes, such as speed of learning and explore/exploit tradeoffs, that illustrate how choices are made and how cognitive processing streams that lead to a choice may go awry. This highlights the important role of computational frameworks in psychiatric research (Adams et al., 2016; Huys et al., 2016, 2021), particularly as this field begins to redefine traditional nosological models and embrace more dimensional, idiographic approaches that emphasize, for example, use of generative models for detection, prediction, and treatment selection (Friston et al., 2017). Relevant to this dissertation, alterations in latent mechanisms may suggest underlying impairments in brain-based computations that can be linked to disordered behavior, even if choice behavior during a task is otherwise unaffected. The role of equifinality in task performance is not well understood, but efforts to utilize these frameworks and methods have begun to illustrate how this may occur. Although task performance was similar between groups in all three studies, we have shown that, **in Study 1**, adolescents with anorexia nervosa restricting subtype (AN-R) used greater model-free learning for reward compared to healthy controls; in **Study 2**, that adults with binge-eating/purging eating disorder phenotypes (BP) were slower to integrate trial-by-trial prediction errors and were more insensitive to valence; in **Study 3**, that adults with restrictive eating disorder phenotypes (R) were faster to integrate both factual and fictive trial-by-trial prediction errors. These results begin

to build granular understanding of how cognitive systems in eating disorders may be altered, how this may differ depending on symptomatology (but necessarily diagnosis), and how psychotherapeutic treatments may be adapted (or selected) to address identified deficits, inefficiencies, or non-optimal processes (Berwian et al., 2025; Niv et al., 2021).

The latter consideration emphasizes distinctions between primarily restricting and binge-eating/purging eating disorder presentations in this work, a finding that was consistent across all three studies. Specifically, **Study 1** showed that adolescents with AN-R were characterized by heavier weighting of model-free learning for reward and attenuated model-free and model-based learning for punishment compared to healthy controls during a two-step Markov decision making task, yet adolescents with anorexia nervosa binge-eating/purging subtype (AN-BP) did not show any differences. This could suggest a lack of maturation in the relative weighting of these systems in adolescents with AN-R, particularly for reward, that may contribute to entrenched restrictive behaviors without engagement in binge-eating/purging cycles. Relatedly, greater eating disorder severity in the AN-R group was seen in individuals with weaker associations between model-based learning for reward and functional connectivity between neural regions linked to reward processing. No such association was observed in the AN-BP group, although it is possible that these systems may remain intact in adolescence and deteriorate in later stages of development, as is seen in adolescents with higher compulsivity (Vaghi et al., 2020).

Study 2 found valence-agnostic attenuation of speed of learning in the BP group compared to both R and undergraduate (UG) comparator groups during a gambling task. Consequently, valence-dependent learning bias in the BP group was also attenuated (i.e., equal weighting of learning from positive and negative outcomes/overweighting of learning from negative outcomes), in stark contrast to the R group which showed higher learning bias (i.e.,

overweighting of learning from positive outcomes). This study highlights phenotypic differences in how valenced information during a risk-related gambling task is prioritized and integrated into choice behavior. The BP group exhibited less sensitivity to prediction errors as they learned about the values of the stimuli that were presented to them and less prioritization between information about wins and losses. Interestingly, lower learning bias was associated with greater binge eating severity, indicating a possible mechanism that could contribute to repeated binge behavior. The R group, however, was characterized by greater sensitivity to prediction errors, particularly those linked to wins. These individuals overweighted learning from wins, compared to losses, a bias that may be considered an avoidance mechanism in the context of risk (Paulus, 2020).

Meanwhile, **Study 3** was conducted in the context of aversive interoceptive learning, highlighting phenotypic differences in learning between the R and BP groups. The R group showed, again, higher sensitivity to prediction errors and, consequently, faster updating of stimuli/cue valuation compared to the BP group, which, similar to **Study 2**, showed less sensitivity and slower updating. Notably, the R group appeared to use both factual updating based on experienced outcomes and fictive updating based on hypothetical outcomes as they learned about the cues. Overweighting of fictive updating would bias valuation of stimuli/cues to become based on outcomes that were not directly experienced, suggesting less reliance on experience. This may provide support for Bayesian accounts of interoceptive processing (e.g., interoceptive inference, interoceptive predictive coding) in psychiatric disorders, in which sensory experience is shaped by non-optimal or maladaptive prior beliefs or expectations, for example hyper-precise prior beliefs (Critchley & Patchitt, 2024; Gu et al., 2019; Khalsa & Lapidus, 2016).

Together, these studies demonstrate dissociable alterations in latent learning and decision-making mechanisms in the context of habit and goal-directed systems, reward and punishment valences, risk, and interoception in individuals with eating disorders. This work suggests that restrictive eating disorder phenotypes may be characterized by increased reliance on reward-related model-free learning in adolescence, augmented speed of learning in both risk and interoceptive contexts, less prioritization of feedback signaling loss under conditions of risk, and greater reliance on fictive outcomes during aversive interoceptive learning; meanwhile, binge-eating/purging eating disorder phenotypes may be characterized by attenuated speed of learning in both risk and interoceptive contexts and more equal prioritization of feedback signaling wins and losses under conditions of risk. Although preliminary, this work could inform future efforts in the development of precision psychiatry practices, as well as in the development of more specific models of eating disorder development, maintenance, and illness trajectory across behavioral phenotypes.

REFERENCES

Aberg, K. C., Doell, K. C., & Schwartz, S. (2016). Linking individual learning styles to approach-avoidance motivational traits and computational aspects of reinforcement learning. *PLoS ONE*, *11*(11), 1–16. https://doi.org/10.1371/journal.pone.0166675

Adamantidis, A. R., Tsai, H. C., Boutrel, B., Zhang, F., Stuber, G. D., Budygin, E. A., Touriño, C., Bonci, A., Deisseroth, K., & de Lecea, L. (2011). Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *Journal of Neuroscience*, *31*(30), 10829–10835. https://doi.org/10.1523/JNEUROSCI.2246-11.2011

Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2016). Computational Psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery and Psychiatry*, *87*(1), 53–63. https://doi.org/10.1136/jnnp-2015-310737

Adhikari, A., Lerner, T. N., Finkelstein, J., Pak, S., Jennings, J. H., Davidson, T. J., Ferenczi, E., Gunaydin, L. A., Mirzabekov, J. J., Ye, L., Kim, S. Y., Lei, A., & Deisseroth, K. (2015). Basomedial amygdala mediates top-down control of anxiety and fear. *Nature*, *527*(7577), 179–185. https://doi.org/10.1038/nature15698

Adoue, C., Jaussent, I., Olié, E., Beziat, S., Eynde, F. V. den, Courtet, P., & Guillaume, S. (2015). A further assessment of decision-making in anorexia nervosa. *European Psychiatry*, *30*(1), 121–127. https://doi.org/10.1016/j.eurpsy.2014.08.004

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*, 24–57. https://doi.org/10.1162/CPSY_a_00002

Allen, M. (2020). Unravelling the Neurobiology of Interoceptive Inference. *Trends in Cognitive Sciences*, *24*(4), 265–266. https://doi.org/10.1016/j.tics.2020.02.002

Andersson, J. L. R., Hutton, C., Ashburner, J. T., Turner, R., & Friston, K. J. (2001). Modeling Geometric Deformations in EPI Time Series. *NeuroImage*, *13*(5), 903–919. https://doi.org/10.1006/nimg.2001.0746

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: A meta-analysis of 36 studies. *Archives of General Psychiatry*, *68*(7), 724–731. https://doi.org/10.1001/archgenpsychiatry.2011.74

Ashburner, J. T. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, *38*(1), 95–113. https://doi.org/10.1016/j.neuroimage.2007.07.007

Ashburner, J. T., & Friston, Karl. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018

Aspen, V., Weisman, H., Vannucci, A., Nafiz, N., Gredysa, D., Kass, A. E., Trockel, M., Jacobi, C., Wilfley, D. E., & Taylor, C. B. (2014). Psychiatric co-morbidity in women presenting across the continuum of disordered eating. *Eating Behaviors*, *15*(4), 686–693. https://doi.org/10.1016/j.eatbeh.2014.08.023

Atiye, M., Miettunen, J., & Raevuori-Helkamaa, A. (2015). A meta-analysis of temperament in eating disorders. *European Eating Disorders Review*, *23*(2), 139–146. https://doi.org/10.1002/erv.2342

Badoud, D., & Tsakiris, M. (2017). From the body's viscera to the body's image: Is there a link between interoception and body image concerns? *Neuroscience and Biobehavioral Reviews*, *77*(December 2016), 237–246. https://doi.org/10.1016/j.neubiorev.2017.03.017

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, *35*(1), 48–69. https://doi.org/10.1038/npp.2009.131

Bardone-Cone, A. M., Wonderlich, S. A., Frost, R. O., Bulik, C. M., Mitchell, J. E., Uppala, S., & Simonich, H. (2007). Perfectionism and eating disorders: Current status and future directions. *Clinical Psychology Review*, *27*(3), 384–405. https://doi.org/10.1016/j.cpr.2006.12.005

Baribault, B., & Collins, A. G. E. (2023). Troubleshooting Bayesian cognitive models. *Psychological Methods*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/met0000554

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews. Neuroscience*, *16*(7), 419–429. https://doi.org/10.1038/nrn3950

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad Is Stronger Than Good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Beck, I., Smits, D. J. M., Claes, L., Vandereycken, W., & Bijttebier, P. (2009). Psychometric evaluation of the behavioral inhibition/behavioral activation system scales and the sensitivity to punishment and sensitivity to reward questionnaire in a sample of eating disordered patients. *Personality and Individual Differences*, *47*(5), 407–412. https://doi.org/10.1016/j.paid.2009.04.007

