UCLA
UCLA Electronic Theses and Dissertations

Title
Identifying decision-making and reinforcement learning deficits in psychosis: Clinical, neural and transdiagnostic implications

Permalink
https://escholarship.org/uc/item/1s96246w

Author
Patel, Pooja Kirit

Publication Date
2023

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Identifying decision-making and reinforcement learning deficits in psychosis: Clinical, neural and transdiagnostic implications

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

by

Pooja Kirit Patel

2023

© Copyright by

Pooja Kirit Patel

2023

ABSTRACT OF THE DISSERTATION

Identifying decision-making and reinforcement learning deficits in psychosis: Clinical, neural and transdiagnostic implications

by

Pooja Kirit Patel

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2023

Professor Katherine H. Karlsgodt, Chair

In schizophrenia (SZ) and psychotic illnesses, negative symptoms (i.e., avolition, anhedonia) contribute to profound social and role impairment and are largely unresponsive to existing pharmacological and psychotherapeutic interventions. One specific process, reinforcement learning (RL), defined as mapping outcomes to certain actions to guide decision-making and behavior based on feedback, has been repeatedly implicated in the etiology of negative symptoms in psychotic illness. Some evidence suggests that schizophrenia is characterized by difficulty learning from positive but not negative feedback, deficits in learning initial associations between stimuli and certain outcomes, and deficits in making decisions under ambiguity (i.e., when probabilities of adverse outcomes are unknown). However, existing work is limited by contradictory findings about whether initial learning of associations is impaired, modeling methods that do not fully account for asymmetries in learning, and inconsistent evidence linking deficits to actual symptomatology in participants.

The goal of this dissertation was to address these limitations in the literature and rigorously investigate negative symptoms, reward-guided decision-making and RL deficits in psychotic illness. I endeavored to characterize moderators of deficit severity and symptom severity across the full spectrum of psychotic presentations. To this end, I adopted a dimensional approach that ensured variability in patient samples through use of a psychosis spectrum sample in Study 1, investigation of possible shared and distinct RL deficits in schizophrenia and bipolar disorder in Study 2, and exploration of how white matter integrity in the brain may be a meaningful predictor of variability in RL in Study 3. In Study 1, I demonstrated that when making decisions under ambiguity, individuals with psychosis can learn to differentiate high risk/low reward from low risk/high reward contexts; however, severity of negative symptoms is associated with a failure to maximize rewards in low-risk situations. In Study 2, I employed a computational RL model that accounts for asymmetries in integrating positive and negative feedback, as well as retention of the values of specific choices over time. While individuals with psychosis are seemingly acquiring initial associations, there appear to be differences in how they use feedback to modify future behaviors. Negative symptoms moderate this difference, such that increased severity is associated greater weighting of negative feedback and lesser weighting of positive feedback. In Study 3, I explored the relationship between computational RL parameters from Study 2 and white matter connectivity in frontoparietal and corticostriatal circuits, two RL-associated circuits; though I did not find any associations between RL parameters and structural brain connectivity, I highlight the need to provide other relevant circuits and provide avenues for future investigation of neural contributions to reinforcement learning. The link between the work presented in this dissertation and broader implications for etiological frameworks for psychotic illness is also discussed.

The dissertation of Pooja Kirit Patel is approved.

Carrie E. Bearden

Melissa J. Sharpe

Michael F. Green

Katherine H. Karlsgodt, Committee Chair

University of California, Los Angeles

2023

This dissertation is dedicated to the brilliant scholars who came before me, particularly those who never saw a career for themselves in the sciences but forged their own paths—their foundational contributions to science, their resilience, and their creativity is a source of endless inspiration. Their hard-won battles gave me the opportunity to dare to dream and pursue my passion for science.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

TABLES **PAGE**

FIGURES

ACKNOWLEDGMENTS

This completion of this dissertation would not be possible without the support of many wonderful people. First and foremost, I would like to thank my advisor Dr. Katherine Karlsgodt for being my mentor, supporting me in being flexible and adaptive in the pursuit of high-quality science, and giving me the opportunity to grow as a scholar. The members of my dissertation committee Drs. Bearden, Sharpe and Green have been unfailingly patient and thoughtful in offering their constructive feedback to strengthen this work. I would also like to acknowledge the support of Dr. Stephanie Groman at University of Minnesota, without whom the innovative analyses in this dissertation would not be possible.

I am grateful for the friendship and guidance from all those in the Clinical and Cognitive Neuroscience Lab: Mary Rshtouni, Haley Wang, Rachel McKinney, Logan Leathem, and Danielle Currin. I cannot imagine succeeding in graduate school without their emotional and technical support, and kindness. I have been lucky enough to find lifelong friends in graduate school, including Christina Hough, Brittany Drake, Julia Schorn, Kelsey Stiles, Sophie Arkin, and Dr. Joao Guassi Moreira. I feel so thankful to have found academic peers who will push me to succeed and inspire me for years to come.

I am indebted to the teams involved in the projects that provided data for this work: Dr. Karlsgodt's team at Zucker Hillside Hospital (Melanie Blair-Thies, Ashley Moyett, Dr. Pamela DeRosse) and the UCLA Consortium for Neuropsychiatric Phenomics (CNP). In particular, I would also like to thank the research participants, many of whom were navigating serious mental health challenges while participating in research protocols. Without their commitment and time, this work would not have been possible.

I am grateful to the friends and family who supported me both before and throughout graduate school. My sisters, Sonali and Rupali, are my ultimate role models, my guideposts, and my fiercest advocates. My parents, Kirit and Yasmin, instilled in me an immense

resilience and tenacity for life— without them I never would have pushed myself out of my comfort zone and pursued graduate school. Navigating the challenges of graduate school would not have been possible without my chosen family— my twin soul and champion Angela Sardo, those who have been by me since childhood (Ava Calvano, Maria LaVerghetta, Emily Cheung-Hynson), those who have helped me weather seasons of chaos (Dee Lowe, Abby Wilson), and those who have helped Los Angeles feel like home (Sophie Kelly, Erin Kaufmann, Sarah Beth Tyrey). You are all the pillars upon which I stand, and I can never thank you enough.

This research was supported by the following: NIMH 1R01MH101506 Multimodal Imagine of Executive and Reward Networks across the Psychosis Spectrum (PI: Karlsgodt), NICHD 1T32HD091059 Brain and Behavioral Development during Adolescence (PIs: Fuligni, Galvan, Dapretto), and the Consortium for Neuropsychiatric Phenomics Coordinating Center UL1-DE019580 and Human Translational Applications (HTAC) core 1PLMH0832271-01 (PI: Bilder).

VITA | POOJA K. PATEL

ACADEMIC AND EMPLOYMENT HISTORY

2015	University of Maryland, College Park B.S. in Biology, Physiology and Neurobiology Concentration
2015	University of Maryland, College Park B.S. in Psychology
2015-2017	University of California, Davis Medical Center Translational Cognitive and Affective Neuroscience Lab Research Assistant
2018	University of California, Los Angeles M.A. in Psychology
2023	Didi Hirsch Mental Health Services, Adult Outpatient APA Clinical Internship

AWARDS, FELLOWSHIPS, AND HONORS

2018	UCLA Graduate Summer Research Mentorship Fellowship
2018-2020	NICHD T32 (1T32HD091059) Brain and Behavior Development during Adolescence
2019	UCLA Psychology Department Stanley Sue Distinguished Research Award

PUBLICATIONS

Currin, D.L., Hart, K.P., Gupta, M.W., **Patel, P.K**., Leathem, L., & Karlsgodt, K.H. (2022). The role of anhedonia in predicting risk-taking behavior in university students. *Journal of Psychiatric Research*. Doi.org/IO.1016/j.jpsychires.2022.09.030

Patel, P.K., Leathem, L.D., Currin, D.L., & Karlsgodt, K.H. (2020). Adolescent Development and Vulnerability to Psychosis. *Biological Psychiatry*. Doi:10.1016/j.biopsych.2020.06.028

Rosenthal, A., Mayo, D., Tully, L., **Patel, P**.**K.,** Ashby, S., Titone, M., DeLeon, E., Meyer, M., Carter, C., Niendam, T. (2019) Contributions of childhood trauma and atypical development to increased clinical symptoms and poor functioning in recent onset psychosis. *Early Intervention in Psychiatry*.

Lopez Garcia, P., Ashby, S., **Patel, P.K**., Pierce, K.M., Meyer, M., Rosenthal, A., Titone, M., Carter, C.S. & Niendam, T.A. (2019). Clinical and Neurodevelopmental Correlates of Aggression in Early Psychosis. *Schizophrenia Research*. doi: 10.1016/j.schres.2019.07.045

Savill, M., Gosdin, M., **Patel, P.K.,** Melnikow, J., Loewy, R. & Niendam, T. (2018). How do early psychosis services define and operationalize the duration of untreated psychosis? *The Journal of Behavioral Health Services & Research*. doi:10.1007/s11414-018-9630-y

Niendam, T.A., Sardo, A., Savill, M., **Patel, P.K**., Xing, G., Loewy, R., Dewa, C. & Melnikow, J. (2019). The rise of early psychosis care in California: An overview of community and university-based services. *Psychiatric Services*. doi:10.1176/appi.ps.201800394

Niendam, T.A., Ray, K.L., Iosif, A.M., Lesh, T.A., Ashby, S.R., **Patel, P.K**., Smucny, J., Ferrer, E., Solomon, M., Ragland, J.D. & Carter, C.S. (2018) Association of Age at Onset and Longitudinal Course of Prefrontal Function in Youth with Schizophrenia. *JAMA Psychiatry* doi:10.1001/jamapsychiatry.2018.2538

Savill M., Sardo A., **Patel, P.K.**, Loewy R., Melnikow J. & Niendam T.A. (2018) Which components of coordinated specialty care for first episode psychosis do program managers see as most important? *Early Intervention in Psychiatry*. doi:10.1111/eip.12690

Burton, A.C., Bissonette G.B., Zhao, A.C., **Patel, P.K.** & Roesch, M. R. (2017). Prior Cocaine Self-Administration Increases Response—Outcome Encoding That Is Divorced from Actions Selected in Dorsal Lateral Striatum. *Journal of Neuroscience*. Doi: 10.1523/JNEUROSCI.089717.2017

SELECTED PRESENTATIONS

Patel, P.K., Groman, S., & Karlsgodt, K.H. Transdiagnostic Computational Modeling of Reinforcement Learning Deficits. Poster Session Presented at: Society of Biological Psychiatry. Annual Conference; 2022 Apr 28-30 ; New Orleans, LA

Patel, P.K., Groman, S., & Karlsgodt, K.H. Reinforcement learning deficits in bipolar disorder and schizophrenia. Accepted as Poster at: Society for Research in Psychopathology. Annual Conference ; 2021 Sept 23-26; Virtual due to COVID-19 Pandemic

Patel, P.K., Blair, M.A., Moyett, A., Arkin, S., DeRosse, P. & Karlsgodt, K.H. Neural and Clinical Correlates of Optimal Risk-Taking in Early Psychosis. Poster Session Presented at: Society of Biological Psychiatry. 74th Annual Conference ; 2019 May 16th-18th ; Chicago, IL

Arkin, S.A., Moyett, A., Blair, M.A., **Patel, P.K.,** Arkin .S., DeRosse, P. & Karlsgodt, K.H. Load Dependent Working Memory Circuitry in Early-Onset Psychosis. Poster Session Presented at: Society of Biological Psychiatry. 74th Annual Conference ; 2019 May 16th-18th; Chicago, IL

Patel, P.K., Blair, M.A., Moyett, A., Arkin, S., DeRosse, P. & Karlsgodt, K.H. Neural and Clinical Correlates of Optimal Risk-Taking in Early Psychosis. Accepted as Poster at: Society of Biological Psychiatry. 74th Annual Conference; 2019 May 16th-18th; Chicago, IL

Patel, P.K., Blair, M.A., Moyett, A., DeRosse, P. & Karlsgodt, K.H. Optimizing risk-taking during the Balloon Analogue Risk Task: Impact of avolition in early psychosis. Poster Session Presented at: Society for Research in Psychopathology. 32nd Annual Conference; 2018 Sept 21st-23rd; Indianapolis, IL

CHAPTER 1

BACKGROUND

Psychotic disorders are characterized by a constellation of complex symptoms including positive symptoms (e.g., delusions, hallucinations, disorganized thought content) and negative symptoms (e.g., anhedonia, avolition) as well as cognitive deficits. Despite the advent of second-generation antipsychotics and psychosis-specific psychosocial interventions, economic and functional outcomes for psychotic illness remain remarkably poor (Addington, Leriger, & Addington, 2003; Pencer, Addington, & Addington, 2005). Negative symptoms continue to impact outcomes even when positive symptoms are managed by pharmacotherapy (Fervaha, Foussias, Agid, & Remington, 2014; Fusar-Poli et al., 2014). The ability to seek and maintain gainful employment and cultivate meaningful relationships can be substantially disrupted by pronounced negative symptoms. Economic burden due to psychosis-related unemployment, suicide, and need for caregiving is in excess of $150 billion in the United States(Cloutier et al., 2016). Interventions designed to reduce negative symptoms would impact ability to seek and maintain gainful employment and engage in rewarding social activities, which could then limit economic burden and improve functional outcomes. However, creating such interventions requires clarification of mechanisms and identification of relevant clinical and neural moderators of functioning and outcome.

Goal-Directed Behavior in Psychosis: Reward Processing Deficits

Given that negative symptoms have long been an unmet treatment need in the psychosis population, understanding factors contributing to negative symptoms has become a high priority. Thus, there is a considerable body of literature focused on identifying how

differences in reward-related learning and decision-making may cause negative symptomatology. Here, I discuss the broader literature on reward processing to provide context for current understand of concepts such as punishment sensitivity, positive feedback sensitivity, and effort expenditure to underscore the importance of examining reinforcement learning as a relevant mechanism underlying both intact and deficit processes observed in individuals with schizophrenia and related conditions.

Reward processing refers to several subcomponent processes dependent on both reward and executive function processes that are critical for decision-making and action selection. Researchers have investigated different aspects of reward processing and sought to link them with negative symptoms. These efforts have employed a wide variety of tasks, paradigms, and theoretical models of processes such as associative learning, response to punishment or non-reward, monetary gain, valuation of effort expended to obtain reward, use of past reward information to guide future behaviors, and use of reward-related feedback when making decisions under ambiguity.

Prior to engaging in action selection, an individual has a certain degree of responsiveness to reward (hedonics) and capacity to predict reward (reward anticipation). Several studies have focused on hedonics and reward anticipation, typically using Monetary Incentive Delay style tasks in which participants are presented with cues that are associated with certain amounts of monetary reward, are required to make a choice related to the cues, and then receive immediate feedback about whether they receive a reward and what the magnitude of that reward is. In schizophrenia and psychosis, hedonics upon receipt of reward and initial responsiveness to reward are relatively intact, but responsivity to cues predictive of reward is reduced (Barch & Dowd, 2010; Zeng et al., 2022).

Recently, building on the idea that there may be difference in how individuals predict and respond to rewards, there has been an emphasis on understanding how individuals

with schizophrenia choose to allocate and expend effort based on individual differences in reward valuation. This process, called effort-cost computation, has been investigated using the Effort Expenditure for Reward Task (EEfRT) (Treadway, Buckholtz, Schwartzman, Lambert, & Zald, 2009). To perform optimally on this task, individuals must appropriately engage reward and cognitive control processes, and choose to expend effort that is relative to both the probability of actually obtaining reward and the magnitude of the reward itself. As a result of this structure, the EEfRT task is believed to better approximate real-world decision making, and in-the-moment evaluations of choices. Schizophrenia is associated with deficits in expending effort commensurate with probability of reward, and decreased willingness to expend effort scales with negative symptom severity (Barch, Treadway, & Schoen, 2014).

Risk-taking, which involves decision-making under conditions of uncertain probability of reward versus punishment/non-reward, necessitates initial reinforcement learning to guide choices and updating of action-outcome mappings in response to feedback. Data from risk-taking paradigms like the Iowa Gambling Task (Shurman, Horan, & Nuechterlein, 2005) and Balloon Analog Risk Task(Reddy et al., 2014) suggest that individuals with schizophrenia may exhibit difficulty making decisions under ambiguity (i.e., when probabilities of adverse outcomes are unknown and associations must be learned).