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, *23*(2), 137–152. https://doi.org/10.1109/TMI.2003.822821

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Berg, K. C., Swanson, S. A., Stiles-Shields, E. C., Eddy, K. T., Peterson, C. B., & Grange, D. Le. (2013). Response patterns on interview and questionnaire versions of the Eating Disorder Examination and their impact on latent structure analyses. *Comprehensive Psychiatry*, *54*(5), 506–516. https://doi.org/https://doi.org/10.1016/j.comppsych.2012.12.006

Bernardoni, F., Geisler, D., King, J. A., Javadi, A. H., Ritschel, F., Murr, J., Reiter, A. M. F., Rössner, V., Smolka, M. N., Kiebel, S., & Ehrlich, S. (2017). Altered Medial Frontal Feedback Learning Signals in Anorexia Nervosa. *Biological Psychiatry*, *83*(3), 235–243. https://doi.org/10.1016/j.biopsych.2017.07.024

Bernardoni, F., Geisler, D., King, J. A., Javadi, A. H., Ritschel, F., Murr, J., Reiter, A. M. F., Rössner, V., Smolka, M. N., Kiebel, S., & Ehrlich, S. (2018). Altered Medial Frontal Feedback Learning Signals in Anorexia Nervosa. *Biological Psychiatry*, *83*(3), 235–243. https://doi.org/10.1016/j.biopsych.2017.07.024

Bernardoni, F., King, J. A., Geisler, D., Ritschel, F., Schwoebel, S., Reiter, A. M. F., Endrass, T., Rossner, V., Smolka, M. N., & Ehrlich, S. (2021). More by Stick Than by Carrot: A Reinforcement Learning Style Rooted in the Medial Frontal Cortex in Anorexia Nervosa. *Journal of Abnormal Psychology*, *130*(7), 736–747. https://doi.org/10.1037/abn0000690

Berner, L. A., Fiore, V. G., Chen, J. Y., Krueger, A., Kaye, W. H., Viranda, T., & Wit, S. de. (2023). Impaired belief updating and devaluation in adult women with bulimia nervosa. *Translational Psychiatry*, *13*(1), 1–9. https://doi.org/10.1038/s41398-022-02257-6

Berner, L. A., Harlé, K. M., Simmons, A. N., Yu, A., Paulus, M. P., Bischoff-Grethe, A., Wierenga, C. E., Bailer, U. F., & Kaye, W. H. (2023). State-specific alterations in the neural computations underlying inhibitory control in women remitted from bulimia nervosa. *Molecular Psychiatry*, (November 2022), 1–8. https://doi.org/10.1038/s41380-023-02063-6

Berner, L. A., Simmons, A. N., Wierenga, C. E., Bischoff-Grethe, A., Paulus, M. P., Bailer, U. F., Ely, A. V., & Kaye, W. H. (2018). Altered interoceptive activation before, during, and after aversive breathing load in women remitted from anorexia nervosa. *Psychological Medicine*, *48*(1), 142–154. https://doi.org/10.1017/S0033291717001635

Berner, L. A., Simmons, A. N., Wierenga, C. E., Bischoff-Grethe, A., Paulus, M. P., Bailer, U. F., & Kaye, W. H. (2019). Altered anticipation and processing of aversive interoceptive

experience among women remitted from bulimia nervosa. *Neuropsychopharmacology*, *44*(7), 1265–1273. https://doi.org/10.1038/s41386-019-0361-4

Berwian, I. M., Hitchock, P., Pisupati, S., Schoen, G., & Niv, Y. (2025). Using computational models of learning to advance cognitive behavioral therapy. *Communications Psychology*, *3*(1), 72. https://doi.org/10.1038/s44271-025-00251-4

Bischoff-Grethe, A., McCurdy, D., Grenesko-Stevens, E., Irvine, L. E. (Zoe), Wagner, A., Yau, W. Y. W., Fennema-Notestine, C., Wierenga, C. E., Fudge, J. L., Delgado, M. R., & Kaye, W. H. (2013). Altered brain response to reward and punishment in adolescents with Anorexia nervosa. *Psychiatry Research - Neuroimaging*, *214*(3), 331–340. https://doi.org/10.1016/j.pscychresns.2013.07.004

Bischoff-Grethe, A., McCurdy, D., Grenesko-Stevens, E., (Zoe) Irvine, L. E., Wagner, A., Wendy Yau, W. Y., Fennema-Notestine, C., Wierenga, C. E., Fudge, J. L., Delgado, M. R., & Kaye, W. H. (2013). Altered brain response to reward and punishment in adolescents with Anorexia nervosa. *Psychiatry Research - Neuroimaging*, *214*(3), 331–340. https://doi.org/10.1016/j.pscychresns.2013.07.004

Bishop, S. J., & Gagne, C. (2018). Anxiety, depression, and decision making: A computational perspective. *Annual Review of Neuroscience*, *41*, 371–388. https://doi.org/10.1146/annurev-neuro-080317-062007

Bolenz, F., & Eppinger, B. (2022). Valence bias in metacontrol of decision making in adolescents and young adults. *Child Development*, *93*(2), e103–e116. https://doi.org/10.1111/cdev.13693

Bolles, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review*, *79*(5), 394–409. https://doi.org/10.1037/h0033120

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual Choice and Learning in a Neural Network Centered on Human Lateral Frontopolar Cortex. *PLOS Biology*, *9*(6), e1001093. https://doi.org/10.1371/journal.pbio.1001093

Braun, D. L., Sunday, S. R., & Halmi, K. A. (1994). Psychiatric comorbidity in patients with eating disorders. *Psychological Medicine*, *24*(4), 859–867. https://doi.org/DOI:%2010.1017/S0033291700028956

Brockmeyer, T., Hahn, C., Reetz, C., Schmidt, U., & Friederich, H.-C. (2015). Approach bias and cue reactivity towards food in people with high versus low levels of food craving. *Appetite*, *95*, 197–202. https://doi.org/10.1016/j.appet.2015.07.013

Brooks, S. J., Rask-Andersen, M., Benedict, C., & Schiöth, H. B. (2012). A debate on current eating disorder diagnoses in light of neurobiological findings: Is it time for a spectrum model? *BMC Psychiatry*, *12*. https://doi.org/10.1186/1471-244X-12-76

Brown, C. S., Nuñez, A., & Wierenga, C. E. (2024). Altered value-based decision-making in anorexia nervosa: A systematic review. *Neuroscience & Biobehavioral Reviews*, *167*, 105944. https://doi.org/10.1016/j.neubiorev.2024.105944

Brown, M., Robinson, L., Campione, G. C., Wuensch, K., Hildebrandt, T., & Micali, N. (2017). Intolerance of Uncertainty in Eating Disorders: A Systematic Review and Meta-Analysis. *European Eating Disorders Review*, *25*(5), 329–343. https://doi.org/10.1002/erv.2523

Brown, T. A., Berner, L. A., Jones, M. D., Reilly, E. E., Cusack, A., Anderson, L. K., Kaye, W. H., & Wierenga, C. E. (2017). Psychometric Evaluation and Norms for the Multidimensional Assessment of Interoceptive Awareness (MAIA) in a Clinical Eating Disorders Sample. *European Eating Disorders Review*, *25*(5), 411–416. https://doi.org/10.1002/erv.2532

Brown, T. A., Vanzhula, I. A., Reilly, E. E., Levinson, C. A., Berner, L. A., Krueger, A., Lavender, J. M., Kaye, W. H., & Wierenga, C. E. (2020). Body mistrust bridges interoceptive awareness and eating disorder symptoms. In *Journal of Abnormal Psychology* (Vol. 129, Issue 5, pp. 445–456). American Psychological Association. https://doi.org/10.1037/abn0000516

Brown, V. M., Chen, J., Gillan, C. M., & Price, R. B. (2020). Improving the Reliability of Computational Analyses: Model-Based Planning and Its Relationship With Compulsivity. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(6), 601–609. https://doi.org/10.1016/j.bpsc.2019.12.019

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596. https://doi.org/10.1038/nn.3961

(Bud) Craig, A. D. (2009). How do you feel—Now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), 59–70. https://doi.org/10.1038/nrn2555

Calhoun, V. D., Wager, T. D., Krishnan, A., Rosch, K. S., Seymour, K. E., Nebel, M. B., Mostofsky, S. H., Nyalakanai, P., & Kiehl, K. (2017). The impact of T1 versus EPI spatial normalization templates for fMRI data analyses. *Human Brain Mapping*, *38*(11), 5331–5342. https://doi.org/10.1002/hbm.23737

Carter, J. C., Blackmore, E., Sutandar-Pinnock, K., & Woodside, D. B. (2004). Relapse in anorexia nervosa: A survival analysis. *Psychological Medicine*, *34*(4), 671–679. https://doi.org/10.1017/S0033291703001168

Carter, J. C., Mercer-Lynn, K. B., Norwood, S. J., Bewell-Weiss, C. V, Crosby, R. D., Woodside, D. B., & Olmsted, M. P. (2012). A prospective study of predictors of relapse in anorexia nervosa: Implications for relapse prevention. *Psychiatry Research*, *200*(2), 518–523. https://doi.org/https://doi.org/10.1016/j.psychres.2012.04.037

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*. https://doi.org/10.1037//0022-3514.67.2.319

Cazé, R. D., & Van Der Meer, M. A. A. (2013). Adaptive properties of differential learning rates for positive and negative outcomes. *Biological Cybernetics*, *107*(6), 711–719. https://doi.org/10.1007/s00422-013-0571-5

Cha, J., Ide, J. S., Bowman, F. D., Simpson, H. B., Posner, J., & Steinglass, J. E. (2016). Abnormal reward circuitry in anorexia nervosa: A longitudinal, multimodal MRI study. *Human Brain Mapping*, *37*(11), 3835–3846. https://doi.org/10.1002/hbm.23279

Chai, X. J., Castañón, A. N., Öngür, D., & Whitfield-Gabrieli, S. (2012). Anticorrelations in resting state networks without global signal regression. *NeuroImage*, *59*(2), 1420–1428. https://doi.org/10.1016/j.neuroimage.2011.08.048

Chan, T. W. S., Ahn, W.-Y., Bates, J. E., Busemeyer, J. R., Guillaume, S., Redgrave, G. W., Danner, U. N., & Courtet, P. (2014). Differential impairments underlying decision making in anorexia nervosa and bulimia nervosa: A cognitive modeling analysis. *International Journal of Eating Disorders*, *47*(2), 157–167. https://doi.org/10.1002/eat.22223

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience and Biobehavioral Reviews*, *55*, 247–267. https://doi.org/10.1016/j.neubiorev.2015.05.005

Christian, C., & Levinson, C. A. (2022). An integrated review of fear and avoidance learning in anxiety disorders and application to eating disorders. *New Ideas in Psychology*, *67*(June), 100964. https://doi.org/10.1016/j.newideapsych.2022.100964

Claes, L., Nederkoorn, C., Vandereycken, W., Guerrieri, R., & Vertommen, H. (2006). Impulsiveness and lack of inhibitory control in eating disorders. *Eating Behaviors*, *7*(3), 196–203. https://doi.org/10.1016/j.eatbeh.2006.05.001

Cloninger, R., Svrakic, D., & Przybeck, T. (1994). The temperament and character inventory (TCI): A guide to its development and use. *Archives of General Psychiatry*.

Collins, A. G. E. (2018). Learning Structures Through Reinforcement. In R. Morris, A. Bornstein, & A. B. T.-G.-D. D. M. Shenhav (Eds.), *Goal-Directed Decision Making: Computations and Neural Circuits* (pp. 105–123). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-812098-9.00005-X

Collins, A. G. E., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience*, *21*(10), 576–586. https://doi.org/10.1038/s41583-020-0355-6

Coniglio, K. A., Becker, K. R., Tabri, N., Keshishian, A. C., Miller, J. D., Eddy, K. T., & Thomas, J. J. (2018). Factorial integrity and validation of the Eating Pathology Symptoms Inventory (EPSI). *Eating Behaviors*, *31*(July), 1–7. https://doi.org/10.1016/j.eatbeh.2018.07.004

Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, *6*(1), 1–8. https://doi.org/10.1002/1098-108X(198701)6:1%3C1::AID-EAT2260060102%3E3.0.CO;2-9

Craig, A. D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. *Nature Reviews. Neuroscience*, *3*(8), 655–666.

Crawley, D., Zhang, L., Jones, E. J. H., Ahmad, J., Oakley, B., Cáceres, A. S. J., Charman, T., Buitelaar, J. K., Murphy, D. G. M., Chatham, C., Ouden, H. den, & Loth, E. (2020). Modeling flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and less optimal learning in autism in each age group. *PLoS Biology*, *18*(10). https://doi.org/10.1371/journal.pbio.3000908

Critchley, H. D., & Patchitt, J. (2024). *Interoception, Insula, and Autonomic Integration: Relevance to the Expression and Treatment of Psychiatric Symptoms* (pp. 1–23). Springer. https://doi.org/10.1007/7854_2024_518

Dang, J., King, K. M., & Inzlicht, M. (2020). Why Are Self-Report and Behavioral Measures Weakly Correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Davenport, P. W., & Vovk, A. (2009). Cortical and subcortical central neural pathways in respiratory sensations. *Respiratory Physiology & Neurobiology*, *167*(1), 72–86. https://doi.org/https://doi.org/10.1016/j.resp.2008.10.001

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Dayan, P. (1993). Improving Generalization for Temporal Difference Learning: The Successor Representation. *Neural Computation*, *5*(4), 613–624. https://doi.org/10.1162/neco.1993.5.4.613

de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I., & Ridderinkhof, K. R. (2012). Corticostriatal connectivity underlies individual differences in the balance

between habitual and goal-directed action control. *Journal of Neuroscience*, *32*(35), 12066–12075. https://doi.org/10.1523/JNEUROSCI.1088-12.2012

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning. *Psychological Science*, *27*(6), 848–858. https://doi.org/10.1177/0956797616639301

Deen, B., Pitskel, N. B., & Pelphrey, K. A. (2011). *Three Systems of Insular Functional Connectivity Identified with Cluster Analysis*. https://dx.doi.org/10.1093/cercor/bhq186

DeGuzman, M., Shott, M. E., Yang, T. T., Riederer, J., & Frank, G. K. W. (2017). Association of Elevated Reward Prediction Error Response With Weight Gain in Adolescent Anorexia Nervosa. *The American Journal of Psychiatry*, *174*(6), 557–565. https://doi.org/10.1176/appi.ajp.2016.16060671

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, *18*(5), 767–772. https://doi.org/10.1038/nn.3981

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., Fox, M. D., Snyder, A. Z., Vincent, J. L., Raichle, M. E., Schlaggar, B. L., & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, *104*(26), 11073–11078. https://doi.org/10.1073/pnas.0704320104

Dougherty, E. N., Bottera, A. R., Forester, G., Schaefer, L. M., Forbes, E. E., & Wildes, J. E. (2024). Prospective associations between cognitive flexibility and eating disorder symptoms in anorexia nervosa and bulimia nervosa. *Psychiatry Research*, *332*(September 2023), 115717. https://doi.org/10.1016/j.psychres.2024.115717

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. https://doi.org/10.7554/eLife.75474

Eddy, K. T., Tabri, N., Thomas, J. J., Murray, H. B., Keshaviah, A., Hastings, E., Edkins, K., Krishna, M., Herzog, D. B., Keel, P. K., & Franko, D. L. (2017). Recovery from anorexia

nervosa and bulimia nervosa at 22-year follow-up. *Journal of Clinical Psychiatry*, *78*(2), 184–189. https://doi.org/10.4088/JCP.15m10393

Eldar, E., Hauser, T. U., Dayan, P., & Dolan, R. J. (2016). Striatal structure and function predict individual biases in learning to avoid pain. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(17), 4812–4817. https://doi.org/10.1073/pnas.1519829113

Eppinger, B., Walter, M., & Li, S. C. (2017). Electrophysiological correlates reflect the integration of model-based and model-free decision information. *Cognitive, Affective and Behavioral Neuroscience*, *17*(2), 406–421. https://doi.org/10.3758/s13415-016-0487-3

Eryilmaz, H., Rodriguez-Thompson, A., Tanner, A. S., Giegold, M., Huntington, F. C., & Roffman, J. L. (2017). Neural determinants of human goal-directed vs. Habitual action control and their relation to trait motivation. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-06284-y

Evans, D. E., & Rothbart, M. K. (2007). Developing a model for adult temperament. *Journal of Research in Personality*, *41*(4), 868–888. https://doi.org/https://doi.org/10.1016/j.jrp.2006.11.002

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, *41*(5), 509–528. https://doi.org/https://doi.org/10.1016/S0005-7967(02)00088-8

Fermin, A. S. R., Friston, K., & Yamawaki, S. (2022). An insula hierarchical network architecture for active interoceptive inference. *Royal Society Open Science*, *9*(6), 220226. https://doi.org/10.1098/rsos.220226

Filoteo, J. V., Paul, E. J., Ashby, F. G., Frank, G. K. W., Helie, S., Rockwell, R., Bischoff-Grethe, A., Wierenga, C., & Kaye, W. H. (2014). Simulating Category Learning and Set Shifting Deficits in Patients Weight-Restored From Anorexia Nervosa. *Neuropsychology*, *28*(5), 741–751. https://doi.org/10.1037/neu0000055

First, M. B., Williams JBW, Karg RS, & Spitzer, R. (2015). Structured Clinical Interview for DSM-5—Research Version (SCID-5 for DSM-5, Research Version; SCID-5-RV). In *The Encyclopedia of Clinical Psychology*.