Reinforcement learning (the process by which actions are mapped to particular outcomes) has been conceptualized as a relatively “upstream” learning process relative to effort-cost computation (Barch & Dowd, 2010) that contributes to action selection and goal-directed behavior (Sutton & Barto, 2018). Reinforcement learning tasks like the Probabilistic Selection Task (PST) have been employed to dissociate “positive feedback sensitivity” versus “negative feedback sensitivity”, defined as preferential selection of highly rewarded stimuli versus preferential avoidance of stimuli with low probability of reward(Frank & Claus,

2006). Schizophrenia has been associated with deficits in learning of contingencies on the PST (Strauss et al., 2011; Waltz, Frank, Robinson, & Gold, 2007), which is comparatively more complex than other tasks where schizophrenia is associated with otherwise intact learning of contingencies. Evidence suggests an asymmetry in feedback sensitivity, with a preserved ability to learn from negative feedback but an impaired ability to learn from positive feedback (Waltz & Gold, 2007). However, it is worth noting that positive and negative feedback sensitivity as defined in non-computational studies of the PST do not allow for parsing of how individuals are using feedback to build representations of the task, how they update action value trial-to-trial based on feedback, or whether they weight feedback differently depending on if its positive or negative.

To investigate these constructs further, probabilistic reversal learning tasks (PRLT) have been used at length to probe reinforcement learning in schizophrenia to understand how individuals use new information or feedback to modify behavior appropriately. In PRLT, individuals must learn stimulus-reward pairs and then modify action selection in the context of switching reward contingencies. In many PRLT paradigms, the number of reversals is determined by the individual's ability to reach criterion (e.g., choosing the higher probability choice in 9 out of 10 trials), at which point the contingencies reverse. Number of reversals has been used to evaluate reinforcement learning deficits in schizophrenia. Previous studies have suggested deficits in schizophrenia during reversal phases of PRLT but possibly intact ability to initially discriminate stimuli; patients who are able to successfully learn contingencies initially take longer to reach criterion on reversal trials and fail to complete as many reversals (Culbreth, Westbrook, Xu, Barch, & Waltz, 2016; Pantelis et al., 1997; Reddy, Horan, & Green, 2016; Waltz & Gold, 2007).

It is evident that a variety of approaches, paradigms, and aspects of learning theory have been applied to gain mechanistic clarity related to negative symptoms. There is

evidence to suggest individuals with psychosis experience selective deficits in anticipating but not in experiencing reward, in using positive feedback to guide behavior but not in using negative feedback or punishment to guide behavior, in allocating appropriate cognitive resources and effort in service of reward, in making decisions under ambiguity, and in modifying behavior when contingencies shift. However, the central feature cutting across all these paradigms is a reliance on reinforcement learning for acquisition of initial associations and for learning new associations after receiving feedback.

Cognitive Correlates of Negative Symptoms

Negative symptoms have long been associated with cognitive deficits in schizophrenia(Bilder et al., 2000; Good et al., 2004; Heydebrand et al., 2004). Working memory in particular is a key cognitive process necessary for the representation and maintenance of action values used to guide future behavior, and deficits of working memory in schizophrenia have been consistently documented (Carter et al., 1996; Deserno, Sterzer, Wüstenberg, Heinz, & Schlagenhauf, 2012; J. Lee & Park, 2005; Park & Holzman, 1992). Working memory has also been implicated in other aspects of reward processing, including effort-cost computation. Poor working memory, over and above reward responsivity or ability to learn associations, predicts decreased willingness to expend effort but only in individuals with less severe negative symptoms (Whitton, Merchant, & Lewandowski, 2020). A number of studies have identified unique contributions of working memory to reinforcement learning deficits in schizophrenia (Collins, Brown, Gold, Waltz, & Frank, 2014; Collins & Frank, 2012; Hager et al., 2015), highlighting the importance of cognitive processes in decision-making processes.

Alterations in cortical regions like the prefrontal cortex (PFC) (Waltz & Gold, 2007) and cognitive control associated regions like anterior cingulate cortex (ACC) have been implicated in the reinforcement learning abnormalities observed in schizophrenia (Culbreth, Gold, Cools, & Barch, 2015). Functional magnetic resonance imaging (fMRI) studies have demonstrated prefrontal dysfunction in schizophrenia which is believed to contribute to behavioral deficits in reinforcement learning and symptomatology (Chung & Barch, 2016; Morris, Quail, Griffiths, Green, & Balleine, 2015).

However, findings about the link between corticostriatal abnormalities and cognitive processes are mixed. Some studies have found no differences in activation of striatal regions during decision-making tasks in schizophrenia (Culbreth, Westbrook, Xu, et al., 2016), but that schizophrenia is associated with dysconnectivity and abnormal functional connectivity patterns between prefrontal and striatal regions. A recent meta-analysis investigated reward anticipation studies in those who are at clinical high risk for psychosis (CHR). This work indicated that there may actually may be increased activity in brain regions associated with cognitive control, such as the medial PFC and ACC, as well as hypoactivity in dorsal striatum during reward anticipation (Zeng et al., 2023), contrary to findings in fully psychotic populations. This suggests that a shift in this pattern to decreased prefrontal activity, and, as an extension of that, altered cognitive contributions to reinforcement learning could be a key part of the transition from risk state to full psychosis.

Negative symptom etiology in psychosis is complex, and it is not yet fully understood. To gain a better understanding of the basis of these important symptoms, there is a need to systematically evaluate and integrate across findings related to reinforcement learning deficits, including contributions of cognitive processes such as working memory, and neural dysfunction in reward and cognitive neural circuits.

Reinforcement Learning Accounts of Psychosis

Having discussed the broader literature on reward processing accounts of psychosis, I provide a deeper review of reinforcement learning specifically in this section. Decades of research have implicated altered reinforcement learning in the pathophysiology of psychosis (Schultz, 1998), with a particular emphasis on the causal role of abnormalities in dopaminergic circuits comprising striatal and frontal brain regions such as the orbitofrontal cortex(Groman et al., 2019; Wallis, 2007). For example, there is an extensive body of literature focused on altered phasic dopaminergic error signaling in the ventral tegmental area(Deserno, Schlagenhauf, & Heinz, 2016; Gold, Waltz, Prentice, Morris, & Heerey, 2008; Juckel et al., 2006; Radua et al., 2015),with research historically focusing on the role of prediction error signaling in the learning of stimulus-reward pairs specifically. Traditional reinforcement learning accounts of negative symptoms posit that there is a critical failure to use reward-related feedback to effectively direct and modify behavior as a result of reduced dopamine neurotransmission in frontal brain regions (Dowd, Frank, Collins, Gold, & Barch, 2016). Inability to adequately predict reward would, in theory, diminish ability to initiate and maintain goal-directed behaviors and decrease engagement in social and role (e.g., school, work) settings. However, schizophrenia and other psychotic disorders are also associated with over-learning of associations between neutral or irrelevant events, which is believed to contribute to positive symptomatology like delusions and hallucinations (Schmack, Rothkirch, Priller, & Sterzer, 2017). Over-learning of irrelevant stimuli has been attributed to dopamine hyperactivity in the striatum (Fusar-Poli & Meyer-Lindenberg, 2013; Sarpal et al., 2015).

Taken together, reinforcement learning disruption in psychosis is complex. Alterations in different dopaminergic circuits may lead to both poor integration of reward-related feedback as well as over-association of otherwise irrelevant information, thus

contributing to both negative and positive symptomatology(Millard, Bearden, Karlsgodt, & Sharpe, 2022). Here I leverage the reinforcement learning framework and focus specifically on how deficits may contribute to negative symptomatology due to their persistence throughout the course of psychotic illness, lack of targeted treatments for negative symptoms, and their impact on functioning.

Despite substantial scientific interest, the development of novel paradigms, and the identification of discrete reward processes, the precise relationship between reward processing, reinforcement learning, and psychosis symptomatology requires further study to move towards a unified etiological framework.

Computational Psychiatry Approaches

Computational psychiatry exists at the nexus of mathematical modeling, neuroscience and cognitive science, and has emerged as a powerful framework for investigating reinforcement learning. Computational models provide estimates of precise decision-making functions that can be integrated across different levels of analyses (e.g., genetic factors, brain circuits, symptoms, behavior) to improve our understanding of cognitive and neural processes(Friston, 2023; Maia & Frank, 2011) that are affected in specific disorders.

Computational models can be conceptualized as mathematical expressions of theories, with parameters believed to correspond to specific psychological, cognitive, and neural mechanisms. Computational modeling can be applied to a variety of tasks, including tasks described earlier in this Chapter. In the reinforcement learning literature, they are many models that have been applied to better understand learning behavior. Broadly, they all posit that individuals (or “agents” in traditional mathematical literature) are navigating

states with a given environment. Within a state, there are a number of actions an individual can take that result in specific outcomes, with outcomes being used in some form to update the value of an action. Individuals must balance exploration, or gathering information, and exploitation, or using known information, to maximize certain outcomes. Existing computational studies of reinforcement learning in schizophrenia have been fruitful, namely in their ability to quantify specific deficits in reward prediction errors and how those deficits correspond to neural signals and cognitive measures but studies vary in the algorithm used which limits which subprocesses can be evaluated.

Computational theoretical models of reinforcement learning posit two decision-making systems: model-free learning (MF-RL), in which value of a behavior is learned from reward history, and model-based learning (MB-RL), in which value of a behavior takes into account environmental structure, predicted changes in environment, and outcomes (Daw, Niv, & Dayan, 2005; Gläscher, Daw, Dayan, & O'Doherty, 2010). To probe differences in MB-RL and MF-RL, previous studies have utilized a 2-stage reinforcement learning task that allow parsing of decision-making based on previously experienced outcomes versus decision-making based on a representation of task structure to maximize future rewards. MF-RL is believed to subserve habitual behavior; MB-RL is believed to subserve goal-directed behavior. Few studies have used 2-stage tasks in clinical populations, but existing evidence from computational studies evaluating MF-RL vs MB-RL suggests specific MB-RL deficits in psychosis, with preserved habitual or MF-RL learning (Culbreth, Westbrook, Daw, Botvinick, & Barch, 2016). Per this framework, MB-RL deficits would by definition result in deficits in goal-directed behavior and negative symptomatology.

Model-free algorithms are further classified as value-based or policy-based (Zhang & Yu, 2020). Policy-based models involve building a decision-making rule and maintaining it in memory; value-based models do not involve storage of explicit policy but rather an implicit

policy of choosing actions with the best value. Commonly used value-based models in the existing computational psychiatry literature in schizophrenia include State-Action-Reward-State-Action (SARSA) and Q-learning models. These models can fit data from tasks like probabilistic reversal learning tasks and probabilistic selection tasks.

SARSA is a type of temporal difference model that accounts for delayed reward with the assumption that an agent seeks to maximize cumulative reward. SARSA models lead to action selection based on a current decision-making rule, and then uses outcomes from the executed action to update the same decision-making rule. SARSA models can also be modified to include an eligibility trace that speeds up learning by updating values for previously visited state-action pairs. Q-learning models are similar to SARSA in that they are temporal difference models; however, Q-learning relies on a decision-making rule for the environment and separate decision-making rule that is updated after executed actions. A commonly employed combined policy and value-based model is Actor-Critic. The “Critic” evaluates the reward value of a state, and the “Actor” chooses an action based on learned stimulus-response pairs. Prediction errors are used to adjust the critic.

Existing literature employing model-free algorithms like Q-learning and Actor-Critic have mixed findings. Some evidence suggests that reinforcement learning in schizophrenia is best captured by an Actor-Critic, where learning is driven solely by prediction errors, rather than Q-learning (Gold et al., 2012; Strauss et al., 2015). Other studies have shown that schizophrenia is characterized by simpler associative learning captured by the Rescorla-Wagner (RW) rule (Schlagenhauf et al., 2014). Across existing computational studies, that differ in the task modeled and the algorithm applied, evidence suggests schizophrenia may be associated with selective deficits in certain forms of reinforcement learning that depend on task demands such as complexity and type of feedback received.

Furthermore, data suggest that increasingly complex models are unlikely to capture patient choice behavior.

Existing studies have largely focused on Q-learning models with one updating parameter, regardless of whether an individual experiences a gain or loss, and regardless of whether one receives positive or negative feedback. In Study 2 and Study 3, we employ a modified Q-learning model that accounts for asymmetries in integrating positive and negative feedback, as well as retention of action value over time. This will allow us to better identify asymmetry in prediction error signaling and have quantified selective deficits in learning from reward but not punishment.

Overview of Current Project

The goal of this dissertation is to investigate negative symptoms, reward-guided decision-making and reinforcement learning deficits in psychotic illness. In this work I focus on two paradigms that require reinforcement learning, and thus initial learning of outcomes of specific choices and integration of feedback to guide future decisions: 1) a risk-taking paradigm in which individuals must integrate feedback and modify behavior appropriately to optimize rewards (Study 1) and 2) a reversal learning paradigm that allows us to model how individuals integrate different kinds of feedback and retain information about previous choices when contingencies shift (Studies 2 & 3).

Through use of these paradigms, we stand to gain a deeper understanding of how reinforcement learning deficits relate to decision-making under uncertainty. We also can learn what features of reinforcement learning may differentiate psychosis from other psychopathology and capture possible asymmetries in prediction error signaling and integration of different kinds of feedback. Underlying all three studies is an emphasis on a

broadly defined spectrum of psychotic symptoms, including how deficits manifest across stage of illness (e.g., early course of illness versus chronic), across and between diagnostic boundaries (e.g., schizophrenia versus bipolar disorder), and across the spectrum of severity (e.g., subclinical psychotic features in healthy controls). In these studies I endeavored to characterize moderators of deficit severity and symptom severity across the full spectrum of psychotic presentations.

Study 1: Individual differences in optimal risk-taking in early psychosis

I first focused on risky decision-making, which in part relies on reinforcement learning to learn associations between specific actions and outcomes. I investigated how risk-taking in optimal contexts (i.e., low probability of loss or punishment, high probability of reward) may be disrupted in the early course of psychotic illness and the extent to which optimal risk-taking deficits scale with symptomatology. The sample included individuals with a variety of psychotic disorder diagnoses, including schizophrenia and bipolar disorder with psychotic features, that showed variability in illness presentation and symptom severity to allow us to probe if differences in optimal risk-taking scale with negative symptomatology.

Study 2: Computational modeling of shared and unique reinforcement learning deficits in schizophrenia and bipolar disorder.

I employed a computational approach to quantify specific reinforcement learning processes that may be disrupted in psychotic illness. I looked across and between schizophrenia and bipolar disorder to identify common and unique deficits in both classes of disorders. I also investigated alterations in reinforcement learning in healthy controls to identify whether there is continuity in reinforcement learning alterations in the subclinical range of the psychosis spectrum.

Study 3: White matter alterations and reinforcement learning deficits

Using the metrics derived in Study 2, I sought to identify the relationship between computational parameters and structural neural metrics of white matter connectivity to integrate across levels of analysis (brain, behavior, symptoms).

CHAPTER 2

STUDY 1: Individual Differences in Optimal Risk-Taking in Early Psychosis

INTRODUCTION

Psychosis typically emerges in adolescence and early adulthood (F. S. Lee et al., 2014), a developmental period associated with substantial neural (Galvan et al., 2006), psychological, and social change (Spear, 2013). As a part of this developmental stage, sampling a diversity of experiences and exploring one's environment is normative and may be critical to social learning and emotional maturation. Such experiences include an inherent amount of risk-taking that can be considered to be normal and even adaptive (Crone & Dahl, 2012; Hauser, Iannaccone, Walitza, Brandeis, & Brem, 2015). Alterations in reward processing and adaptive risk-taking in adolescence and early adulthood may then be especially deleterious, leading to failure to meet developmental milestones foundational to functioning in adulthood, including living independently and seeking and maintaining gainful employment.

Alterations in risky decision-making have been evaluated at length in chronic schizophrenia samples, using a variety of paradigms such as the Iowa Gambling Task (IGT) (Bechara, Damasio, Tranel, & Damasio, 1997) or the Balloon Analogue Risk Task (BART) (Lejuez et al., 2002). The IGT requires individuals to choose gambles from four decks of cards, with two decks offering larger monetary rewards and two decks offering smaller monetary rewards. Some cards carry risk of a monetary penalty, and participants must identify an optimal strategy to minimize penalty; importantly, every card carries immediate monetary reward, but every card does not carry risk of penalty. Alternatively, the BART involves inflating a virtual balloon to gain reward, with reward increasing with each

successive inflation. However, the balloon will explode if inflated too many times; the number of inflations before explosion varies probabilistically. Participants can “cash out” at any time on a given trial. Unlike the IGT, loss is *certain* on each and every trial if an individual inflates the balloon too many times. Because of this key difference, the BART allows for exploration of how much risk of eventual loss an individual is willing to tolerate to optimize reward. Furthermore, by using a very simple and relatable real life negative outcome (the balloon popping), the BART was developed to provide a more realistic index of risk taking, relative to the more abstract risks of the IGT.