Fischer, A. G., & Ullsperger, M. (2013). Real and Fictive Outcomes Are Processed Differently but Converge on a Common Adaptive Mechanism. *Neuron*, *79*(6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Fladung, A. K., Schulze, U. M. E., Schöll, F., Bauer, K., & Grön, G. (2013). Role of the ventral striatum in developing anorexia nervosa. *Translational Psychiatry*, *3*(July), 2–7. https://doi.org/10.1038/tp.2013.88

Foerde, K., Daw, N. D., Rufin, T., Walsh, B. T., Shohamy, D., & Steinglass, J. E. (2021). Deficient Goal-Directed Control in a Population Characterized by Extreme Goal Pursuit. *Journal of Cognitive Neuroscience*, *33*(3), 463–481. https://doi.org/10.1162/jocn_a_01655

Foerde, K., & Steinglass, J. E. (2017). Decreased feedback learning in anorexia nervosa persists after weight restoration. *International Journal of Eating Disorders*, *50*(4), 415–423. https://doi.org/10.1002/eat.22709

Forbush, K. T., Wildes, J. E., Pollack, L. O., Dunbar, D., Luo, J., Patterson, K., Petruzzi, L., Pollpeter, M., Miller, H., Stone, A., Bright, A., & Watson, D. (2013). Development and validation of the Eating Pathology Symptoms Inventory (EPSI). In *Psychological Assessment* (Vol. 25, Issue 3, pp. 859–878). American Psychological Association. https://doi.org/10.1037/a0032639

Frank, G. K. W. (2013). Altered brain reward circuits in eating disorders: Chicken or egg? *Current Psychiatry Reports*, *15*(10). https://doi.org/10.1007/s11920-013-0396-x

Frank, G. K. W., Deguzman, M. C., Shott, M. E., Laudenslager, M. L., Rossi, B., & Pryor, T. (2018). Association of Brain Reward Learning Response with Harm Avoidance, Weight Gain, and Hypothalamic Effective Connectivity in Adolescent Anorexia Nervosa. *JAMA Psychiatry*, *75*(10), 1071–1080. https://doi.org/10.1001/jamapsychiatry.2018.2151

Frank, G. K. W., Reynolds, J. R., Shott, M. E., Jappe, L., Yang, T. T., Tregellas, J. R., & O'Reilly, R. C. (2012). Anorexia nervosa and obesity are associated with opposite brain reward response. *Neuropsychopharmacology*, *37*(9), 2031–2046. https://doi.org/10.1038/npp.2012.51

Frank, G. K. W., Reynolds, J. R., Shott, M. E., & O'Reilly, R. C. (2011). Altered temporal difference learning in bulimia nervosa. *Biological Psychiatry*, *70*(8), 728–735. https://doi.org/10.1016/j.biopsych.2011.05.011

Frazier, J. A., Chiu, S., Breeze, J. L., Makris, N., Lange, N., Kennedy, D. N., Herbert, M. R., Bent, E. K., Koneru, V. K., Dieterich, M. E., Hodge, S. M., Rauch, S. L., Grant, P. E., Cohen, B. M., Seidman, L. J., Caviness, V. S., & Biederman, J. (2005). Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder. *American Journal of Psychiatry*, *162*(7), 1256–1265. https://doi.org/10.1176/appi.ajp.162.7.1256

Friston, K. J. (2016). Computational Nosology and Precision Psychiatry. *Computational Psychiatry*, (2016), 2–23. https://doi.org/10.7551/mitpress/9780262035422.003.0011

Friston, K. J., Redish, A. D., & Gordon, J. A. (2017). Computational Nosology and Precision Psychiatry. *Computational Psychiatry (Cambridge, Mass.)*, *1*, 2–23. https://doi.org/10.1162/CPSY_a_00001

Friston, Karl. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *3*(3), 165–189. https://doi.org/10.1002/hbm.460030303

Gehrlach, D. A., Dolensek, N., Klein, A. S., Roy Chowdhury, R., Matthys, A., Junghänel, M., Gaitanos, T. N., Podgornik, A., Black, T. D., Reddy Vaka, N., Conzelmann, K. K., & Gogolla, N. (2019). Aversive state processing in the posterior insular cortex. *Nature Neuroscience*, *22*(9), 1424–1437. https://doi.org/10.1038/s41593-019-0469-1

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis*, *1*(3), 515–534. https://doi.org/10.1214/06-BA117A

Gelman, A., Lee, D., & Guo, J. (2015). Stan: A Probabilistic Programming Language for Bayesian Inference and Optimization. *Journal of Educational and Behavioral Statistics*, *40*(5), 530–543. https://doi.org/10.3102/1076998615606113

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Gershman, S. J. (2015). Do learning rates adapt to the distribution of rewards? *Psychonomic Bulletin and Review*, *22*(5), 1320–1327. https://doi.org/10.3758/s13423-014-0790-3

Gershman, S. J. (2018). The successor representation: Its computational logic and neural substrates. *Journal of Neuroscience*, *38*(33), 7193–7200. https://doi.org/10.1523/JNEUROSCI.0151-18.2018

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goaldirected control. *eLife*, *5*(MARCH2016), 1–24. https://doi.org/10.7554/eLife.11305

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective and Behavioral Neuroscience*, *15*(3), 523–536. https://doi.org/10.3758/s13415-015-0347-6

Glashouwer, K. A., Bloot, L., Veenstra, E. M., Franken, I. H. A., & de Jong, P. J. (2014). Heightened sensitivity to punishment and reward in anorexia nervosa. *Appetite*, *75*, 97–102. https://doi.org/10.1016/j.appet.2013.12.019

Glasser, M. F., Smith, S. M., Marcus, D. S., Andersson, J. L. R., Auerbach, E. J., Behrens, T. E. J., Coalson, T. S., Harms, M. P., Jenkinson, M., Moeller, S., Robinson, E. C., Sotiropoulos, S. N., Xu, J., Yacoub, E., Ugurbil, K., & Van Essen, D. C. (2016). The Human Connectome Project's neuroimaging approach. *Nature Neuroscience*, *19*(9), 1175–1187. https://doi.org/10.1038/nn.4361

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., & Jenkinson, M. (2013). The Minimal Preprocessing Pipelines for the Human Connectome Project and for the WU-Minn HCP Consortium. *Neuroimage*, *80*, 105–12404. https://doi.org/10.1016/j.neuroimage.2013.04.127.

Godier, L. R., de Wit, S., Pinto, A., Steinglass, J. E., Greene, A. L., Scaife, J., Gillan, C. M., Walsh, B. T., Simpson, H. B., & Park, R. J. (2016). An investigation of habit learning in Anorexia Nervosa. *Psychiatry Research*, *244*, 214–222. https://doi.org/10.1016/j.psychres.2016.07.051

Goldstein, J. M., Seidman, L. J., Makris, N., Ahern, T., O'Brien, L. M., Caviness, V. S., Kennedy, D. N., Faraone, S. V, & Tsuang, M. T. (2007). Hypothalamic Abnormalities in Schizophrenia: Sex Effects and Genetic Vulnerability. *Biological Psychiatry*, *61*(8), 935–945. https://doi.org/https://doi.org/10.1016/j.biopsych.2006.06.027

Grange, D. le, & Loeb, K. L. (2007). Early identification and treatment of eating disorders: Prodrome to syndrome. *Early Intervention in Psychiatry*, *1*(1), 27–39. https://doi.org/10.1111/j.1751-7893.2007.00007.x

Groenewegen, H. J., Wright, C. I., Beijer, A. V. J., & Voorn, P. (1999). Convergence and segregation of ventral striatal inputs and outputs. *Annals of the New York Academy of Sciences*, *877*, 49–63. https://doi.org/10.1111/j.1749-6632.1999.tb09260.x

Gu, X., FitzGerald, T. H. B., & Friston, K. J. (2019). Modeling subjective belief states in computational psychiatry: Interoceptive inference as a candidate framework. *Psychopharmacology*, *236*(8), 2405–2412. https://doi.org/10.1007/s00213-019-05300-5

Gueguen, M. C., Schweitzer, E. M., & Konova, A. B. (2021). Computational theory-driven studies of reinforcement learning and decision-making in addiction: What have we learned? *Current Opinion in Behavioral Sciences*, *38*, 40–48. https://doi.org/10.1016/j.cobeha.2020.08.007

Guillaume, S., Gorwood, P., Jollant, F., Eynde, F. V. D., Courtet, P., & Richard-Devantoy, S. (2015). Impaired decision-making in symptomatic anorexia and bulimia nervosa patients: A meta-analysis. *Psychological Medicine*, *45*(16), 3377–3391. https://doi.org/10.1017/S003329171500152X

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003

Hagan, K. E., & Forbush, K. T. (2021). Reward learning in unmedicated women with bulimia nervosa: A pilot investigation. *Journal of Psychiatric Research*, *136*, 63–70. https://doi.org/10.1016/j.jpsychires.2021.01.046

Hagan, K. E., Walsh, B. T., & Haynos, A. F. (2024). *The explore / exploit trade-off: An ecologically valid and translational framework that can advance mechanistic understanding of eating disorders*. (November 2023), 1–7. https://doi.org/10.1002/eat.24173

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., Ernst, M., & Pine, D. S. (2010). Attention bias modification treatment: A meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, *68*(11), 982–990. https://doi.org/10.1016/j.biopsych.2010.07.021

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, *19*(1), 117–126. https://doi.org/10.1038/nn.4173

Harris, E. C., & Barraclough, B. (1998). Excess mortality of mental disorder. *British Journal of Psychiatry*, *173*(JULY), 11–53. https://doi.org/10.1192/bjp.173.1.11

Harrison, A., O'Brien, N., Lopez, C., & Treasure, J. (2010). Sensitivity to reward and punishment in eating disorders. *Psychiatry Research*, *177*(1–2), 1–11. https://doi.org/10.1016/j.psychres.2009.06.010

Harrison, A., Treasure, J., & Smillie, L. D. (2011). Approach and avoidance motivation in eating disorders. *Psychiatry Research*, *188*(3), 396–401. https://doi.org/10.1016/j.psychres.2011.04.022

Harrison, O. K., Köchli, L., Marino, S., Luechinger, R., Hennel, F., Brand, K., Hess, A. J., Frässle, S., Iglesias, S., Vinckier, F., Petzschner, F. H., Harrison, S. J., & Stephan, K. E. (2021). Interoception of breathing and its relationship with anxiety. *Neuron*, *109*(24), 4080-4093.e8. https://doi.org/10.1016/j.neuron.2021.09.045

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive Reward Signals in the Anterior Cingulate Cortex. *Science*, *324*(5929), 948–950. https://doi.org/10.1126/science.1168488