Previous studies using the BART in schizophrenia samples show that, on average, schizophrenia is associated with less risk-taking (i.e., fewer inflations) compared to healthy controls and other psychiatric groups, resulting in fewer earned points. However, the ability to make “safe” choices (i.e., “cash out” before explosion) and to avoid punishment appear to be intact (Boka et al., 2020; Cheng, Tang, Li, Lau, & Lee, 2012; Reddy et al., 2014). While individuals with schizophrenia learn initial contingencies and are able to modify behavior to some extent, risk propensity is blunted overall compared to controls (Brown et al., 2015).

Performance on the risk-taking tasks including the BART is reliant on reinforcement learning, as individuals must learn to associate choices with specific outcomes, and then use feedback (e.g., gaining points and/or monetary reward, explosions, and loss of points and/or monetary reward) to modify choice behavior on successive trials. Failure to appropriately learn associations and use feedback to optimize reward has been implicated in negative symptomatology in schizophrenia and psychosis. The negative symptom of avolition, defined as diminished goal-directed behavior and reduced environmental engagement, may contribute to variability in optimal risk-taking specifically. Individuals with pronounced avolition may be less able to use feedback to build representations of different

risk contexts to effectively modify their behavior. Here I have focused on avolition, with the goal of gaining a greater understanding of this relationship.

Evidence for the relationship between altered risk-taking and negative symptomatology overall is mixed, with several studies reporting no significant relationship between symptoms and task metrics (Brown et al., 2015; Reddy et al., 2014). However, few studies have examined alterations in risk-taking specifically in the early course of psychotic illness or in spectrum-based samples that ostensibly have greater clinical variability than schizophrenia-only or chronic samples. Additionally, existing studies focusing on evaluating differences in risk-taking between patients with schizophrenia and healthy controls have not explicitly considered the role of context. Most studies use variant of the BART with one type of balloon with a single risk level (probability of explosion on each inflation), as opposed to a version that includes different balloons with differing probabilities of explosion that require participants to discriminate between stimuli (i.e., different colored balloons) and modify behavior depending on the degree of risk carried by a certain balloon.

It is important to understand how people behave in circumstances when greater risk-taking might be optimal, or how decision-making behavior can vary depending on the degree of risk (i.e., in higher or lower risk scenarios). If an individual is unable to modulate their level of risk-taking based on risk context, it may lead to engaging in maladaptive risk-taking (e.g., substance use or unsafe sex practices) where there are serious potential negative outcomes, over more adaptive forms of risk-taking (e.g., prosocial risk-taking, environmental exploration) which might take place in a context of more potential long term positive outcomes and less severe negative outcomes. The extent to which the ability to appropriately modulate behavior based on degree of risk is impaired in psychosis remains underexplored, along with whether the relationship between negative symptoms and risk-taking behavior varies based on risk context.

While understanding the link between symptoms and behavior is of interest, the existing data on the relationship between altered risk-taking and broadly construed negative symptomatology is mixed, with several studies reporting no significant relationship between overall negative symptom scores and task metrics from the BART (Brown et al., 2015; Cheng et al., 2012; Luk et al., 2019). This may be because different aspects of negative symptoms have different relationships with task performance. Here, I use a version of the BART that contains both high and low risk conditions, to probe the relationship between symptoms, behavior, and context. I proposed that the negative symptom of avolition, defined as diminished goal-directed behavior and reduced environmental engagement, may particularly contribute to variability in optimal risk-taking. For example, I expected that individuals with pronounced avolition will be less able to use information about the degree of risk to effectively modify their behavior.

I aimed to characterize differences in risk-taking behavior between early psychosis (EP) participants and healthy controls (HC). Optimal risk-taking was evaluated using two BART metrics: (1) inflations on low-risk trials with high probability of reward and low probability of punishment with more successful inflations indicating good performance and (2) number of explosions (i.e., adverse outcomes) across all trials, with more explosions indicating poor performance. In addition to evaluating overall risk-propensity and number of adverse outcomes, I also evaluated differences in behavior following punishment. I hypothesized that, consistent with prior studies suggesting overall blunted risk propensity in schizophrenia, EP will show less risk-propensity across both high- and low-risk trials. Additionally, I hypothesized that EP would show more explosions across all trials and no differences in response to punishment compared to HC. Taken together, reduced risk-propensity regardless of the degree or risk and more explosions across the task would reflect suboptimal risk-taking behavior in EP, with response to punishment conserved.

Within the patient group, I hypothesized risk-taking deficits in EP would scale with symptomatology, such that avolition would be associated with reduced inflations across both high- and low-risk contexts, more explosions, and fewer inflations following punishment.

METHODS

Participants

The sample included 56 early psychosis participants (EP) and 52 age-matched healthy controls (HC) between the ages of 16 and 25 who participated in a longitudinal study at Zucker Hillside Hospital and the Feinstein Institute for Medical Research in New York (Table 1). The study protocol was approved by the Institutional Review Board of Northwell Health. Participants ages 18 and older were provided with written informed consent. Minor participants were provided with written assent along with parental written informed consent. Inclusion criteria for the EP group included duration of illness less than two years, current treatment with atypical antipsychotic medication, and a primary psychotic illness diagnosis including Schizophrenia, Schizoaffective, Bipolar Disorder with Psychotic Features, Major Depressive Disorder with Psychotic Features, and Unspecified Psychotic Disorder. Inclusion criteria for the HC group included no primary psychotic illness or current mood disorder diagnosis in the participating individual or their first-degree relatives. Exclusion criteria for both EP and HC groups included IQ less than 70, insufficient fluency in the English language, and history of neurological disorders or significant head trauma.

Diagnostic Interview

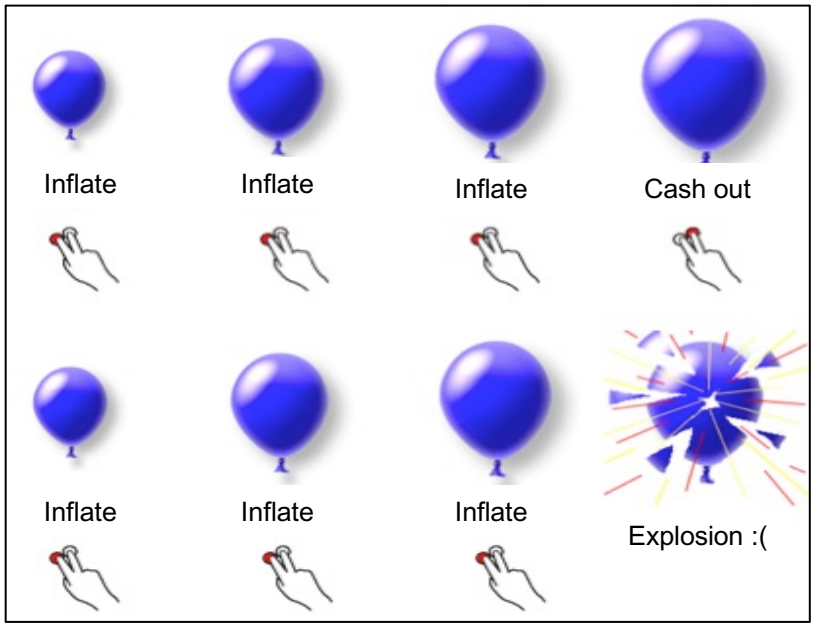
All EP and HC participants were evaluated for current and lifetime Axis-I disorders using the Structured Clinical Interview for the DSM-IV (First & Gibbon, 2004). All diagnostic

interviews were conducted by trained graduate-level assessors. All information collected during the SCID was compiled into a summary for each participant, which was then reviewed by at least two faculty psychologists to reach consensus on diagnoses.

Balloon Analogue Risk Task (BART)

Participants completed a version of the BART, a risk-propensity task, modified for use in the Consortium for Neuropsychiatric Phenomics (CNP) (Poldrack et al., 2016). On each trial, participants inflated a balloon via button press to gain points. The balloon on each trial was initially 152 x 152 pixels and increased by 2 x 2 pixels for each successive inflation. The likelihood of the balloon exploding on a given inflation was determined probabilistically; if the balloon popped, no points were gained (*Figure 1*). Trials were either high-risk, where fewer inflations were required before the balloon popped (probability ranging from 1 to 32 inflations), or low-risk, where more inflations were required for the balloon to pop (probability ranging from 1 to 128 inflations). Degree of risk was indicated by balloon color, the mapping of which was unknown to the participant at the start of the task. Participants were instructed to earn as many points as possible, and to decide during each trial whether to “cash out”; if the balloon exploded before cashing out, no points would be gained. All participants completed 40 trials (20 high-risk and 20 low-risk balloons, randomized across the task). Points gained were displayed after each trial, with a cumulative total presented at the end of the task.

Figure 1. BART Schematic. Participants inflate a virtual balloon via button press. In the trial depicted at the top of the figure, the participant inflated the balloon and chose to cash out, leading to retention of points. In the trial at the bottom of the figure, the participant continued inflating until the balloon exploded, leading to no points gained on that trial.



To perform optimally and maximize points, individuals must learn the association between balloon color and risk, and subsequently alter the number of inflations and "cash out" selectively according to trial type to minimize punishment (i.e., balloon explosions). Therefore, an optimal performer would inflate more on low-risk trials, where there is greater likelihood of reward and lower likelihood of punishment.

The first metric derived from the BART is mean adjusted inflations (MAI), which averages the number of button presses (i.e., inflations) across trials where the balloon did not pop, with higher values indicating greater risk-propensity. On low-risk trials, greater risk-propensity is adaptive and optimal to take advantage of the opportunity to gain points. On high-risk trials, reduced risk-propensity is optimal to avoid punishment and retain as many points as possible. The second metric is total explosions (TE) which is the number of trials with an adverse outcome (i.e., explosions) across all trials in the task, serving as an

indicator of suboptimal performance, The third metric is mean adjusted inflations after an explosion (IAE), calculated as the number of inflations on a trial immediately following an explosion, serving as an indicator of behavioral change following a punishment.

Clinical and cognitive measures

EP participants completed an additional clinical interview including measures of symptomatology. Negative symptom ratings were obtained via the Scale for Assessment of Negative Symptoms (SANS)(Andreasen, 1989). The Global Rating of Avolition score from the SANS was the primary clinical variable of interest. The SANS also includes an Avolition-Apathy subscale which includes items assessing for grooming, role functioning, and physical anergia. While the Avolition-Apathy subscale simply averages its items, the Global Rating of Avolition captures the full severity of an individual's amotivated state. Thus, I used the Global Rating of Avolition score in lieu of the Avolition-Apathy subscale as an overall indicator of severity, across domains of amotivated behavior. Scores on the Global Rating of Avolition rating range from 0 (No Avolition) to 5 (Severe).

IQ was estimated using combined scaled scores on the Wechsler Adult Intelligence Scale (WASI)- Third Edition. The WASI-III was administered as part of a larger neuropsychological battery, including subtests of the MATRICS Cognitive Consensus Battery (MCCB) (Nuechterlein et al., 2008).

Statistical analysis

All statistical analyses were conducted in Stata v. 16.1. Age was centered at the mean across both EP and HC groups. Sex was dummy coded with 0 indicating male biological sex. BART metrics were evaluated for non-normality using the Shapiro-Wilk test.

Non-normal variables were square root transformed to correct for skew, kurtosis and non-normality, and re-evaluated to confirm normality.

ANCOVAs were conducted to identify differences in risk-taking behavior, as indexed by BART, between the HC and EP groups. Age and sex were evaluated for inclusion as covariates given documented associations with risk-taking behavior (Braams et al., 2015). Additional post-hoc t-tests were conducted to determine if there were differences in behavior based on risk condition. Within the EP group, multiple regression analyses were conducted to determine the relationship between avolition, as indexed by the SANS, and risk-taking behavior, as indexed by the BART, to evaluate the relationship of symptomatology and risk behaviors.

RESULTS

Demographic and Clinical Characteristics

Demographic and clinical characteristics are shown in Table 1. The EP group included individuals with a range of diagnoses and variability in clinical presentation. As in many previous studies, the EP sample was predominantly male.

Comparison of Early Psychosis and Healthy Control: MAI, TE, IAE

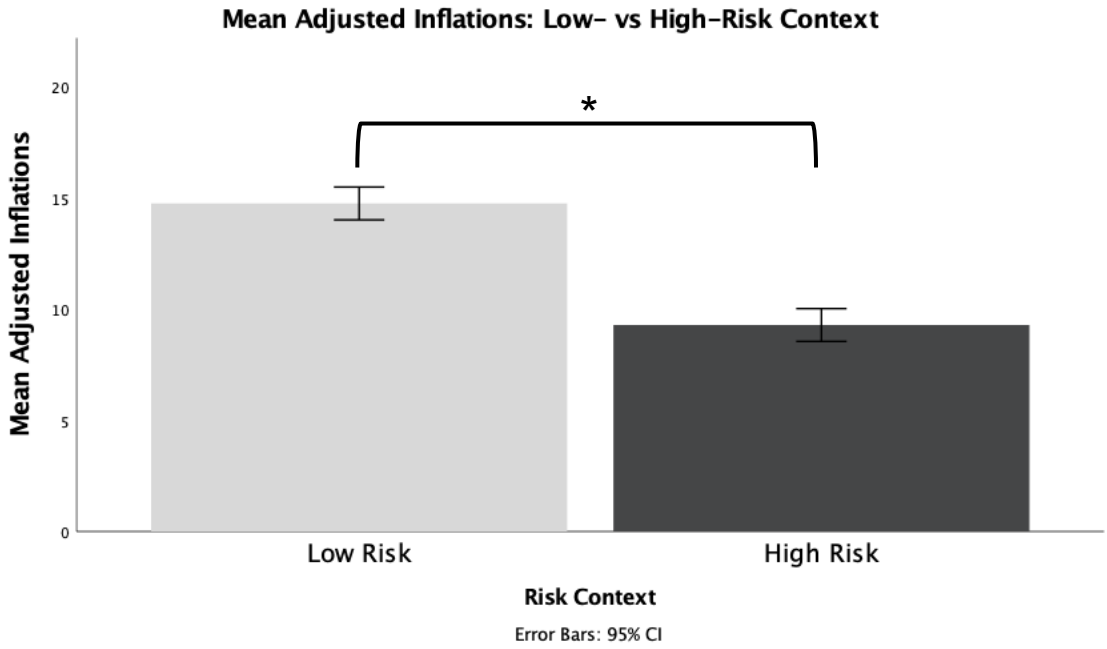
First, I conducted analyses collapsed across EP and HC. In the full sample, collapsed across group, participants inflated more on low-risk trials relative to high-risk trials ($t(108)=7.14$, $p < 0.001$) (*Figure 2*). This indicates that the differences in risk probabilities were meaningful and detectable by participants and suggests that both EP and HC participants on average appropriately modified their behavior as a function of trial type and risk context. In addition, across both risk conditions and participant groups, female

participants inflated less than male participants (t(107)=-2.39, p=0.009), which is consistent with prior literature indicating greater risk-propensity in males. Therefore, sex was included as covariate for all analyses.

Table 1. Study 1 Demographics

	Group	
	HC (N=52)	**EP (N=56)**
Sex*		
% Male	40.4	71.9
% Female	59.6	28.1
IQ*		
	106.8 (1.4)	99.5 (1.9)
Race		
% Asian or Pacific Islander	25	10.5
% Black or African American	15.4	40.4
% Native American	0	1.8
% White	53.8	33.3
% Mixed Race or Other	5.8	14
Age		
Minimum	16.2	16.2
Maximum	26.9	26.9
Average (SD)	20.7 (2.0)	21.7 (2.5)
Diagnosis		
% Schizophrenia Spectrum	--	67.8
% BP with psychotic features	--	28.6
% MDD with psychotic features	--	2.4
% Other Specified Psychotic Disorder	--	1.2

Figure 2. Inflations in High versus Low-Risk Contexts. * p < 0.05



Notably, age was not correlated with any task metrics and thus was not included as a covariate in any analyses. Previous studies have suggested a quadratic relationship between age and risk behavior in youth. However, risk-taking peaks in mid-adolescence, both behaviorally, in terms of increased risk-taking, and neurally, as indexed by nucleus accumbens activity in response to reward (Braams et al., 2015). The study sample consists of individuals in later adolescence and early adulthood, and no individuals under 16, which may have contributed to the lack of age effects.

While the overall ANCOVA model evaluating differences in MAI across trial types between EP and HC was significant ($F(2,106=3.14$, $p=0.047$), group was not a significant predictor of differences over and above sex ($p=0.43$). Similarly, an ANCOVA evaluating differences in MAI specifically on low-risk trials was significant ($F(2,106)=4.21$, $p=0.019$), but group was not a significant predictor of differences over and above sex ($p=0.42$). The

ANCOVA evaluating differences in MAI on high-risk trials was non-significant (F(2,106)=0.21, p=0.52).