Haynos, A. F., Berg, K. C., Cao, L., Crosby, R. D., Lavender, J. M., Utzinger, L. M., Wonderlich, S. A., Engel, S. G., Mitchell, J. E., Le Grange, D., Peterson, C. B., & Crow, S. J. (2017). Trajectories of higher- and lower-order dimensions of negative and positive affect relative to restrictive eating in anorexia nervosa. In *Journal of Abnormal Psychology* (Vol. 126, Issue 5, pp. 495–505). American Psychological Association. https://doi.org/10.1037/abn0000202

Haynos, A. F., Hall, L. M. J., Lavender, J. M., Peterson, C. B., Crow, S. J., Klimes-Dougan, B., Cullen, K. R., Lim, K. O., & Camchong, J. (2019). Resting state functional connectivity of networks associated with reward and habit in anorexia nervosa. *Human Brain Mapping*, *40*(2), 652–662. https://doi.org/10.1002/hbm.24402

Haynos, A. F., Koithan, E., & Hagan, K. E. (2023). Learned industriousness as a translational mechanism in anorexia nervosa. *Nature Reviews Psychology*, *2*(2), 112–126. https://doi.org/10.1038/s44159-022-00134-z

Haynos, A. F., Lavender, J. M., Nelson, J., Crow, S. J., & Peterson, C. B. (2020). Moving towards specificity: A systematic review of cue features associated with reward and punishment in anorexia nervosa. *Clinical Psychology Review*, *79*(March), 101872. https://doi.org/10.1016/j.cpr.2020.101872

Haynos, A. F., Widge, A. S., Anderson, L. M., & Redish, A. D. (2022). Beyond Description and Deficits: How Computational Psychiatry Can Enhance an Understanding of Decision-Making in Anorexia Nervosa. *Current Psychiatry Reports*, *24*(1), 77–87. https://doi.org/10.1007/s11920-022-01320-9

Homan, P., Levy, I., Feltham, E., Gordon, C., Hu, J., Li, J., Pietrzak, R. H., Southwick, S., Krystal, J. H., Harpaz-Rotem, I., & Schiller, D. (2019). Neural computations of threat in the aftermath of combat trauma. *Nature Neuroscience*, *22*(3), 470–476. https://doi.org/10.1038/s41593-018-0315-x

Huang, H., Thompson, W., & Paulus, M. P. (2017). Computational Dysfunctions in Anxiety: Failure to Differentiate Signal From Noise. *Biological Psychiatry*, *82*(6), 440–446. https://doi.org/10.1016/j.biopsych.2017.07.007

Huettel, S. A. (2006). Behavioral, but not reward, risk modulates activation of prefrontal, parietal, and insular cortices. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(2), 141–151. https://doi.org/10.3758/CABN.6.2.141

Huys, Q. J. M., Browning, M., Paulus, M. P., & Frank, M. J. (2021). Advances in the computational understanding of mental illness. *Neuropsychopharmacology*, *46*(1), 3–19. https://doi.org/10.1038/s41386-020-0746-4

Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Computational Biology*, *7*(4). https://doi.org/10.1371/journal.pcbi.1002028

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, *19*(3), 404–413. https://doi.org/10.1038/nn.4238

Jacquemot, A. M. M. C., & Park, R. (2020). The Role of Interoception in the Pathogenesis and Treatment of Anorexia Nervosa: A Narrative Review. *Frontiers in Psychiatry*, *11*. https://doi.org/10.3389/fpsyt.2020.00281

Janssen, L. K., Mahner, F. P., Schlagenhauf, F., Deserno, L., & Horstmann, A. (2020). Reliance on model-based and model-free control in obesity. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-79929-0

Jappe, L. M., Frank, G. K. W., Shott, M. E., Rollin, M. D. H., Pryor, T., Hagman, J. O., Yang, T. T., & Davis, E. (2011). Heightened sensitivity to reward and punishment in anorexia nervosa. *International Journal of Eating Disorders*, *44*(4), 317–324. https://doi.org/10.1002/eat.20815

Jenkinson, P. M., Koukoutsakis, A., Panagiotopoulou, E., Vagnoni, E., Demartini, B., Gambini, O., Christakou, A., Fotopoulou, A., Jenkinson, P. M., Koukoutsakis, A., Panagiotopoulou, E., Vagnoni, E., Demartini, B., Nistico, V., Gambini, O., Christakou, A., & Jenkinson, P. M. (2023). General Body Appearance Values Modulate Risk Aversion in Eating Restriction Body Appearance Values Modulate Risk Aversion in Eating Restriction. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001445

Jenkinson, P. M., Taylor, L., & Laws, K. R. (2018). Self-reported interoceptive deficits in eating disorders: A meta-analysis of studies using the eating disorder inventory. *Journal of Psychosomatic Research*, *110*(April), 38–45. https://doi.org/10.1016/j.jpsychores.2018.04.005

Jonker, N. C., Glashouwer, K. A., Hoekzema, A., Ostafin, B. D., & Jong, P. J. D. (2020). Heightened self-reported punishment sensitivity, but no differential attention to cues signaling punishment or reward in anorexia nervosa. *PLoS ONE*, *15*(3), 1–17. https://doi.org/10.1371/journal.pone.0229742

KAUFMAN, J., BIRMAHER, B., BRENT, D., RAO, U. M. A., FLYNN, C., MORECI, P., WILLIAMSON, D., & RYAN, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial Reliability and Validity Data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(7), 980–988. https://doi.org/https://doi.org/10.1097/00004583-199707000-00021

Kaye, W. H., Bulik, C. M., Thornton, L., Barbarich, N., & Masters, K. (2004). Comorbidity of anxiety disorders with anorexia and bulimia nervosa. *American Journal of Psychiatry*, *161*(12), 2215–2221. https://doi.org/10.1176/appi.ajp.161.12.2215

Kaye, W. H., Fudge, J. L., & Paulus, M. (2009). New insights into symptoms and neurocircuit function of anorexia nervosa. *Nature Reviews Neuroscience*, *10*(8), 573–584. https://doi.org/10.1038/nrn2682

Kaye, W. H., Wierenga, C. E., Bailer, U. F., Simmons, A. N., & Bischoff-Grethe, A. (2013). Nothing tastes as good as skinny feels: The neurobiology of anorexia nervosa. In *Trends in Neurosciences* (Vol. 36, Issue 2, pp. 110–120). Elsevier. https://doi.org/10.1016/j.tins.2013.01.003

Keating, C., Tilbrook, A. J., Rossell, S. L., Enticott, P. G., & Fitzgerald, P. B. (2012). Reward processing in anorexia nervosa. *Neuropsychologia*, *50*(5), 567–575. https://doi.org/10.1016/j.neuropsychologia.2012.01.036

Keel, P. K., Kennedy, G. A., Rogers, M. L., Joyner, K. J., Bodell, L. P., Forney, K. J., & Duffy, M. E. (2022). Reliability and Validity of a Transdiagnostic Measure of Reward Valuation Effort. *Psychological Assessment*, *34*(5), 419–430. https://doi.org/10.1037/pas0001107

Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, *7*(5). https://doi.org/10.1371/journal.pcbi.1002055

Khalsa, S. S., Berner, L. A., & Anderson, L. M. (2022). Gastrointestinal Interoception in Eating Disorders: Charting a New Path. *Current Psychiatry Reports*, (0123456789). https://doi.org/10.1007/s11920-022-01318-3

Khalsa, S. S., Craske, M. G., Li, W., Vangala, S., Strober, M., & Feusner, J. D. (2015). Altered interoceptive awareness in anorexia nervosa: Effects of meal anticipation, consumption and bodily arousal. *International Journal of Eating Disorders*, *48*(7), 889–897. https://doi.org/10.1002/eat.22387

Khalsa, S. S., Hassanpour, M. S., Strober, M., Craske, M. G., Arevian, A. C., & Feusner, J. D. (2018). Interoceptive anxiety and body representation in anorexia nervosa. *Frontiers in Psychiatry*, *9*(SEP), 1–12. https://doi.org/10.3389/fpsyt.2018.00444

Khalsa, S. S., & Lapidus, R. C. (2016). Can interoception improve the pragmatic search for biomarkers in psychiatry? *Frontiers in Psychiatry*, *7*(JUL), 1–19. https://doi.org/10.3389/fpsyt.2016.00121

Klabunde, M., Collado, D., & Bohon, C. (2017). An interoceptive model of bulimia nervosa: A neurobiological systematic review. *Journal of Psychiatric Research*, *94*, 36–46. https://doi.org/10.1016/j.jpsychires.2017.06.009

Klump, K. L., Strober, M., Bulik, C. M., Thornton, L., Johnson, C., Devlin, B., Fichter, M. M., Halmi, K. A., Kaplan, A. S., Woodside, D. B., Crow, S., Mitchell, J., Rotondo, A., Kell, P. K., Berrettini, W. H., Plotnicov, K., Pollice, C., Lilenfeld, L. R., & Kaye, W. H. (2004). Personality characteristics of women before and after recovery from an eating disorder. *Psychological Medicine*, *34*(8), 1407–1418. https://doi.org/10.1017/S0033291704002442

Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1655), 20130474. https://doi.org/10.1098/rstb.2013.0474

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818. https://doi.org/10.1038/nn.3100

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, *25*(1), 178–206. https://doi.org/10.3758/s13423-016-1221-4

Kubanek, J., Snyder, L. H., & Abrams, R. A. (2015). Reward and punishment act as distinct factors in guiding behavior. *Cognition*, *139*, 154–167. https://doi.org/10.1016/j.cognition.2015.03.005