An ANCOVA evaluating differences in TE between EP and HC was trending (F(2, 106)=2.69, p=0.072), but group was a nonsignificant predictor (p=0.47). An ANCOVA evaluating differences in IAE on low-risk trials was significant, but group was a nonsignificant predictor (p=0.57). An ANCOVA evaluating differences in IAE on high-risk trials was nonsignificant (F(2,106)=0.08, p=0.92).

Early Psychosis Group Analysis

To determine whether individual differences in negative symptoms might explain some of the variability in risk propensity within the patient group, multiple regression analyses were conducted to evaluate the relationship between avolition (via the Global Rating of Avolition of the SANS) and MAI, IAE, and TE in the BART (*Table 2*). Global Rating of Avolition and sex significantly predicted mean adjusted inflations across both risk-contexts of the BART (F(2,52)=3.69, p=0.032), where avolition significantly predicted inflations over and above sex (b=-0.198, p=0.019).

Next, I sought to clarify whether the effect of avolition on inflations varied by risk context, to understand whether avolition impacted EP participants' ability to modify their behavior according to risk level. Two post-hoc regressions were conducted to evaluate the impact of avolition in high-risk versus low-risk trials. For high-risk trials, a multiple regression with global ratings of avolition and sex did not significantly predict MAI (F(2,52)=0.92, p=0.41). For low-risk trials, avolition and sex did significantly predict MAI (F(2,52)=3.68, p=0.032), with avolition significantly predicting MAI over and above sex (b=-0.25, p=0.015)

such that an increase in ratings of avolition was associated with fewer inflations on average on low-risk trials.

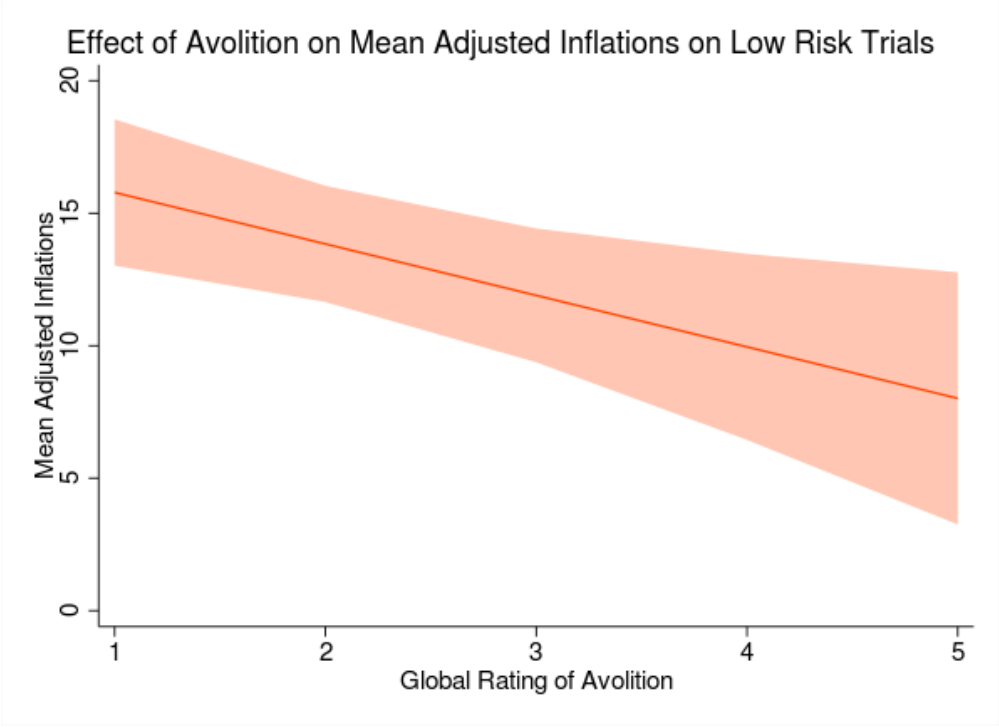
Table 2. Avolition and BART Performance in EP

	MAI: Across All Trials					
Variable	B	SE B	F	df	p	Adj R^2
Global Rating of Avolition	-0.198*	0.0815	3.69	2,52	0.032	0.0701
Sex	.370*	0.247				
Constant	3.461*	0.241				

	MAI: Low-Risk Trials					
Variable	B	SE B	F	df	p	Adj R^2
Global Rating of Avolition	-0.245*	0.0971	3.68	2,52	0.032	0.0730
Sex	0.466	0.310				
Constant	3.722*	0.289				

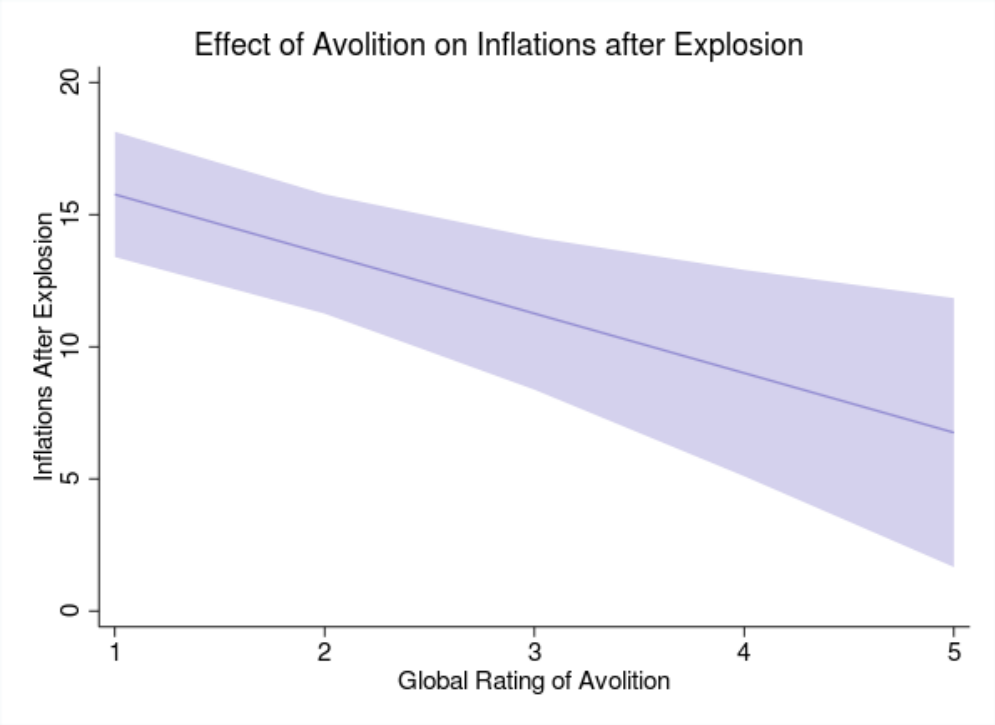
	IAE: Low-Risk Trials					
Variable	B	SE B	F	df	p	Adj R^2
Global Rating of Avolition	-0.314*	.0952		2,51	0.0002	0.186
Sex	0.937*	.0952				
Constant	3.454*	.0300				*** p < .05**

Figure 3. Avolition and Mean Adjusted Inflations on Low-Risk Trials on the BART. Depicted with 95% CI.



I conducted a multiple regression to determine if avolition was associated with the number of trials with a punishment outcome. Avolition and sex did not significantly predict TE ($F(2,52)=2.06$, $p=0.14$). I then conducted multiple regressions to determine if avolition was associated with reduced inflations following punishment in both low-risk and high-risk trials. Avolition and sex did not significantly predict IAE on high-risk trials ($F(2,51)=0.91$, $p=0.41$). However, avolition and sex did significantly predict IAE on low-risk trials ($F(2,52)=9.96$, $p=0.0002$), with avolition predicting IAE on low-risk trials over and above sex ($b=-0.31$, $p=0.002$), such that an increase in ratings of avolition was associated with few IAE on low-risk trials (Figure 4).

Figure 4. Avolition and BART Mean Adjusted Inflations After Explosion on Low-Risk Trials.
Depicted with 95% CI.



Exploratory Analyses

IQ and Task Performance

EP and HC significantly differed in IQ, consistent with previous studies in psychosis populations (t(107)= 3.05, p=0.002)(Aylward et al., 1984). IQ was not considered as a covariate initially due to concerns that it would minimize meaningful variability between groups related to my questions of interest. In supplementary analyses including IQ as a covariate in ANCOVAs evaluating EP vs. HC differences in MAI, TE, and IAE, all models remained non-significant. In the EP group analyses, with the inclusion of IQ as a covariate, the effect of avolition on MAI and IAE after explosions on low-risk trials remained significant.

Exploratory Factor Analysis

Because BART metrics were correlated with one another, an exploratory factor analysis with no limitations on retained factors was run in Stata to investigate if MAE, TE, and IAE index different constructs (*Table 3*). All three BART metrics showed high communality (communality= 1 - uniqueness), and TE showed the most uniqueness of all three metrics.

Table 3. Study 1 Exploratory Factor Analysis

Variable	**Factor 1**	**Factor 2**	**Uniqueness**
Total Explosions	0.7874	0.2216	0.0735
Mean Adjusted Inflations	0.9625	0.0025	0.3309
Inflations After Explosion- Low Risk	0.8711	-0.2031	0.1999

CONCLUSION

Group Differences

Contrary to my hypotheses, results suggested that at the group level, EP and HC do not significantly differ in BART task engagement as indexed by MAI, TE, or IAE on either high- or low-risk trials. Instead, findings indicate that alterations in risky decision-making may not be unilateral in psychosis, with the ability to modulate behavior based on degree of risk and punishment potentially conserved in individuals with psychosis. Regardless of group, participants on average inflated more in the low-risk trials suggesting that both EP and HC individuals understood the task and successfully learned to modulate behavior based on balloon and trial type. These findings are partially consistent with existing literature suggesting that participants with psychosis can learn contingencies and modify behavior (Culbreth et al., 2015; Strauss et al., 2011) but are inconsistent with literature suggesting overall blunted risk-propensity relative to healthy controls (Boka et al., 2020; Cheng et al., 2012; Luk et al., 2019; Tikàsz et al., 2019). This may be in part due to greater clinical variability within the spectrum-based sample, with some EP with less severe symptomatology performing at a level similar to HC.

Individual Differences in EP

When exploring variability within the EP group, clinical ratings of avolition accounted for a significant amount of variability in risk propensity over and above sex. The effect of avolition on task engagement was significant on low-risk but not on high-risk trials. This may suggest that when the likelihood of an adverse outcome (i.e., balloon exploding and loss of points) is high, individuals with elevated avolition may be appropriately risk-averse, similar to individuals without elevated avolition and healthy controls. In conditions of low-risk with

maximum likelihood of reward, there is an opportunity to adaptively exploit the task context to gain more points. In this context, individuals with elevated avolition ratings do not optimize performance, which is reflected in fewer inflations despite the decreased chance of negative outcome and increased opportunity for higher reward. This is consistent with the notion that individuals with more avolition may be less likely, in daily life, to take advantage of opportunities and engage with the environment to obtain rewarding outcomes. However, avolition was not associated with TE, the second metric of optimal risk-taking. Based on EFA results, TE may index a different risk-related construct than MAI and had the greatest uniqueness amongst BART metrics. Further rigorous investigation of symptom moderators of risk behavior must be conducted before concluding that negative symptoms moderate risk behavior in low-, but not high-risk situations.

Avolition was associated with IAE, the measure of behavior in response to punishment. Specifically, avolition was associated with reduced IAE on low-risk rather than high-risk trials. By nature of having a lower probability of explosion, low-risk trials have a low incidence of adverse outcomes or punishment. In theory, an optimal performer could be less sensitive to punishment in low-risk trials (which could be interpreted as coincidental or spurious) and more sensitive to punishment on high-risk trials in order to maximize points on future trials. My findings could indicate that EP with elevated avolition may weight punishment heavily when making decisions on subsequent trials even when the likelihood of actual punishment is low. These findings are consistent with literature suggesting greater punishment sensitivity in schizophrenia (Brown et al., 2015; Luk et al., 2019; Waltz et al., 2007).

Taken together, findings suggest that suboptimal risk-taking and response to punishment in EP may be influenced by negative symptom severity, specifically avolition. However, the relationship between avolition and risk-taking varies depending on the degree

of risk for a given trial. Risk-taking deficits may not be unilateral in psychotic illness, with substantial within-group variability affecting risk-taking in a context-dependent manner.

Limitations and Future Directions

I sought to use a broad-spectrum approach, examining the relationship between negative symptoms and risk-taking across different psychotic disorder diagnoses. Negative symptomatology can cut across psychotic disorder diagnoses and may be a clinical correlate that meaningfully predicts risk-taking behavior over and above diagnosis. I was not sufficiently powered to compare psychotic mood disorders and schizophrenia spectrum disorders. Secondly, the study did not include comprehensive measure of real-world adaptive or optimal risk-taking. Existing measures frequently assess risk across domains, include life threatening and/or sexual risk-taking, rather than specifically prosocial risk-taking or exploratory behavior that is adaptive in the adolescent and early adulthood period. In addition, a wider age range would have allowed us to observe EP and HC differences in the broader developmental trajectory of adaptive risk-taking.

Future studies may seek to identify the mechanistic links between reduced real-world adaptive risk-taking, task-based measures of optimal risk-taking, and psychosocial functioning in psychotic illness. Additionally, longitudinal studies may allow the precise mapping of developmental trajectories of risk-taking behavior in early psychosis compared to typical development.

CHAPTER 3

STUDY 2: Computational modeling of shared and unique reinforcement learning deficits in schizophrenia and bipolar disorder.

INTRODUCTION

In Chapter 2, I detailed an investigation of how individuals early in the course of psychotic illness tolerate risk and make decisions under different degrees of uncertainty. I demonstrated that psychosis participants were able to distinguish high- and low-risk contexts, and on average modified their behavior appropriately to optimize reward. Within the psychosis group, there was considerable heterogeneity, and negative symptomatology was associated with decreased risk-taking in the low-risk context and decreased risk-taking after punishment on low-risk trials. However, I was not able to assess *how* individuals with psychosis may be updating value representations trial-to-trial based on positive vs. negative feedback, and whether reinforcement learning subprocesses scaled with symptomatology.

Computational psychiatry exists at the nexus of mathematical modeling, neuroscience, and cognitive science, and has emerged as a powerful framework for investigating reinforcement learning. Computational models provide estimates of precise decision-making functions that can be integrated across different levels of analyses (e.g., brain, behavior) to improve our understanding of cognitive and neural processes (Maia & Frank, 2011) that are affected in specific disorders such as schizophrenia and bipolar disorders. In providing discrete parameters, computational reinforcement learning models allow for dissociation of specific deficits in reinforcement learning. This may reveal disorder specific, or symptom specific, deficits in the underlying cognitive architecture in different psychiatric disorders.

Here, I choose to focus on schizophrenia spectrum and bipolar disorders, which

share a number of overlapping symptoms, etiological factors and treatments (Badner & Gershon, 2002; Ellison-Wright & Bullmore, 2010). For example, schizoaffective disorder and bipolar disorder with psychotic features are both characterized by prominent, impairing, and distressing mood (i.e., depressive symptoms and/or manic symptoms sufficient for a major depressive episode or manic episode classification per the DSM-V) and psychotic symptoms. The key distinction between these diagnoses is whether mood episodes ever occur in the absence of psychotic symptoms (bipolar disorder with psychotic features) or if psychotic symptoms ever occur in the absence of mood symptoms (schizoaffective disorder).

There is limited evidence to suggest that schizoaffective disorder and bipolar disorder with psychotic features have fully distinct etiology, or distinct clinical presentation particularly when individuals of either diagnosis experience mood and psychosis concurrently. Even when considering just non-affective psychosis, individuals with a schizophrenia diagnosis still endorse considerable mood symptoms (Conley, Ascher-Svanum, Zhu, Faries, & Kinon, 2007); likewise, a large percentage (~58%) of individuals with bipolar disorder without psychotic features still endorse some psychotic-like experiences particularly in the form of unusual beliefs (Dunayevich & Keck Jr, 2000). Many researchers have posited a continuum between schizophrenia and bipolar disorder, particularly in light of empirical evidence suggesting a substantial percentage of schizophrenia and bipolar cases falling on a continuum of what is conventionally diagnosed as schizoaffective disorder (Keshavan et al., 2011; Pearlson, 2015).

In both schizophrenia and bipolar disorder, reinforcement learning alterations are believed to contribute to social and role impairment, which in turn substantially limit functional recovery. However, whether specific aspects of reinforcement learning are differentially impacted in schizophrenia and bipolar disorder remains unclear. Some

evidence suggests that over and above diagnostic group, symptom severity may predict alterations in reinforcement learning (Strauss et al., 2015). Research evaluating reward processing more broadly suggests possible opposite responses to reward and punishment, differences in risky reward-seeking behavior and more pronounced deficits in associative learning in schizophrenia compared to bipolar disorder. However, real-world reward-seeking behavior, particularly in disadvantageous situations like substance use, and impulsive behavior are elevated in the clinical presentations of *both* schizophrenia and bipolar disorder. Genome-wide association studies (GWAS) have identified shared genetic tendency towards risky reward-seeking behavior in both schizophrenia and bipolar disorder (Hindley et al., 2021).