Lapidus, R. C., Puhl, M., Kuplicki, R., Stewart, J. L., Paulus, M. P., Rhudy, J. L., Feinstein, J. S., Khalsa, S. S., & Investigators, on behalf of the T. 1000. (2020). Heightened affective response to perturbation of respiratory but not pain signals in eating, mood, and anxiety disorders. *PLOS ONE*, *15*(7), e0235346. https://doi.org/10.1371/journal.pone.0235346

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, *1*(4), 1–9. https://doi.org/10.1038/s41562-017-0067

Levinson, C. A., & Williams, B. M. (2020). Eating disorder fear networks: Identification of central eating disorder fears. *International Journal of Eating Disorders*, *53*(12), 1960–1973. https://doi.org/10.1002/eat.23382

Li, J., Shi, C., Li, L., & Collins, A. G. E. (2024). Dynamic noise estimation: A generalized method for modeling noise fluctuations in decision-making. *Journal of Mathematical Psychology*, *119*, 102842. https://doi.org/10.1016/j.jmp.2024.102842

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences*, *104*(22), 9493–9498. https://doi.org/10.1073/pnas.0608842104

Makris, N., Goldstein, J. M., Kennedy, D., Hodge, S. M., Caviness, V. S., Faraone, S. V, Tsuang, M. T., & Seidman, L. J. (2006). Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia Research*, *83*(2), 155–171. https://doi.org/https://doi.org/10.1016/j.schres.2005.11.020

Marcus, D. S., Harms, M. P., Snyder, A. Z., Jenkinson, M., Wilson, J. A., Glasser, M. F., Barch, D. M., Archie, K. A., Burgess, G. C., Ramaratnam, M., Hodge, M., Horton, W., Herrick, R., Olsen, T., McKay, M., House, M., Hileman, M., Reid, E., Harwell, J., … Van Essen, D. C. (2013). Human Connectome Project informatics: Quality control, database services, and data visualization. *NeuroImage*, *80*, 202–219. https://doi.org/10.1016/j.neuroimage.2013.05.077

Martin, E., Dourish, C. T., Rotshtein, P., Spetter, M. S., & Higgs, S. (2019). Interoception and disordered eating: A systematic review. *Neuroscience and Biobehavioral Reviews*, *107*(August), 166–191. https://doi.org/10.1016/j.neubiorev.2019.08.020

Martinelli, M. K., Schreyer, C. C., Vanzhula, I. A., & Guarda, A. S. (2024). Impulsivity and reward and punishment sensitivity among patients admitted to a specialized inpatient eating disorder treatment program. *Frontiers in Psychiatry*, *15*. https://doi.org/10.3389/fpsyt.2024.1325252

Matton, A., de Jong, P., Goossens, L., Jonker, N., Van Malderen, E., Vervaet, M., De Schryver, N., & Braet, C. (2017). Sensitivity for Cues Predicting Reward and Punishment in Young Women with Eating Disorders. *European Eating Disorders Review*, *25*(6), 501–511. https://doi.org/10.1002/erv.2541

Mcdannald, M. A., Takahashi, Y. K., Lopatina, N., Pietras, B. W., Jones, J. L., & Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *European Journal of Neuroscience*, *35*(7), 991–996. https://doi.org/10.1111/j.1460-9568.2011.07982.x

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally Dissociable Influences on Learning Rate in a Dynamic Environment. *Neuron*, *84*(4), 870–881. https://doi.org/10.1016/j.neuron.2014.10.013

Mehling, W. E., Price, C., Daubenmier, J. J., Acree, M., Bartmess, E., & Stewart, A. (2012). The Multidimensional Assessment of Interoceptive Awareness (MAIA). *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0048230

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, *214*(5), 655–667. https://doi.org/10.1007/s00429-010-0262-0

Michely, J., Eldar, E., Erdman, A., Martin, I. M., & Dolan, R. J. (2022). Serotonin modulates asymmetric learning from reward and punishment in healthy human volunteers. *Communications Biology*, *5*(1), 1–9. https://doi.org/10.1038/s42003-022-03690-5

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2022). Value representations in the rodent orbitofrontal cortex drive learning, not choice. *eLife*, *11*, 1–27. https://doi.org/10.7554/eLife.64575

Mitchell, J. E., & Crow, S. (2006). Medical complications of anorexia nervosa and bulimia nervosa. *Current Opinion in Psychiatry*, *19*(4).

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1523/jneurosci.16-05-01936.1996

Monteleone, A. M., Monteleone, P., Esposito, F., Prinster, A., Volpe, U., Cantone, E., Pellegrino, F., Canna, A., Milano, W., Aiello, M., Di Salle, F., & Maj, M. (2017). Altered processing of rewarding and aversive basic taste stimuli in symptomatic women with anorexia nervosa and bulimia nervosa: An fMRI study. *Journal of Psychiatric Research*, *90*, 94–101. https://doi.org/10.1016/j.jpsychires.2017.02.013

Monteleone, P., Scognamiglio, P., Monteleone, A. M., Perillo, D., & Maj, M. (2014). Cortisol awakening response in patients with anorexia nervosa or bulimia nervosa: Relationships to sensitivity to reward and sensitivity to punishment. *Psychological Medicine*, *44*(12), 2653–2660. https://doi.org/10.1017/S0033291714000270

Morris, L. S., Kundu, P., Dowell, N., Mechelmans, D. J., Favre, P., Irvine, M. A., Robbins, T. W., Daw, N., Bullmore, E. T., Harrison, N. A., & Voon, V. (2016). Fronto-striatal organization: Defining functional and microstructural substrates of behavioural flexibility. *Cortex*, *74*, 118–133. https://doi.org/10.1016/j.cortex.2015.11.004

Moutoussis, M., Shahar, N., Hauser, T. U., & Dolan, R. J. (2018). Computation in Psychotherapy, or How Computational Psychiatry Can Aid Learning-Based Psychological Therapies. *Computational Psychiatry*, *2*(0), 50. https://doi.org/10.1162/cpsy_a_00014

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, *9*(11), 856–869. https://doi.org/10.1038/nrn2478

Nair, A., Rutledge, R. B., & Mason, L. (2020). Under the Hood: Using Computational Psychiatry to Make Psychological Therapies More Mechanism-Focused. *Frontiers in Psychiatry*, *11*(March), 1–7. https://doi.org/10.3389/fpsyt.2020.00140

Neveu, R., Fouragnan, E., Barsumian, F., Carrier, E., Lai, M., Nicolas, A., Neveu, D., & Coricelli, G. (2016). Preference for safe over risky options in binge eating. *Frontiers in Behavioral Neuroscience*, *10*(MAR), 1–9. https://doi.org/10.3389/fnbeh.2016.00065

Newman, M. G., & Llera, S. J. (2011). A novel theory of experiential avoidance in generalized anxiety disorder: A review and synthesis of research supporting a contrast avoidance model of worry. *Clinical Psychology Review*, *31*(3), 371–382. https://doi.org/10.1016/j.cpr.2011.01.008

Nieto-Castanon, A. (2020). *Handbook of functional connectivity Magnetic Resonance Imaging methods in CONN*. Hilbert Press.

Nieto-Castanon, A., & Whitfield-Gabrieli, S. (2022). *CONN functional connectivity toolbox: RRID SCR_009550, release 22* (22nd ed.). Hilbert Press. https://doi.org/10.56441/hilbertpress.2246.5840

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*, *32*(2), 551–562. https://doi.org/10.1523/JNEUROSCI.5498-10.2012

Niv, Y., Hitchcock, P., Berwian, I. M., & Schoen, G. (2021). Toward Precision Cognitive-Behavioral Therapy via Reinforcement Learning Theory. In L. M. Williams & L. M. Hack (Eds.), *Precision Psychiatry: Using Neuroscience Insights to Inform Personally Tailored, Measurement-Based Care* (1st ed., p. 199). American Psychiatric Pub.

Norem, J. K., & Cantor, N. (1986). Defensive Pessimism. Harnessing Anxiety as Motivation. *Journal of Personality and Social Psychology*, *51*(6), 1208–1217. https://doi.org/10.1037/0022-3514.51.6.1208

Nussenbaum, K., & Hartley, C. A. (2024). Understanding the development of reward learning through the lens of meta-learning. *Nature Reviews Psychology*, *3*(6), 424–438. https://doi.org/10.1038/s44159-024-00304-1

Oberndorfer, T., Simmons, A., McCurdy, D., Strigo, I., Matthews, S., Yang, T., Irvine, Z., & Kaye, W. (2013). Greater anterior insula activation during anticipation of food images in women recovered from anorexia nervosa versus controls. *Psychiatry Research - Neuroimaging*, *214*(2), 132–141. https://doi.org/10.1016/j.pscychresns.2013.06.010

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249. https://doi.org/10.1016/j.tics.2005.03.010

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., & Gross, J. J. (2004). For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage*, *23*(2), 483–499. https://doi.org/10.1016/j.neuroimage.2004.06.030

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. https://doi.org/10.1038/82959

O'Doherty, J. P., Lee, S. W., & McNamee, D. (2015). The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, *1*, 94–100. https://doi.org/10.1016/j.cobeha.2014.10.004

O'Hara, C. B., Campbell, I. C., & Schmidt, U. (2015). A reward-centered model of anorexia nervosa: A focused narrative review of the neurological and psychophysiological literature. *Neuroscience and Biobehavioral Reviews*, *52*, 131–152. https://doi.org/10.1016/j.neubiorev.2015.02.012

Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The Curse of Planning: Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive. *Psychological Science*, *24*(5), 751–761. https://doi.org/10.1177/0956797612463080

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 20941–20946. https://doi.org/10.1073/pnas.1312011110

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *Journal of Cognitive Neuroscience*, *27*(2), 319–333. https://doi.org/10.1162/jocn_a_00709