Previous studies suggest that individuals with schizophrenia have specific reinforcement learning deficits, including impairments learning from positive but not negative feedback (Strauss et al., 2011) and deficits in the ability to maintain action value representations (Gold et al., 2008; Waltz & Gold, 2007, 2015). Reinforcement learning deficits specific to associating actions and rewards may underlie negative symptoms such as avolition and anhedonia, with failure to use feedback to inform behavior impeding goal-directed behavior and engagement in rewarding activities (Strauss et al., 2011).

Existing computational studies of reinforcement learning in schizophrenia vary in the model used, and several studies have a substantial percentage of participants whose data cannot be fit with more complex computational models intended to capture learning asymmetries or participant beliefs about task structure. Additionally, existing studies vary in schizophrenia medication status, which is particularly relevant given the potential for antipsychotics to impact reward-related dopaminergic signaling in the brain. Taken together, these inconsistencies limit ability to generalize findings and the ability of these datasets to inform one another.

In studies investigating reinforcement learning in bipolar disorder, evidence suggests trait hypersensitivity to reward-related stimuli and cues both behaviorally and neurally, which in turn contributes to vulnerability for mania (Alloy, Olino, Freed, & Nusslock, 2016). Reward hypersensitivity may be mood-independent, with previous work identifying behavioral and neural reward hypersensitivity in healthy individuals with hypomanic traits (Damme, Young, & Nusslock, 2017), euthymic bipolar disorder patients and even in depressed bipolar disorder patients (Chase et al., 2013). However, findings are mixed across studies. There is evidence of greater resting-state functional connectivity amongst reward-relevant regions overall, but state-related decreases in hypersensitivity in bipolar depression (Satterthwaite et al., 2015). Computational modeling provides a unique opportunity to rigorously test the reward hypersensitivity theory by explicitly probing reward-mediated learning and comparing these measures to loss-related behavior. Despite the benefits of the computational approaches, this framework has been underutilized in the investigation of reward hypersensitivity in bipolar disorder. Critically, reward hypersensitivity may distinguish schizophrenia and bipolar disorder, but conclusions about shared versus distinct deficits are limited by a lack of transdiagnostic computational studies.

While differences in integrating positive feedback may distinguish schizophrenia and bipolar disorder, cognitive deficits may be a transdiagnostic mechanism that leads to altered integration of feedback, suboptimal retention of action values, and deficits in action selection. My objective was to evaluate shared and unique deficits in reinforcement learning in schizophrenia and bipolar disorder, using a publicly available dataset from the Consortium for Neuropsychiatric Phenomics (CNP). Furthermore, I sought to identify clinical and cognitive correlates associated with identified deficits. I first established general patterns of group differences using traditional summary statistics across all trials of the task. I then further probed differences in summary statistics by fitting a computational model

using trial-by-trial data with separate parameters for action value updating after positive feedback and action value updating after negative feedback, and a value representation retention parameter; the use of this model in other clinical samples suggested a high likelihood that all participants' data could be fit with this model.

I hypothesized that value updating of action values after positive feedback (i.e., gains), but not value updating after negative feedback (i.e., losses), would be *lower* in schizophrenia compared to healthy controls and bipolar disorder. Alternatively, in bipolar disorder value updating after positive feedback, but not negative feedback, would be *higher* relative to healthy controls and schizophrenia. Across diagnostic groups, I predicted that lower working memory performance in bipolar disorder and schizophrenia would be associated with a worse retention of value representations.

I also pursued an exploratory aim and leveraged the large healthy control group to assess the relationship between variability in subclinical symptoms present in healthy controls and reinforcement learning parameters. Evidence suggests subclinical levels of psychotic symptoms and mood symptoms are present in otherwise healthy controls; spectrum-based approaches are consistent with neurodevelopmental framework wherein individuals possess varying degrees of genetic risk, which in turn results in a range of endophenotypes and confers risk for broader phenotypes than are associated with clinically significant psychosis alone (Cannon & Keller, 2006). Previous studies have found alterations in cognition, brain and behavior even in individuals with subthreshold symptoms (Barber, Lindquist, DeRosse, & Karlsgodt, 2018; Hegarty et al., 2019; Karlsgodt, Niendam, Bearden, & Cannon, 2009). Specifically, I examined the impact of impulsivity, bipolar traits, and anhedonia, which have been associated with reinforcement learning in schizophrenia and bipolar disorder populations.

METHODS

Project Infrastructure

This study leveraged a publicly available dataset from the Consortium for Neuropsychiatric Phenomics (CNP). The CNP investigated neuropsychological phenotypes and mechanisms on a genome-wide and phenome-wide scale. The CNP consisted of two research cores, the Human Translational Applications Core (HTAC) and Translational Methods & Facilities Core, along with a coordinating center. The HTAC LA2K study was a large project that recruited a sample of approximately 1200 adults from the Los Angeles area to participate in broad phenotyping (clinical and behavioral interviews, cognitive testing) and genetic testing. Within that study was the LA5C neuroimaging study, which included approximately 50 individuals each with schizophrenia (SZ), bipolar disorder (BP), and attention deficit/hyperactivity disorder. I used data from all healthy controls from the LA2K study and data from SZ and BP collected as part of the LA5C study nested within the parent LA2K study. All data were collected at UCLA.

Participants included 1101 healthy controls (HC) from the community, 58 SZ, and 49 BP (*Table 4*). Participants aged 21-50 were recruited via community advertisement in Los Angeles; racial and ethnic distribution was consistent with the diverse greater Los Angeles area. Exclusion criteria included: (1) neurological disease, (2) head injury with loss of consciousness of cognitive sequelae, (3) substance dependence within past 6 months, (4) contraindications for MRI (e.g., claustrophobia, pregnancy, metal implant), (5) vision impairment that sufficiently impeded ability to complete tasks, (6) left-handedness. Healthy controls were additionally excluded for history of major mental illness or ADHD, and current mood or anxiety disorder. Participants completed a urine drug screen (testing for cannabis, amphetamines, opioids, cocaine, benzodiazepines) on the day of testing, and were

excluded if results were positive.

Clinical Measures and Working Memory Measure

Diagnoses were obtained via SCID-IV. SZ and BP participants completed the Brief Psychiatric Rating Scale (BPRS)(Ventura, Nuechterlein, Subotnik, & Gilbert, 1995), Scale for the Assessment of Negative Symptoms (SANS)(Andreasen, 1989), Scale for the Assessment of Positive Symptoms (SAPS) (Andreasen, 1984), Young Mania Rating Scale (YMRS) (Young, Biggs, Ziegler, & Meyer, 2000), and Hamilton Psychiatric Rating Scale for Depression (HAMD-28) (Williams, 1988).

Additionally, SZ, BP and HC completed additional clinical and trait measures, including the Hopkins Symptom Checklist (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974), Barratt Impulsiveness Scale (BIS-11) (Patton, Stanford, & Barratt, 1995); Dickman Functional and Dysfunctional Impulsivity Scale(Claes, Vertommen, & Braspenning, 2000), Scale for Traits that Increase Risk for Bipolar II Disorder, Hypomanic Personality Scale (HPS) (Eckblad & Chapman, 1986) and the Chapman Scales (Perceptual Aberrations, Social Anhedonia, Physical Anhedonia) (Chapman, Chapman, & Raulin, 1976). These additional measures can capture variability in traits associated with schizophrenia and bipolar disorder even in healthy controls.

Table 4. Demographics of Full CNP Sample

	Group		
	HC (N=1101)	**SZ (N=58)**	**BP (N=49)**
Sex			
% Male	46.7	75.9	57.1
% Female	53.3	24.1	42.9
Race			
% Native American	0.9	3.4	2.0
% Black or African American	22.2	22.4	8.2
% Asian	0.4	1.7	0
% White	73.8	65.5	75.5
% Native Hawaiian/Pacific Islander	0.1	3.4	0
% Mixed Race or Other	1.7	6.9	14.3
Age			
Minimum	21	21	21
Maximum	50	49	50
Average (SD)	31.24 (8.47)	35.84 (8.61)	35.29 (9.01)
Diagnosis			
% Schizophrenia	--	79.3	--
% Schizoaffective	--	20.7	--
% BP 1 without psychotic features	--	--	93.3
% BP 1 with psychotic features	--	--	6.1

Participants also completed a spatial working memory capacity task (SCAP) (Glahn et al., 2002) both inside the scanner and as part of behavioral testing. I used data from behavioral testing sessions as only a subset of controls completed scans. The SCAP varies working memory load; participants were shown an array of 1, 3, 5, or 7 yellow circles pseudorandomly around a central fixation. After delay, participants were shown a green circle and had to indicate whether the green circle was in the same position as one of the yellow circles in the previously displayed target array. I used participant data from the SCAP as a performance-based measure of working memory ability in analyses of cognitive

correlates of reinforcement learning deficits. Given limited variability in working memory capacity, I used percent correct as the primary index of working memory.

Medications

Antipsychotic medications, commonly prescribed in both schizophrenia and bipolar disorder, typically function as dopamine antagonists. Antipsychotic medication may subsequently alter midbrain dopaminergic signaling and dopamine-mediated reward behaviors. To be eligible for the study, schizophrenia participants had to be medicated with antipsychotics. Medication data was collected during clinical assessment; however, medication history was self-reported leading to incomplete medication histories for all participants. In the bipolar group, medications prescribed varied including antipsychotics, mood stabilizers, anticonvulsants, and antidepressants, all of which have the potential to impact reward and cognitive neural circuitry. Approximately 45% of the bipolar sample were prescribed antipsychotics.

I analyzed medication as a potential confound and examined the effect of medication in secondary analyses. I calculated chlorpromazine-equivalent dose for antipsychotics (calculated (Woods, 2003) as in our existing studies) based on provided medication histories. In the schizophrenia group, seven participants (approx. 13%) had incomplete medication data and chlorpromazine equivalents were not able to be calculated. In the bipolar group, given the small number of participants in the other medication categories other than antipsychotics, the study was not sufficiently powered to evaluate the impact of each class of medication.

CNP Reinforcement Learning Task

The CNP Reinforcement Learning Task (CNP RL) was programmed in e-Prime 2.0. The task consisted of a probabilistic selection task (PST) (Frank, Seeberger, & O'reilly,

2004) followed by a probabilistic reversal learning task (PRLT)(Swainson et al., 2000). During the task, participants were instructed to choose between two abstract stimuli; feedback was probabilistic such that one stimulus was more associated with positive feedback than the other (e.g., one stimulus is correct 80% of the time, the other is correct 20% of the time). Feedback appeared as “Correct” in green font, or “Incorrect” in red font above the stimuli. The task was self-paced, and responses were made by left or right keyboard arrow button press. Participants completed four blocks: (1) Training 1, (2) PST (28 trials), (3) Training 2 (40 trials), and (4) PRLT (40 trials).

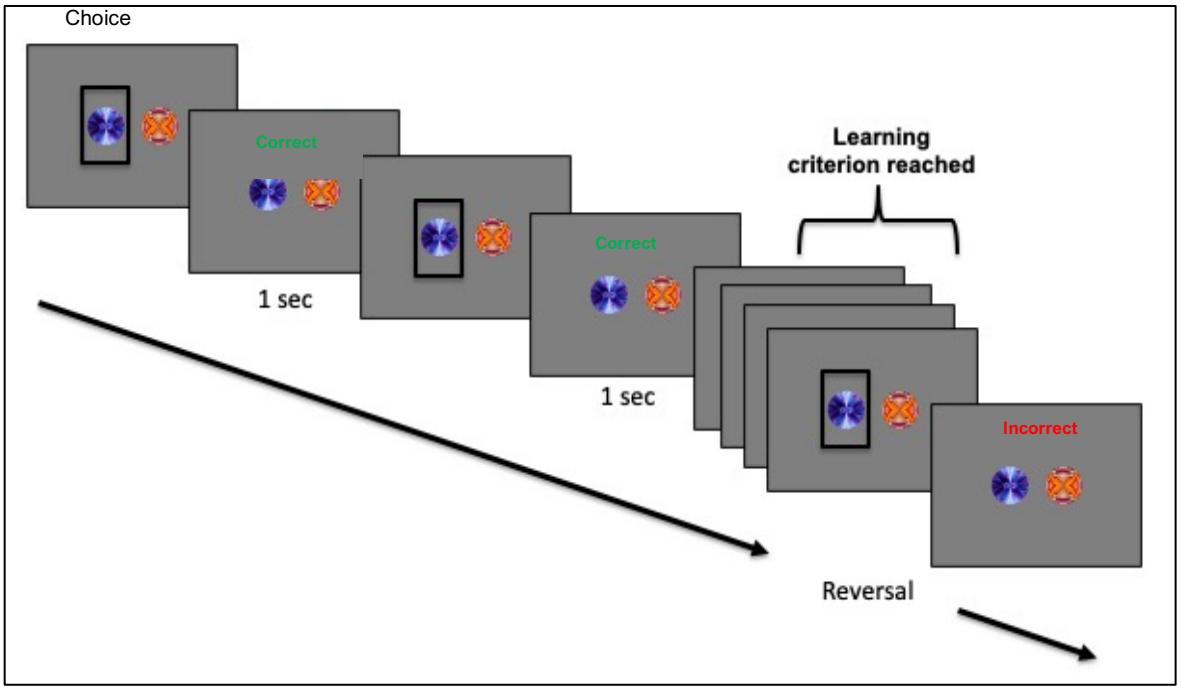
During the first training phase of the task, participants were presented with pairs of fractal stimuli with varying probabilities of reward associated with each stimulus (100/0 deterministic pair, 80/20 probabilistic pair, 70/30 probabilistic pair, 60/40 probabilistic pair). A learning criterion was enforced (70%, 65%, 60%, and 55% for each respective pair). Accuracy was calculated once 60 trials were completed, and then was continuously calculated for each subsequent trial until criterion had been reached for each pair. If criterion was not reached at 80 trials, participants completed an additional 80 trials (max 160 trials) before moving to PST trials.

During PST trials, to determine if participants favor learning from positive feedback or negative feedback, cards were then recombined across pairs, such that every stimulus is presented in a pair with every other stimulus (e.g.,70/20, 70/40, 70/60, etc.). Feedback sensitivity is evaluated based on response to the novel stimulus pairs, such that a positive bias is characterized by preferential selection of high probability stimuli and negative bias is characterized by preferential avoidance of low probability stimuli.

Following completion of the PST, participants completed another training phase of 40 trials of the original pairs (10 trials each of 100/0, 80/20, 70/30, and 60/40 pairs). The PRLT, similar to other reversal learning tasks, requires participants to modify behavior in response

to changing stimulus-outcome contingencies (*Figure 5*). Following the second training, the contingencies in the 100/0 and 70/30 pairs reversed, whereas the 80/20 and 60/40 contingencies remained consistent. Participants completed 10 trials of each pair following reversal (40 trials total during the PRLT).

Figure 5. PRLT Schematic. Participants choose stimuli and receive probabilistic feedback; for example, the left stimulus above is correct 70% of the time. After reaching a learning criterion, participants completed probabilistic selection trials and additional training trials, followed by a reversal in the 100/0 and 70/30 pairs. Intertrial interval was 500 msec.



CNP RL Cleaning Rules

Trials were eliminated for unrealistically low (e.g., 100ms) or very long (e.g., 6000ms) reaction times. Because PST and PRLT phases of the task only begin once a learning criterion has been reached, it is assumed all participants with PST and PRLT task data appropriately learned the rules of PST and PRLT phases of the task. For summary data analysis, individuals were only included if they had responses for every trial of the PST as each recombination is presented for only one trial.

Analysis of PRLT summary data

To establish patterns of behavior across all trials, I evaluated group differences in task summary statistics from both PST and PRLT phases of the RL task. For PST trials, positive feedback sensitivity and negative feedback sensitivity were defined in accordance with Frank et al, 2004; positive feedback sensitivity is the proportion of times the highest probability stimulus is chosen during the recombination phase (i.e., 80% stimulus chosen from 80/70, 80/60, 80/40, 80/30, and 80/0 pairs) and negative feedback sensitivity is the proportion of times the lowest probability stimulus was avoided during the recombination phase (i.e., 20% stimulus avoided in 20/100, 20/70, 20/60, 20/40, and 20/30 pairs).