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*(2), 97–98. https://doi.org/10.1038/nn802

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., Durr, A., & Pessiglione, M. (2012). Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning. *Neuron*, *76*(5), 998–1009. https://doi.org/10.1016/j.neuron.2012.10.017

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms9096

Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S. J. (2016). The Computational Development of Reinforcement Learning during Adolescence. *PLoS Computational Biology*, *12*(6), 1–25. https://doi.org/10.1371/journal.pcbi.1004953

Palminteri, S., & Lebreton, M. (2022). The computational roots of positivity and confirmation biases in reinforcement learning. *Trends in Cognitive Sciences*, *26*(7), 607–621. https://doi.org/10.1016/j.tics.2022.04.005

Patterson, T. K., & Knowlton, B. J. (2018). Subregional specificity in human striatal habit learning: A meta-analytic review of the fMRI literature. *Current Opinion in Behavioral Sciences*, *20*, 75–82. https://doi.org/10.1016/j.cobeha.2017.10.005

Paulus, M. P. (2020). Driven by Pain, Not Gain: Computational Approaches to Aversion-Related Decision Making in Psychiatry. *Biological Psychiatry*, *87*(4), 359–367. https://doi.org/10.1016/j.biopsych.2019.08.025

Paulus, M. P., & Stein, M. B. (2006). An Insular View of Anxiety. *Biological Psychiatry*, *60*(4), 383–387. https://doi.org/10.1016/j.biopsych.2006.03.042

Paulus, M. P., & Stein, M. B. (2010). Interoception in anxiety and depression. *Brain Structure & Function*, *214*(5–6), 451–463. https://doi.org/10.1007/s00429-010-0258-9

Pearce, A. L., Fuchs, B. A., & Keller, K. L. (2022). The role of reinforcement learning and value-based decision-making frameworks in understanding food choice and eating behaviors. *Frontiers in Nutrition*, *9*(November), 1–13. https://doi.org/10.3389/fnut.2022.1021868

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier.

Petzschner, F. H., Garfinkel, S. N., Paulus, M. P., Koch, C., & Khalsa, S. S. (2021). Computational Models of Interoception and Body Regulation. *Trends in Neurosciences, Special Issue: The Neuroscience of Interoception*, *44*(1), 63–76. https://doi.org/10.1016/j.tins.2020.09.012

Pezzulo, G., Rigoli, F., & Chersi, F. (2013). The mixed instrumental controller: Using value of information to combine habitual choice and mental simulation. *Frontiers in Psychology*, *4*(MAR), 1–15. https://doi.org/10.3389/fpsyg.2013.00092

Pike, A. C., & Robinson, O. J. (2022). Reinforcement Learning in Patients with Mood and Anxiety Disorders vs Control Individuals: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *79*(4), 313–322. https://doi.org/10.1001/jamapsychiatry.2022.0051

Potter, T. C. S., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience*, *25*, 272–280. https://doi.org/10.1016/j.dcn.2016.10.005

R Development Core Team. (2015). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https://doi.org/10.1007/978-1-4020-6850-8

Radzikowska, M., Pike, A. C., & Hall-McMaster, S. (2025). Computational Perspectives on Cognition in Anorexia Nervosa: A Systematic Review. *Computational Psychiatry*, *9*(1), 100–121. https://doi.org/10.5334/cpsy.128

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357

Redish, A. D., & Gordon, J. A. (2016). *Computational psychiatry: New perspectives on mental illness* . The MIT Press.

Reilly, E. E., Brown, T. A., DeJesus, C. R., Kaye, W. H., & Wierenga, C. E. (2025). Exploring Reciprocal Associations Between Self-Reported Anxiety and Eating Disorder Symptoms Longitudinally: A Bivariate Latent Change Score Approach. *International Journal of Eating Disorders*, *58*(2), 336–348. https://doi.org/10.1002/eat.24329

Reilly, E. E., Lavender, J. M., Berner, L. A., Brown, T. A., Wierenga, C. E., & Kaye, W. H. (2019). Could repetitive negative thinking interfere with corrective learning? The

example of anorexia nervosa. *International Journal of Eating Disorders*, *52*(1), 36–41. https://doi.org/10.1002/eat.22997

Reilly, E. E., Rockwell, R. E., Ramirez, A. L., Anderson, L. K., Brown, T. A., Wierenga, C. E., & Kaye, W. H. (2020). Naturalistic outcomes for a day-hospital programme in a mixed diagnostic sample of adolescents with eating disorders. *European Eating Disorders Review*, *28*(2), 199–210. https://doi.org/10.1002/erv.2716

Reiter, A. M. F., Heinze, H. J., Schlagenhauf, F., & Deserno, L. (2017). Impaired Flexible Reward-Based Decision-Making in Binge Eating Disorder: Evidence from Computational Modeling and Functional Neuroimaging. *Neuropsychopharmacology*, *42*(3), 628–637. https://doi.org/10.1038/npp.2016.95

Ritschel, F., Geisler, D., King, J. A., Bernardoni, F., Seidel, M., Boehm, I., Vettermann, R., Biemann, R., Roessner, V., Smolka, M. N., & Ehrlich, S. (2017). Neural correlates of altered feedback learning in women recovered from anorexia nervosa. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-04761-y

Rosenbaum, G. M., Grassie, H. L., & Hartley, C. A. (2022). Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. *eLife*, *11*, 1–36. https://doi.org/10.7554/eLife.64620

Sala, M., Brosof, L. C., & Levinson, C. A. (2019). Repetitive negative thinking predicts eating disorder behaviors: A pilot ecological momentary assessment study in a treatment seeking eating disorder sample. *Behaviour Research and Therapy*, *112*, 12–17. https://doi.org/10.1016/j.brat.2018.11.005

Sala, M., & Levinson, C. A. (2016). The longitudinal relationship between worry and disordered eating: Is worry a precursor or consequence of disordered eating? *Eating Behaviors*, *23*, 28–32. https://doi.org/10.1016/j.eatbeh.2016.07.012

Santo-Angles, A., Fuentes-Claramonte, P., Argila-Plaza, I., Guardiola-Ripoll, M., Almodóvar-Payá, C., Munuera, J., McKenna, P. J., Pomarol-Clotet, E., & Radua, J. (2021). Reward and fictive prediction error signals in ventral striatum: Asymmetry between factual and counterfactual processing. *Brain Structure and Function*, *226*(5), 1553–1569. https://doi.org/10.1007/s00429-021-02270-3

Schaefer, L. M., & Steinglass, J. E. (2021). Reward Learning Through the Lens of RDoC: a Review of Theory, Assessment, and Empirical Findings in the Eating Disorders. *Current Psychiatry Reports*, *23*(1), 14–17. https://doi.org/10.1007/s11920-020-01213-9

Schaumberg, K., Reilly, E. E., Gorrell, S., Levinson, C. A., Farrell, N. R., Brown, T. A., Smith, K. M., Schaefer, L. M., Essayli, J. H., Haynos, A. F., & Anderson, L. M. (2021). Conceptualizing eating disorder psychopathology using an anxiety disorders framework: Evidence and implications for exposure-based clinical research. *Clinical Psychology Review*, *83*(November 2020), 101952. https://doi.org/10.1016/j.cpr.2020.101952

Scholz, V., Waltmann, M., Herzog, N., Reiter, A., Horstmann, A., & Deserno, L. (2023). Cortical grey matter mediatesincreases in model-based control and learning from positive feedback from adolescence to adulthood. *The Journal of Neuroscience*, *43*(12), JN-RM-1418-22. https://doi.org/10.1523/jneurosci.1418-22.2023

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565–573. https://doi.org/10.1016/j.tics.2013.09.007

Seth, A. K., & Critchley, H. D. (2013). Extending predictive processing to the body: Emotion as interoceptive inference. *Behavioral and Brain Sciences*, *36*(3), 227–228. https://doi.org/10.1017/S0140525X12002270

Shott, M. E., Filoteo, J. V., Jappe, L. M., Pryor, T., Maddox, W. T., Rollin, M. D. H., Hagman, J. O., & Frank, G. K. W. (2012). Altered implicit category learning in anorexia nervosa. *Neuropsychology*, *26*(2), 191–201. https://doi.org/10.1037/a0026771

Simmons, W. K., Avery, J. A., Barcalow, J. C., Bodurka, J., Drevets, W. C., & Bellgowan, P. (2013). Keeping the body in mind: Insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Human Brain Mapping*, *34*(11), 2944–2958. https://doi.org/10.1002/hbm.22113

Smith, R., Kuplicki, R., Feinstein, J., Forthman, K. L., Stewart, J. L., Paulus, M. P., Investigators, T. 1000, & Khalsa, S. S. (2020). A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLOS Computational Biology*, *16*(12), e1008484.