For the PRLT, I evaluated differences in switches in choice behavior after first correction following reversal, which can reflect differences in behavior modification following feedback. Given that the PRLT phase of the task had one reversal, I was not able to evaluate group differences in total number of reversals.

Variables were entered into ANCOVAs covarying for age and sex to evaluate group differences in positive feedback sensitivity, negative feedback sensitivity and switches after first correction; post-hoc tests were conducted to determine the nature of group differences using Bonferroni correction for multiple comparisons. Age and sex were centered to group mean for all analysis, as well as any predictors that took on non-meaningful zero values.

Computational modeling of trial-by-trial data

Individual choice data was fit using a forgetting reinforcement learning model to estimate retention of value representations (γ), updating after positive feedback (i.e., "Correct") (Δ_+), and updating after negative feedback (i.e., "Incorrect") (Δ_0) (Barraclough, Conroy, & Lee, 2004; Groman, Rich, Smith, Lee, & Taylor, 2018). Δ_0 takes on a negative value, with interpretation of differences in Δ_0 based on absolute value of the parameter.

Parameters were estimated using maximum likelihood with the fminsearch function in MATLAB (2019a).

$$If\ R = 1 \qquad Q_{t+1} = \gamma Q_t(a_t) + \Delta_+$$

$$If\ R = 0 \qquad Q_{t+1} = \gamma Q_t(a_t) + \Delta_0$$

$$\Delta = \alpha(R - Q_t(a_t))$$

Estimated parameters (Δ_+, Δ_0, γ) were then entered as outcome variables in a MANCOVA with age and sex as covariates, conducted in Stata v 16.1. Age and sex were centered to group for individuals included in the computational sample. Post-hoc univariate tests were conducted to determine the nature of significant group differences, using Bonferroni correction for multiple comparisons.

I conducted post-hoc analyses with Δ_+, Δ_0, and γ parameters fit separately for training versus reversal phases of the task. I hypothesized that in the reversal phase, high retention may be less adaptive, where high retention during training would be ideal as participants learn to discriminate pairs. I sought to evaluate if SZ showed worse retention overall, or if it was specific to task phase. I conducted a repeated-measures MANCOVA to evaluate whether there was a within-subjects effect of task phase, and whether there was an interaction between group and task phase.

RESULTS

Group Differences in Summary Statistics

On the PST, ANCOVA detected statistically significant differences in overall positive feedback sensitivity ($F(3,1198) = 7.273$, $p<0.001$), such that SZ show significantly lower probability of choosing stimuli associated with positive feedback relative to HC ($p=0.003$). BP did not differ significantly from SZ or HC (*Figure 6*). There were no significant group

differences in negative feedback sensitivity.

On the PRLT, ANCOVA detected statistically significant differences in number of switches following reversal, such that HC performed significantly fewer switches following 1st correction compared to SZ ($p<0.001$) (*Figure 7*). There were no significant group differences in switches following 1st correction between HC and BP groups, or BP and SZ groups. Within SZ and BP groups, there were no significant correlations between summary statistics, symptoms, or working memory performance.

Figure 6: Group Differences in PST Positive Feedback Sensitivity. * $p < 0.05$ corrected for multiple comparisons.

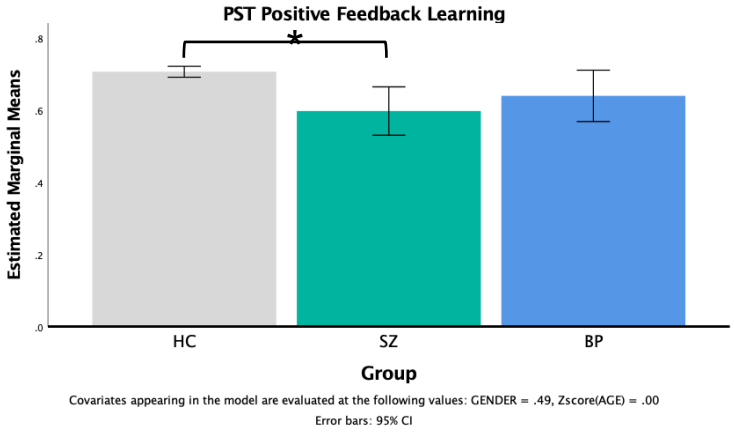
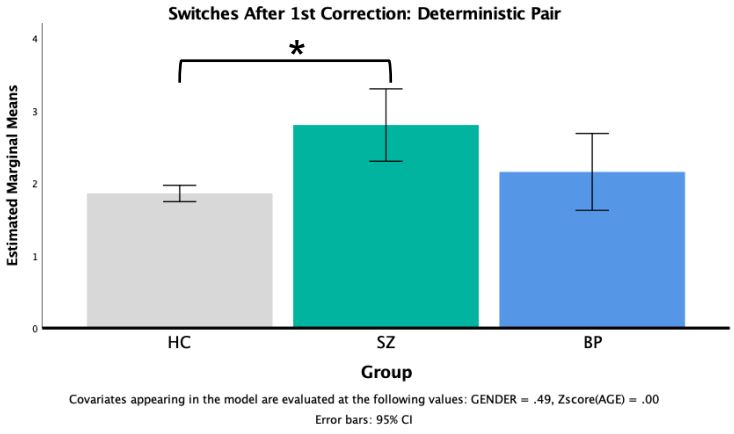


Figure 7: Group Differences in PRLT Switches After First Correction Post-Reversal. * $p < 0.05$ corrected for multiple comparisons.



Group Differences in RL Computational Parameters

Retention (γ) is bounded, with values between 0 and 1. However, Δ_+ and Δ_0 are not bounded and can therefore take on extreme values. Outlier exclusion was based on z-scores (z-score > 3.2), resulting in a slightly smaller sample (*Table 5*).

MANCOVA detected statistically a significant difference in Δ_+, Δ_0, and γ between HC, SZ, and BP groups, controlling for age and sex ($F(6, 2{,}306)=2.32$ p =.031, Wilks' Λ = .988, partial η^2 = .006). Post-hoc univariate tests detected that retention of action values (γ) was significantly greater in BP compared to HC (p=.048), with no significant differences between HC and SZ, or SZ and BP (*Figure 8*). Value updating after positive feedback (Δ_+) was significantly lesser in SZ compared to HC (p=.046), with no significant differences between HC and BP, or SZ and BP (*Figure 9*).

Table 5. CNP Demographics- Participants with Usable Computational Data

	Group		
	HC (N=1058)	**SZ (N=55)**	**BP (N=48)**
Sex			
% Male	46.9	75.9	57.1
% Female	53	24.1	42.9
Race			
% Native American	0.9	3.6	2.1
% Black or African American	22.2	21.4	8.3
% Asian	0.4	1.7	0
% White	73.7	64.3	75.5
% Native Hawaiian/Pacific Islander	0.1	3.4	0
% Mixed Race or Other	2.2	9	14.6
Age			
Minimum	21	21	21
Maximum	50	49	50
Average (SD)	31.40 (8.47)	35.89 (8.59)	35.56 (8.91)
Diagnosis			
% Schizophrenia	--	82.1	--
% Schizoaffective	--	17.9	--
% BP 1 without psychotic features	--	--	93.8
% BP 1 with psychotic features	--	--	6.2

Figure 8: Group Differences in γ (Retention). * p < 0.05 corrected for multiple comparisons.

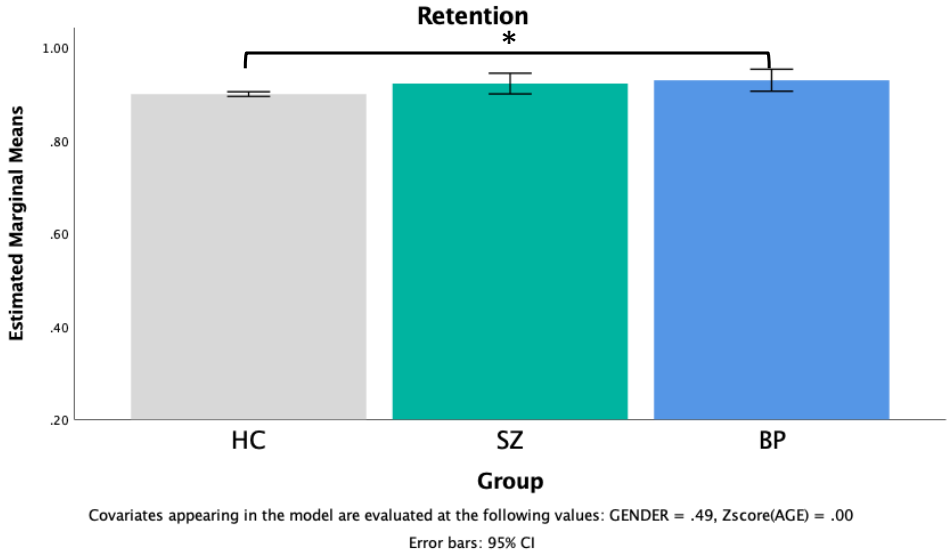
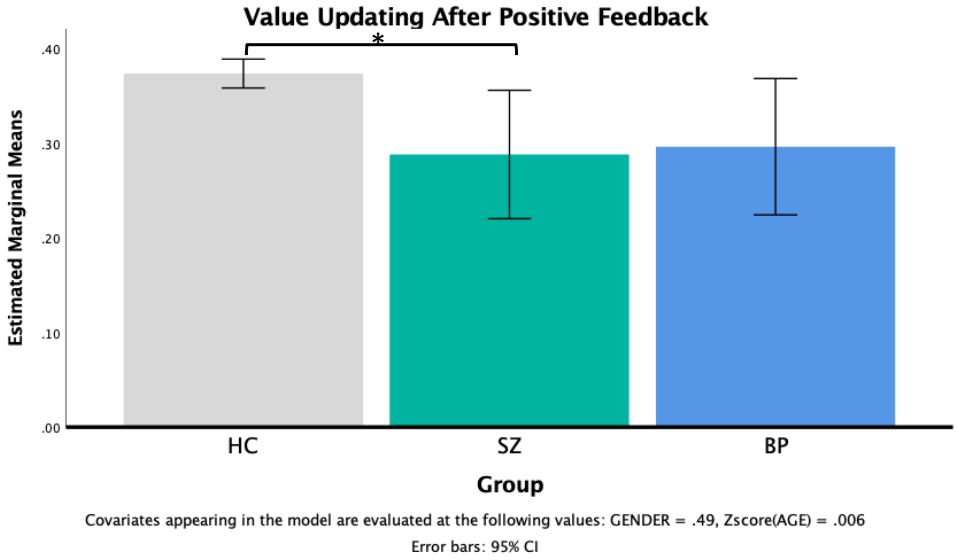


Figure 9: Group Differences in Δ_+ (Value Updating After Positive Feedback). * p < 0.05 corrected for multiple comparisons.



Symptom Predictors of RL Computational Parameters

In SZ participants, more severe negative symptoms were associated with reduced updating of action values after positive feedback (Δ_+), over and above age and sex ($F(3,51)=2.93$ $p=0.042$) (*Figure 10*). In SZ participants, more severe negative symptoms

were associated with greater updating of action values after negative feedback (Δ_0), over and above age and sex, (F(3,51)=2.81, p=0.048) (with more negative values equating to greater updating, resulting in a negative slope) (*Figure 11*).

Figure 10. Negative Symptom Severity in SZ and Δ_+ (Value Updating after Positive Feedback). Depicted with 95% CI.

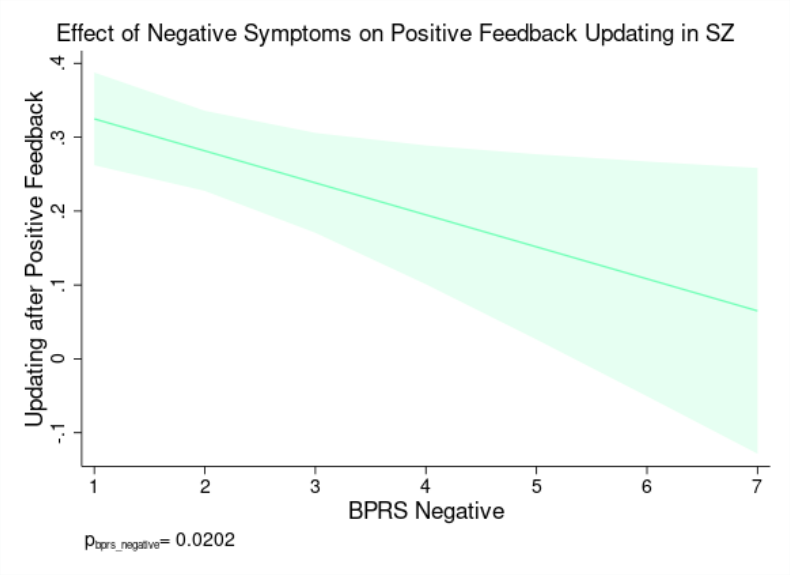
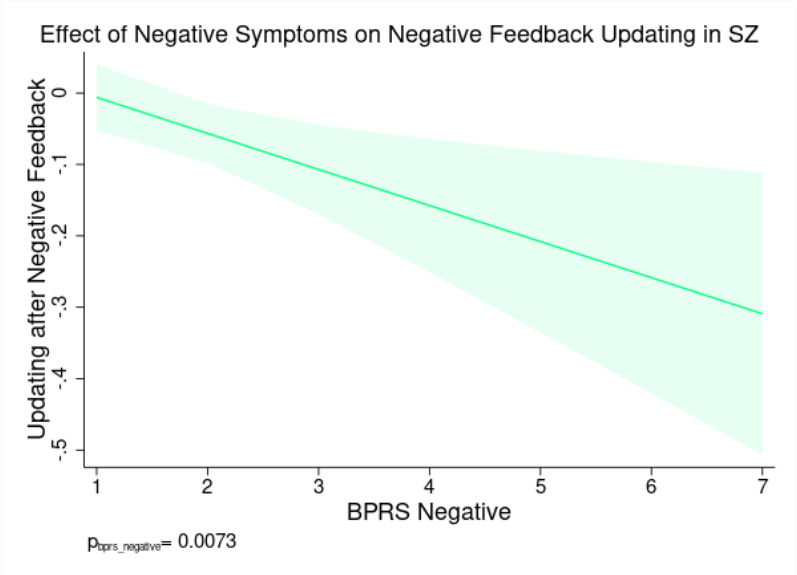
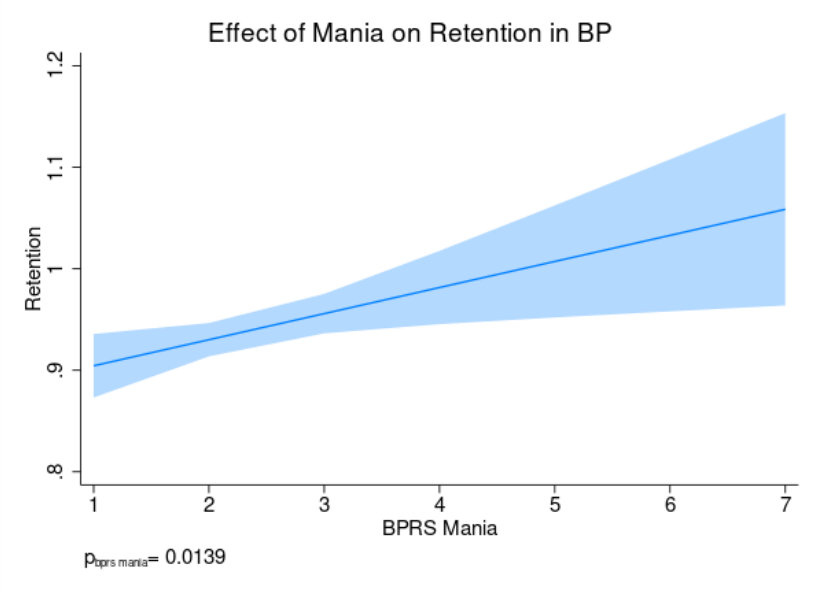


Figure 11. Negative Symptom Severity in SZ and Δ_0 (Value Updating after Negative Feedback). Depicted with 95% CI.



In BP, mania was not significantly predictive of Δ_+ or Δ_0. However, more severe mania was associated with greater retention of action values (γ), over and above age and sex (F(3,44)=3.25, p=0.03) (*Figure 12*).

Figure 12. Mania Severity in BP and γ (Retention). Depicted with 95% CI.