Spielberger, C. D., & Vagg, P. R. (1984). Psychometric Properties of the STAI: A Reply to Ramanaiah, Franzen, and Schill. *Journal of Personality Assessment*, *48*(1), 95–97. https://doi.org/10.1207/s15327752jpa4801_16

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature Neuroscience*, *18*(5), 620–627. https://doi.org/10.1038/nn.3982

Steding, J., Boehm, I., King, J. A., Geisler, D., Ritschel, F., Seidel, M., Doose, A., Jaite, C., Roessner, V., Smolka, M. N., & Ehrlich, S. (2019). Goal-directed vs. Habitual instrumental behavior during reward processing in anorexia nervosa: An fMRI study. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-49884-6

Steinglass, J. E., & Walsh, B. T. (2016). Neurobiological model of the persistence of anorexia nervosa. *Journal of Eating Disorders*, *4*(1), 1–7. https://doi.org/10.1186/s40337-016-0106-2

Steinglass, J., & Walsh, B. T. (2006). Habit learning and anorexia nervosa: A cognitive neuroscience hypothesis. *The International Journal of Eating Disorders*, *39*(4), 267–275. https://doi.org/10.1002/eat.20244

Steward, T., Mestre-Bach, G., Vintró-Alcaraz, C., Agüera, Z., Jiménez-Murcia, S., Granero, R., & Fernández-Aranda, F. (2017). Delay Discounting of Reward and Impulsivity in Eating Disorders: From Anorexia Nervosa to Binge Eating Disorder. *European Eating Disorders Review*, *25*(6), 601–606. https://doi.org/10.1002/erv.2543

Stuss, D. T., & Levine, B. (2002). Adult Clinical Neuropsychology: Lessons from Studies of the Frontal Lobes. *Annual Review of Psychology*, *53*(Volume 53, 2002), 401–433. https://doi.org/10.1146/annurev.psych.53.100901.135220

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning an introduction*. MIT Press.

Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A., & Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nature Neuroscience*, *15*(9), 1281–1289. https://doi.org/10.1038/nn.3188

Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human pavlovian-instrumental transfer. *Journal of Neuroscience*, *28*(2), 360–368. https://doi.org/10.1523/JNEUROSCI.4028-07.2008

Tolman, E. C. (1948). Cognitive maps in rats and men. In *Psychological Review* (Vol. 55, pp. 189–208). American Psychological Association. https://doi.org/10.1037/h0061626

Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences*, *31*(6), 837–862. https://doi.org/https://doi.org/10.1016/S0191-8869(00)00183-5

Touroutoglou, A., Hollenbeck, M., Dickerson, B. C., & Barrett, L. F. (2012). Dissociable large-scale networks anchored in the right anterior insula subserve affective experience and attention. *NeuroImage*, *60*(4), 1947–1958. https://doi.org/10.1016/j.neuroimage.2012.02.012

Uniacke, B., Timothy Walsh, B., Foerde, K., & Steinglass, J. (2018). The Role of Habits in Anorexia Nervosa: Where We Are and Where to Go From Here? *Current Psychiatry Reports*, *20*(8). https://doi.org/10.1007/s11920-018-0928-5

Vaghi, M. M., Moutoussis, M., Váša, F., Kievit, R. A., Hauser, T. U., Vértes, P. E., Shahar, N., Romero-Garcia, R., Kitzbichler, M. G., Bullmore, E. T., & Dolan, R. J. (2020). Compulsivity is linked to reduced adolescent development of goal-directed control and frontostriatal functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(41), 25911–25922. https://doi.org/10.1073/pnas.1922273117

Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience*, *27*(15), 4019–4026. https://doi.org/10.1523/JNEUROSCI.0564-07.2007

Van den Bergh, O., Witthöft, M., Petersen, S., & Brown, R. J. (2017). Symptoms and the body: Taking the inferential leap. *Neuroscience and Biobehavioral Reviews*, *74*, 185–203. https://doi.org/10.1016/j.neubiorev.2017.01.015

Van der Meer, M. A. A., & Redish, A. D. (2011). Ventral striatum: A critical look at models of learning and evaluation. *Current Opinion in Neurobiology*, *21*(3), 387–392. https://doi.org/10.1016/j.conb.2011.02.011

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC (with Discussion). *Bayesian Analysis*, *16*(2), 667–718. https://doi.org/10.1214/20-BA1221

Verharen, J. P. H., Danner, U. N., Schröder, S., Aarts, E., Elburg, A. A. van, & Adan, R. A. H. (2019). Insensitivity to Losses: A Core Feature in Patients With Anorexia Nervosa? *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(11), 995–1003. https://doi.org/10.1016/j.bpsc.2019.05.001

Villano, W. J., Kraus, N. I., Reneau, T. R., Jaso, B. A., Ross Otto, A., & Heller, A. S. (2023). Individual differences in naturalistic learning link negative emotionality to the development of anxiety. *Science Advances*, *9*(1), 1–16. https://doi.org/10.1126/sciadv.add2976

Voon, V. (2015). Cognitive biases in binge eating disorder: The hijacking of decision making. *CNS Spectrums*, *20*(6), 566–573. https://doi.org/10.1017/S1092852915000681

Voon, V., Baek, K., Enander, J., Worbe, Y., Morris, L. S., Harrison, N. A., Robbins, T. W., Rück, C., & Daw, N. (2015). Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Translational Psychiatry*, *5*(11), e670–e670. https://doi.org/10.1038/tp.2015.165

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W., Harrison, N. A., Wood, J., Daw, N. D., Dayan, P., Grant, J. E., & Bullmore, E. T. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry*, *20*(3), 345–352. https://doi.org/10.1038/mp.2014.44

Wächter, T., Lungu, O. V., Liu, T., Willingham, D. T., & Ashe, J. (2009). Differential effect of reward and punishment on procedural learning. *Journal of Neuroscience*, *29*(2), 436–443. https://doi.org/10.1523/JNEUROSCI.4132-08.2009

Wagner, A., Aizenstein, H., Mazurkewicz, L., Fudge, J., Frank, G. K., Putnam, K., Bailer, U. F., Fischer, L., & Kaye, W. H. (2008). Altered insula response to taste stimuli in individuals recovered from restricting-type anorexia nervosa. *Neuropsychopharmacology*, *33*(3), 513–523. https://doi.org/10.1038/sj.npp.1301443

Wagner, A., Aizenstein, H., Venkatraman, V. K., Bischoff-Grethe, A., Fudge, J., May, J. C., Frank, G. K., Bailer, U. F., Fischer, L., Putnam, K., & Kaye, W. H. (2010). Altered striatal response to reward in bulimia nervosa after recovery. *International Journal of Eating Disorders*, *43*(4), 289–294. https://doi.org/10.1002/eat.20699

Walsh, B. T. (2013). The enigmatic persistence of anorexia nervosa. *American Journal of Psychiatry*, *170*(5), 477–484. https://doi.org/10.1176/appi.ajp.2012.12081074

Wan Lee, S., Shimojo, S., & O'Doherty, J. P. (2014). Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. *Neuron*, *81*(3), 687–699. https://doi.org/10.1016/j.neuron.2013.11.028

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, *2*(3), 125–141. https://doi.org/10.1089/brain.2012.0073

Wierenga, C. E., Bischoff-Grethe, A., Berner, L. A., Simmons, A. N., Bailer, U., Paulus, M. P., & Kaye, W. H. (2020). Increased anticipatory brain response to pleasant touch in women remitted from bulimia nervosa. *Translational Psychiatry*, *10*(236), 1–14. https://doi.org/10.1038/s41398-020-00916-0

Wierenga, C. E., Bischoff-Grethe, A., Brown, C. S., & Brown, G. G. (2025). Reinforcement learning in women remitted from anorexia nervosa: Preliminary examination with a hybrid reinforcement learning/drift diffusion model. *Journal of the International Neuropsychological Society*, *31*(2), 127–137. https://doi.org/10.1017/S1355617725000013

Wierenga, C. E., Brown, C. S., & Reilly, E. E. (2025). Reinforcement Learning and Decision Making in Anorexia Nervosa. *Current Psychiatry Reports*, *27*(12), 711–722. https://doi.org/10.1007/s11920-025-01643-3

Wierenga, C. E., Ely, A., Bischoff-Grethe, A., Bailer, U. F., Simmons, A. N., & Kaye, W. H. (2014). Are extremes of consumption in eating disorders related to an altered balance between reward and inhibition? *Frontiers in Behavioral Neuroscience*, *8*, 410. https://doi.org/10.3389/fnbeh.2014.00410

Wierenga, C. E., Reilly, E., Bischoff-Grethe, A., Kaye, W. H., & Brown, G. G. (2022). Altered Reinforcement Learning from Reward and Punishment in Anorexia Nervosa: Evidence from Computational Modeling. *Journal of the International Neuropsychological Society*, *28*(10), 1003–1015. Cambridge Core. https://doi.org/10.1017/S1355617721001326

Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, *17*(8), 513–523. https://doi.org/10.1038/nrn.2016.56

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081. https://doi.org/10.1037/a0038199

Worbe, Y., Palminteri, S., Savulich, G., Daw, N. D., Fernandez-Egea, E., Robbins, T. W., & Voon, V. (2016). Valence-dependent influence of serotonin depletion on model-based choice strategy. *Molecular Psychiatry*, *21*(5), 624–629. https://doi.org/10.1038/mp.2015.46

Wu, M., Brockmeyer, T., Hartmann, M., Skunde, M., Herzog, W., & Friederich, H. C. (2016). Reward-related decision making in eating and weight disorders: A systematic review and meta-analysis of the evidence from neuropsychological studies. *Neuroscience and Biobehavioral Reviews*, *61*, 177–196. https://doi.org/10.1016/j.neubiorev.2015.11.017

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin*, *139*(2), 497–518. https://doi.org/10.1037/a0029383

Zhang, Y., Paik, J., & Pirolli, P. (2015). Reinforcement Learning and Counterfactual Reasoning Explain Adaptive Behavior in a Changing Environment. *Topics in Cognitive Science*, *7*(2), 368–381. https://doi.org/10.1111/tops.12143

Zucker, N. L., & Bulik, C. M. (2020). On bells, saliva, and abdominal pain or discomfort: Early aversive visceral conditioning and vulnerability for anorexia nervosa. *International Journal of Eating Disorders*, *53*(4), 508–512. https://doi.org/10.1002/eat.23255