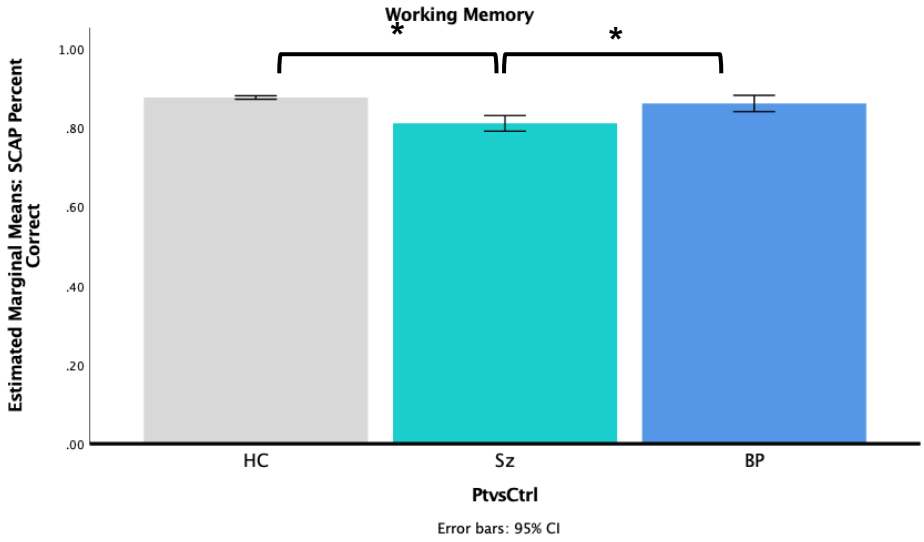
Working Memory and RL Computational Parameters

ANCOVA covarying for age and sex detected significant group differences in working memory performance as indexed by percent correct on the SCAP (F(2,1152)=13.425, p < 0.001). Post-hoc univariate tests revealed that SZ had significantly lower SCAP performance compared to HC (p < 0.001) and BP (p=0.002) (*Figure 13)*. There were no significant differences in working memory performance between HC and BP.

Due to significant working memory performance differences between SZ and BP, patient group was included as a predictor in working memory analyses to determine if there was an effect of working memory performance on γ over and above patient group. A linear

regression predicting γ from patient group, age, sex, and working memory performance was non-significant (F(4,94)=0.54, p=0.71).

Figure 13: Group Differences in Working Memory Performance. * p < 0.05 corrected for multiple comparisons.



Medication Effects

There were no significant correlations between medication, as indexed by CPZ equivalents, and any computational parameters. As a result, medication was not controlled for or evaluated as a confound in any additional analyses.

Exploratory Analyses

Group Differences in RL Parameters: Schizophrenia and Bipolar Disorder

I removed SZ individuals with a schizoaffective diagnosis and BP individuals with a psychotic features specifier, and examined if there were group differences between schizophrenia and bipolar disorder once individuals with substantial overlap in

symptomatology were removed. Similar to my results including all diagnoses, MANCOVA detected statistically a significant difference in Δ_+, Δ_0, and γ between HC, SZ, and BP groups, controlling for age and sex (F(6, 2,306)=2.32 p =.031, Wilks' Λ = .988, partial η^2 = .006). Post-hoc univariate tests detected that retention of action values (γ) was significantly greater in BP compared to HC (p=.031), with no significant differences between HC and SZ, or SZ and BP. Value updating after positive feedback (Δ_+) was significantly lesser in SZ compared to HC (p=.039), with no significant differences between HC and BP, or SZ and BP. There were no differences in Δ_0.

RL Computational Parameters: Training versus Reversal Trials

In general, the greater the number of trials, the better the solution that the algorithm can find. By looking at each phase of the task separately, there were substantially fewer trials to be modeled and a large number of participant data could not be fit (*Table 6*).

A repeated measures MANCOVA detected a significant within-subject effect of task phase (training vs reversal) on Δ_+, Δ_0 and γ. The interaction between group and task phase was nonsignificant, indicating no significant difference in the pattern of change in Δ_+, Δ_0 and γ from training to reversal phases based on group. Between-subject effects in Δ_+, Δ_0 and γ were nonsignificant. Within-subjects effects for $\Delta_0$were nonsignificant. However, there was a significant within-subjects effect of task phase on γ, such that collapsed across group, individuals increased their retention in the reversal phase compared to the training phase (p<.001) (*Figure 14*). There was also a significant within-subjects effect of task phase on Δ_+, such that collapsed across group, value updating after positive feedback decreased in reversal trials compared to training trials (*Figure 15*).

Healthy Controls: Sub-clinical Traits and RL Parameters

There were no significant correlations between RL parameters (Δ_+, Δ_0 and γ), bipolar traits as indexed by the Scale for Traits that Increase Risk for Bipolar II Disorder, and psychotic-like experiences as indexed by the Chapman Scales (Perceptual Aberrations, Social Anhedonia, Physical Anhedonia) (Chapman et al., 1976).

Table 6. CNP Demographics- Participants with Usable Computational Data: Training vs Reversal

	Group		
	HC (N=672)	**SZ (N=35)**	**BP (N=29)**
Sex			
% Male	48.2	75	69
% Female	51.8	25	31
Race			
% Native American	0.1	0	10.3
% Black or African American	23.1	30.6	3.4
% Asian	0.4	2.8	0
% White	73.5	58.3	75.9
% Native Hawaiian/Pacific Islander	0.1	2.8	0
% Mixed Race or Other	2.2	2.8	10.3
Age			
Minimum	21	22	22
Maximum	50	49	47
Average (SD)	31.14 (8.44)	34.89 (8.15)	34.54 (7.55)
Diagnosis			
% Schizophrenia	--	82.1	--
% Schizoaffective	--	17.9	--
% BP 1 without psychotic features	--	--	93.8
% BP 1 with psychotic features	--	--	6.2

Figure 14. Retention (y)- Training versus Reversal Trials. *p < 0.05

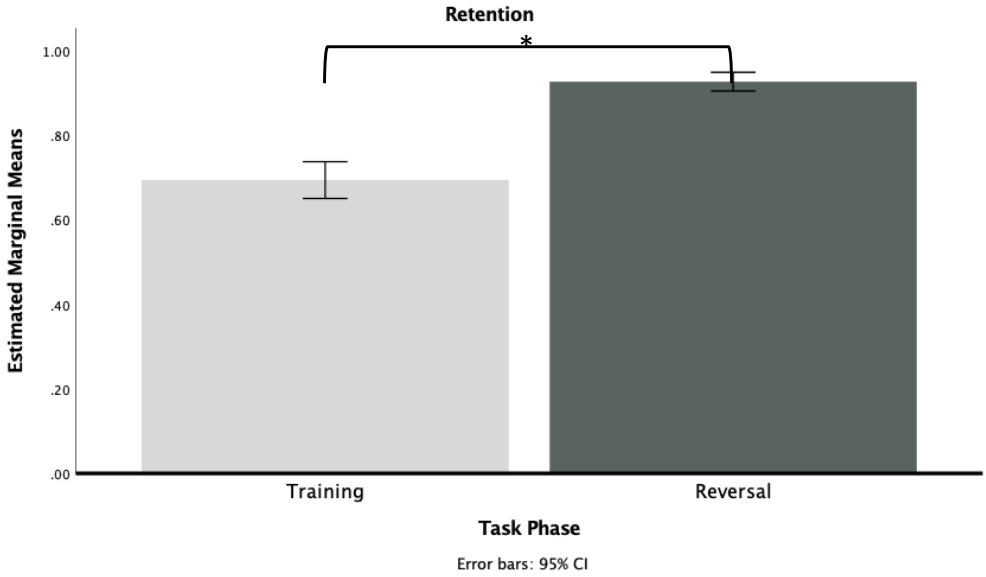
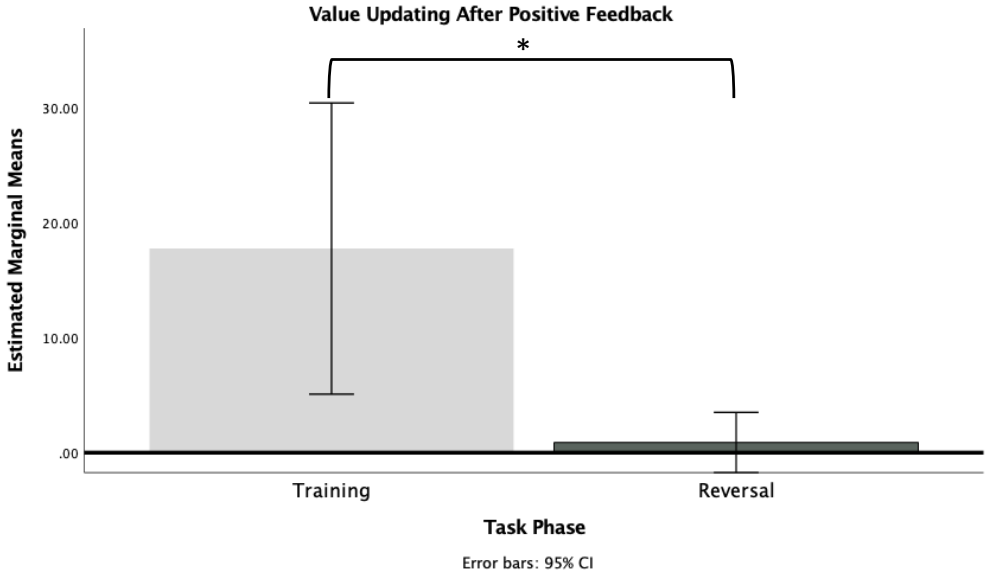


Figure 15. Value Updating after Positive Feedback (Δ_+)- Training versus Reversal Trials. *p <0.05



CONCLUSION

Summary Statistics

Analysis of summary statistics was consistent with existing literature suggesting that schizophrenia is associated with reduced positive feedback sensitivity relative to controls, preserved negative feedback sensitivity, and an overall less organized response when modifying behavior following feedback. However, while the bipolar group was qualitatively intermediate compared to healthy control and schizophrenia groups, there were no significant differences between bipolar and schizophrenia, or bipolar and healthy controls. There were no significant relationships between summary statistics, symptom measures, or working memory performance, which partially could be attributed summary statistics providing coarser metrics of participant reinforcement learning processes.

The objective of this study was not to replicate existing findings using summary statistics, but these findings further highlight selective and not unilateral deficits, particularly in schizophrenia, that bear further investigation. They also highlight that even at a broader level, there are no clear differences in performance between schizophrenia and bipolar; we were able to extend upon these findings and identify if actual symptomatology is more predictive of reinforcement learning deficits over and above diagnostic category in our computational analyses.

Computational Analyses

Computational analyses provided more detailed information about the nature of reinforcement learning deficits in schizophrenia and bipolar disorder by allowing us to evaluate *how* individuals were using feedback to guide future behavior. Furthermore, we could then link these specific processes with symptomatology to characterize within-group variability in our patient groups. Schizophrenia may be associated with selective deficits

updating value based on positive but not negative feedback relative to healthy controls; however, there were no differences in any parameters between schizophrenia and bipolar groups. While reinforcement learning is believed to be differentially impacted in schizophrenia and bipolar disorder, results indicate considerable heterogeneity within patient groups and substantial overlap in the distributions of parameters between schizophrenia and bipolar disorder. The lack of clear differences between schizophrenia and bipolar is consistent with dimensional approaches focused on deficits core to actual symptomatology that can cut across disorders, rather than deficits found on average within separate diagnostic groups.

Consistent with my hypotheses, elevated avolition in schizophrenia moderates updating after positive feedback, with diminished ability to update after positive feedback as avolition increases in severity. However, avolition was also associated with *increased* updating after negative feedback. These results may indicate that schizophrenia participants with pronounced negative symptoms have a less organized strategy in response to both positive and negative feedback, suggesting that greater severity of negative symptoms may be related to difficulties integrating feedback broadly. This general failure to appropriately use feedback, regardless of valence, could contribute to disorganized representations of action value that are updated ineffectively and behavioral inertia. On the other hand, those with less pronounced negative symptoms may be better able to “rely” on learning from negative feedback to drive behavioral change and respond to their environments, even if behaviors remain ineffective.

Contrary to my hypotheses, severity of manic symptoms in bipolar disorder was not associated with increased value updating after positive feedback. BP with elevated mania did show greater retention of action values. While retention of action values does facilitate continued engagement in rewarded behaviors, very high retention could reflect a suboptimal

strategy wherein participants stay with the same decision even when contingencies shift and they should alter their behavior accordingly. In real-world settings, this tendency may be seen in continued engagement in risky behavior in mania, even when those behaviors are no longer rewarding or likely to yield positive outcomes.

There was no significant relationship between retention and working memory performance in SZ and BP, even when controlling for baseline group differences in working memory performance. I chose to use a performance-based metric of working memory that previously has been shown to differentiate patient groups from healthy controls (Poldrack et al., 2016). However, it may not fully index working memory as a construct, which may limit its predictive utility. Retention may also be mediated by other cognitive constructs that I did not examine, including attention. Additionally, use of a chronic sample may have led to less variability in working memory functioning overall, further limiting my ability to probe how working memory differences account for variability in reinforcement learning parameters like retention.

Post-hoc analyses evaluating value updating after positive feedback updating, value updating after negative feedback, and retention in training versus reversal trials indicated that value updating after positive feedback and retention were greater in reversal trials compared to training trials. This suggests that across groups, individuals were changing their integration of positive feedback at different phases of the task. Greater retention and greater value updating after positive feedback may facilitate performance on the reversal stage of the task and support individuals in modifying choice behavior to optimize positive feedback. However, training versus reversal trial results should be interpreted very cautiously as only the half of the sample was able to be fit due to fewer trials per task phase.

Exploratory analyses examining the relationship between sub-clinical traits and

computational parameters in healthy controls were non-significant, suggesting no continuity across the spectrum of severity in the symptom-parameter relationships demonstrated in the patient groups. However, the impact of sub-clinical traits on performance has been shown to have a very small effect size. Though the healthy control sample was large, I still may have been underpowered to detect these relationships. Additionally, studies demonstrating a relationship between sub-clinical traits and performance typically use measures like the Community Assessment of Psychic Experiences (CAPE) (Mark & Toulopoulou, 2016) or the Prodromal Questionnaire- Brief (PQ-B) (Loewy, Pearson, Vinogradov, Bearden, & Cannon, 2011), which distinguish psychotic-like experiences from their associated distress and provide more information about the overall structure of psychotic-like experiences. The traits measures included in the CNP do not provide the same degree of specificity and have not been used as extensively in the literature to probe sub-threshold symptoms.

Limitations and Future Directions

Findings in the bipolar group should be interpreted in light of recruitment across phase of illness (i.e., mania, depression, and euthymia) in the CNP sample. Bipolar disorder is associated with trait reward hypersensitivity, which would suggest that behavioral responses in pursuit of and in response to reward may be stable across phase of illness (Alloy et al., 2016). However, state-related changes during depressive episodes in bipolar disorder may also impact reinforcement learning (Satterthwaite et al., 2015). Given the small bipolar sample size, limited information about phase of illness at time of actual testing, and small effect sizes, I was not able to evaluate the impact of phase of illness on reinforcement learning.

The computational model fit in this study does not account for a participant's beliefs

or updating of beliefs. “Beliefs” are probability distributions about particular events that may or may not be consciously accessible. The mean of the distribution corresponding to the expected state. Individuals may have more or less variance within these probability distributions, which is framed as uncertainty (more variance) versus precision (less variance. Within the predictive coding framework, beliefs are organized hierarchically with prediction errors serving as low level beliefs (i.e., uncertainty related to information) that propagate upwards through the hierarchy (i.e., uncertainty related to environmental volatility) (Valton, Romaniuk, Steele, Lawrie, & Seriès, 2017). Belief-updating models can be particularly relevant when attempting to characterize behavior on reversal learning tasks, where individuals must update beliefs in response to feedback indicating a switch in previously established contingencies.

Previous studies specifically in schizophrenia have accounted for belief updating using models such as Hierarchical Gaussian Filters (HGF) (Deserno et al., 2020) and Hidden Markov Models (HMM) (Schlagenhauf et al., 2014). Results suggest that schizophrenia may be associated with greater estimates of environmental volatility, and tighter coupling between levels of beliefs leading to poor integration of feedback and worse performance overall. However, HGF and HMM are substantially more complex to fit, and previous work has also shown that these models do not fit the choice behavior of a large number of patients (ranging from ~30% to ~50% of the sample) (Deserno et al., 2020; Schlagenhauf et al., 2014), resulting in those patients being excluded from analyses. Eliminating substantial portions of the sample, particularly in a non-random way, as is often the case in studies using more complex models, was inconsistent with the goal of understanding sources of heterogeneity across the entire patient sample. Furthermore, evidence suggests that simpler models better capture patient choice behavior.

My results suggest that individuals in the patient groups were able to learn

associations necessary for the task, but I was not able to identify which cognitive processes mediate differences in learning processes in schizophrenia and bipolar disorder. Future studies may choose to examine additional cognitive correlates (e.g., attention) of computational parameters to gain mechanistic clarity and to identify which aspects of cognition critically support updating after feedback and retention of action values over time.

CHAPTER 4

STUDY 3: White Matter Alterations and Reinforcement Learning Deficits

INTRODUCTION

Computational models of reinforcement learning are theory-driven, with parameters believed to map on to psychological decision-making processes such as prediction and retention of action values. Moreover, these computations can be linked to specific biological mechanisms, such as dopaminergic prediction error signaling (Schultz, 1998). However, the relationship between computationally derived parameters and structural neural metrics has yet to be rigorously investigated in schizophrenia or bipolar disorder.

Striatal regions such as the nucleus accumbens (NAcc) and prefrontal regions, such as the orbitofrontal cortex (OFC), are believed to critically support reinforcement learning (Frank & Claus, 2006; Groman et al., 2019; Wallis, 2007). Dopamine release in the NAcc via the mesolimbic pathway has been associated with reward response and goal-directed behavior. In the schizophrenia literature, the role of the OFC has historically been described as maintaining representations of expected value (Waltz & Gold, 2015); however preclinical studies implicate the OFC in the representation of the current state (i.e. a cognitive representation of relevant contingencies). An individual may learn several contingencies, and the OFC may facilitate navigation to the most relevant set of contingencies for the given task (Schuck, Cai, Wilson, & Niv, 2016; Sharpe et al., 2019).

I previously established that there were no significant differences in value updating after positive feedback or retention between schizophrenia and bipolar disorder. However, heterogeneity in computational reinforcement learning processes was associated with different aspects of symptomatology typically associated with schizophrenia compared to

bipolar disorder. Schizophrenia and bipolar disorder not only have shared genetic risk, but may also show similar patterns of deficits in white matter connectivity that serve as a possible transdiagnostic predictor of computational reinforcement learning processes (Koshiyama, Fukunaga, Okada, Morita, Nemoto, Usui, et al., 2020).

Altered white matter integrity in corticostriatal and frontoparietal circuitry that has been observed in schizophrenia and bipolar disorder (Kochunov et al., 2020; D.-K. Lee et al., 2020) could, therefore, impact reinforcement learning and probabilistic reversal learning performance. To our knowledge, no studies to date have evaluated the relationship between white matter alterations and computational reinforcement learning parameters in schizophrenia and bipolar disorder, despite the relevance of corticostriatal and frontoparietal tracts in supporting key process for reinforcement learning.

In bipolar disorder, it has been suggested that reward hypersensitivity plays a causal role in mania and hypomania. In healthy controls, structural connectivity between the OFC and NAcc scaled with hypomanic traits, such that greater white matter integrity corresponded to elevated hypomanic traits (Damme et al., 2017). Hypomania and is associated with reward hypersensitivity (Martin & Potts, 2004); as such, findings about structural connectivity between the OFC and NAcc could be extended to suggest a relationship between value updating after positive feedback/reward and increased corticostriatal integrity. Previous work has also identified a relationship between trait impulsivity and greater connectivity in the accumbofrontal tract, the anatomical connection between the OFC and NAcc, in healthy controls (Ikuta, del Arco, & Karlsgodt, 2018). The accumbofrontal tract has primarily been described using histology (Rigoard et al., 2011) and was only recently found to be detectable in human brains using diffusion tensor imaging (Karlsgodt et al., 2015). As such, few studies have examined the accumbofrontal tract in clinical populations specifically.

The superior longitudinal fasciculus (SLF) is a major white matter tract with projections across wide areas of the human brain. The SLF includes anatomical connection between frontal and parietal cortices in the brain, and as such has been a tract of interest related to associative learning, and various cognitive functions such as working memory, visuospatial ability and verbal memory (Koshiyama, Fukunaga, Okada, Morita, Nemoto, Yamashita, et al., 2020). Schizophrenia has been associated reduced SLF integrity, which has in turn predicts working memory deficits (Karlsgodt et al., 2008).

I sought to identify and compare structural neural correlates of reinforcement learning processes across diagnostic groups. I investigated whether white matter alterations in reward and cognitive circuits moderate reinforcement learning alterations within schizophrenia and bipolar disorder by leveraging the advantages of computational models, which allow for precise mapping of parameters to biological and psychological processes. I focused on two specific white matter tracts based on prior literature: the accumbofrontal tract (Karlsgodt et al., 2015) and the superior longitudinal fasciculus. I hypothesized that across diagnoses*, heightened* action value updating after positive feedback will be associated with increased white matter connectivity between the OFC and NAcc. I also hypothesized that across diagnoses, greater retention in value representation will be associated with increased frontoparietal white matter connectivity.

METHODS

Sample

I utilized previously collected data from a publicly available dataset, the Consortium for Neuropsychiatric Phenomics (CNP). The CNP investigated neuropsychological phenotypes and mechanisms on a genome-wide and phenome-wide scale. The CNP

consisted of two research cores, the Human Translational Applications Core (HTAC) and Translational Methods & Facilities Core, along with a coordinating center. The HTAC LA2K study was a large project that recruited a sample of approximately 1200 adults from the Los Angeles area to participate in broad phenotyping (clinical and behavioral interviews, cognitive testing) and genetic testing. Within that study was the LA5C neuroimaging study, which included approximately 50 individuals each with schizophrenia (SZ), bipolar disorder (BP), and attention deficit/hyperactivity disorder. All data were collected at UCLA.

Imaging data was collected in all schizophrenia (SZ) and bipolar disorder (BP) participants, and the subset of 138 healthy controls (HC) participants who were part of LA5C. The CNP inclusion criteria are detailed in Chapter 3 (Study 2) included: (1) neurological disease, (2) head injury with loss of consciousness of cognitive sequelae, (3) substance dependence within past 6 months, (4) contraindications for MRI (e.g., claustrophobia, pregnancy, metal implant), (5) vision impairment that sufficiently impeded ability to complete tasks, and (6) left-handedness.

In addition to study-wide exclusion criteria, additional quality control criteria for imaging diagnostics were enforced to determine participant inclusion in the final imaging sample. Diagnostic criteria were the following: (1) Global in-mask signal: Mean (mean range= 4-12) showing a gradual decline and overall negative slope, (2) Log power spectrum of global signal time course (mean range= 2-9) showing a gradual decline and overall negative slope, (3) Signal-to-Noise (SNR) ratio (mean range 2-8) displaying no isolated spikes, (4) Mean slice intensity by time showing a range of colors, no white spots, no white lines, no large bands of color, and (5) Motion parameters: Translation (mean range 0.2-1.2mm) showing no large spikes that extend the overall range by >1mm, spikes >2mm-3mm.

Task and Parameters

I used parameters extracted from the computational models fit described in Study 2 from the CNP PRLT (*Figure 5*), previously described in detail in Chapter 3. Parameters of interest include retention of value representations (γ), and updating after positive feedback (i.e., "Correct") (Δ_+) (Barraclough et al., 2004; Groman et al., 2018). For exploratory analyses, I also used medication data and symptom and trait measures, also described in Chapter 3.

Image Acquisition

Data were collected at two different UCLA facilities: the Ahmanson Lovelace Brain Mapping Center (ALBMC) and the Staglin Center for Cognitive Neuroscience (CCN). Each were equipped with a 3T Siemens Trio MRI scanner. Functional T2*-weighted echoplanar images (EPIs) were collected with the following parameters: slice thickness = 4 mm, 34 slices, TR = 2 s, TE = 30 ms, flip angle = 90°, matrix 64 x 64, FOV = 192 mm, oblique slice orientation. Additionally, a T2-weighted matched-bandwidth high-resolution anatomical scan (same slice prescription as EPI) and MPRAGE were collected. The parameters for the high-resolution scan were: 4mm slices, TR/TE=5000/34, 4 averages, matrix = 128x128, 90-degree flip angle. The parameters for MPRAGE were the following: TR = 1.9 s, TE = 2.26 ms, FOV = 250, matrix = 256 x 256, sagittal plane, slice thickness = 1 mm, 176 slices. DTI parameters were: 64 directions, 2mm slices, TR/TE=9000/93, 1 average, 96x96 matrix, 90-degree flip angle, axial slices, b=1000.

Processing of Imaging Data

DTI Preprocessing:

Standard DTI processing was done in FSL (FMRIB Software Library(Smith et al., 2004), including eddy_correct and BET). CNP diffusion data was collected with a single

phase-encoding direction (Anterior→Posterior), which is incompatible with susceptibility distortion correction with topup in FSL, which requires reverse phase-encoding scans. For eddy current correction, eddy requires outputs from topup. Data were preprocessed using FSL eddy_correct instead of eddy.

After susceptibility distortion correction and eddy current correction, fractional anisotropy (FA), L1, L2 and L3 images were calculated using FSL DTIFit, which fits a tensor model at each voxel. I then used tract based spatial statistics (TBSS)(Smith et al., 2006). A TBSS skeleton was created based on the FA data; other modalities (radial, axial, and mean diffusivity) were projected onto the TBSS skeleton, as recommended by FSL. I then masked the skeleton using ROIs of interest from the JHU white matter atlas, in addition to the accumbofrontal tract as defined in Karlsgodt et al(Karlsgodt et al., 2015).This yielded an FA value per tract per subject that was entered into statistical models. I focused on the accumbofrontal tract (Karlsgodt et al., 2015)(AF), a corticostriatal tract between the orbitofrontal cortex and nucleus accumbens, and superior longitudinal fasciculus (SLF), a frontoparietal tract associated with working memory (Karlsgodt et al., 2008) (*Figure 16*).

Diffusion Tensor Imaging (DTI) Quality Assurance:

DTI Quality Assurance (QA) included visual inspection of the fractional anisotropy map for unusual features, inspection of the color map to check whether the directions of major tracts in the brain were appropriately colored (e.g. anterior-posterior tracts should be green, right-left tracts should be red, superior-inferior tracts should be blue), evaluation of the percentage of voxels missing in the brain and comparing against the fractional anisotropy map to determine whether large regions, and in particular, tract data is cropped, and visual inspection of each volume of the raw data for artifacts. Data was then given an overall quality score, based on the degree of cropping (0=no cropping, 1=minor cropping, 2=severe unusable cropping), motion flags, tensor direction flags, and artifact flags. In a

small percentage of participants, there were scanner-related vibration artifacts. All participants with vibration artifacts were excluded from analyses. Following QA, the final number of participants with useable DTI data and usable computational RL data were as follows: 95 healthy controls, 43 schizophrenia participants, and 40 bipolar participants.

Figure 16. White Matter Masks. Top- right and left superior longitudinal fasciulus. Bottom- right and left accumbofrontal tract.

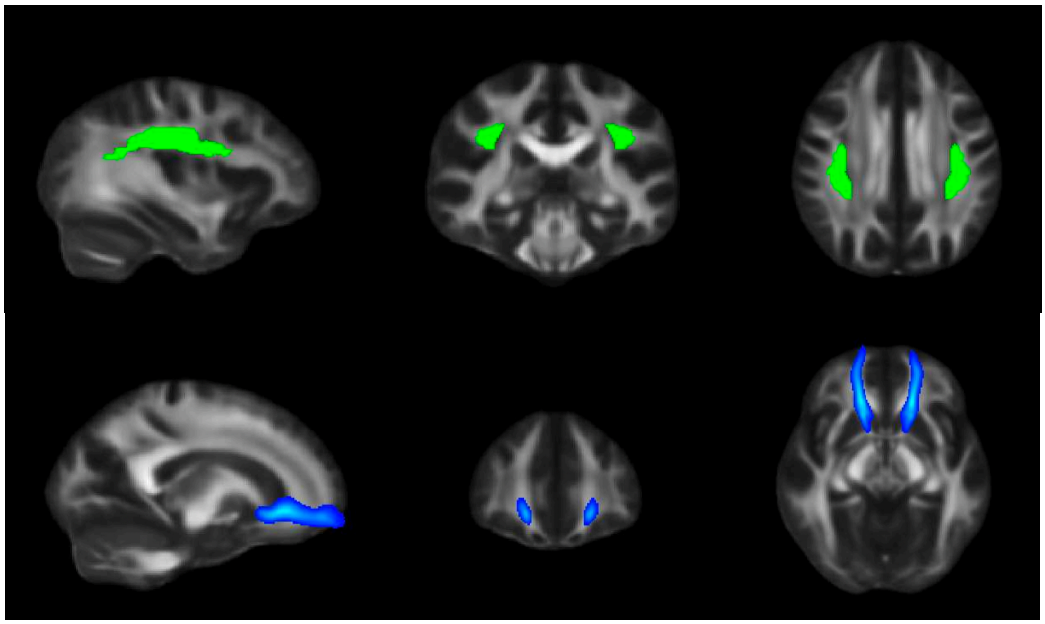


Table 7. Demographics of CNP DTI Sample

	Group		
	HC (N=95)	**SZ (N=43)**	**BP (N=40)**
Sex			
% Male	48.4	74.4	62.5
% Female	51.6	25.6	37.5
Race			
% Native American	1.1	2.3	2.5
% Black or African American	20	25.6	10
% Asian	1.1	2.3	0
% White	76.8	60.5	70.0
% Native Hawaiian/Pacific Islander	0	3.4	0
% Mixed Race or Other	1.1	7	17.5
Age			
Minimum	21	21	21
Maximum	50	49	50
Average (SD)	31.38 (8.69)	35.65 (8.77)	35.12 (9.01)
Diagnosis			
% Schizophrenia	--	76.7	--
% Schizoaffective	--	23.3	--
% BP 1 without psychotic features	--	--	95
% BP 1 with psychotic features	--	--	5

Establishing the relationship between neural correlates and computational parameters

All statistical analyses were conducted in Stata v 16.1. I first established group differences in FA values for right and left SLF, and right and left AF by conducting a MANCOVA predicting FA values from group, covarying for mean centered age and gender. I then conducted post-hoc univariate tests to determine the nature of group differences with Bonferroni correction for multiple comparisons.

I then collapsed across diagnostic groups SZ and BP, particularly in light of recent evidence suggesting shared white matter alterations in schizophrenia and bipolar disorder (Kochunov et al., 2020). I conducted a linear regression predicting frontoparietal white matter integrity from the decay parameter (γ) covarying for age, sex and scanner and a

linear regression predicting corticostriatal white matter integrity from value updating after positive feedback (Δ_+) covarying for age, sex and scanner. I conducted secondary post-hoc analyses including diagnostic status to determine the effect of action value updating and retention over and above group.

RESULTS

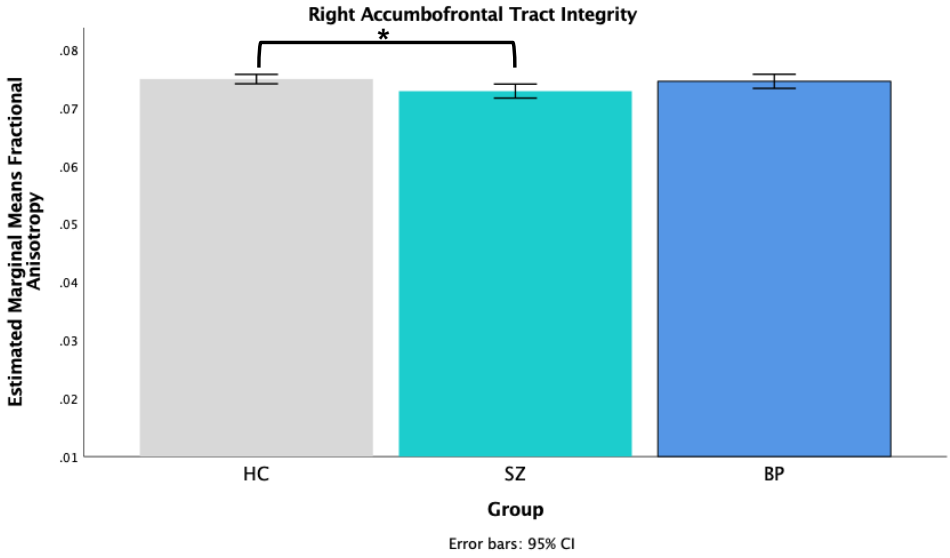
Group Differences in FA Values

MANCOVA detected a statistically significant group difference in right SLF, left SLF, right AF, and left AF FA values in HC, BP, and SZ groups, covarying for age, sex and scanner ($F(8,38)=2.115$ $p=0.03$, Wilks' $\Lambda = .907$, partial $\eta^2 = .048$). Post-hoc univariate tests detected a statistically significant difference in FA values, such that SZ showed significantly lower right AF FA ($p<0.001$), relative to HC (**Figure 17**). SZ showed reduced FA in the right and left compared to HC, but results did not survive Bonferroni correction for multiple comparisons. There were no significant differences in any FA values (right SLF, left SLF, right AF, left AF) between BP and HC, or BP and SZ.

SZ and BP: SLF Integrity and Retention of Action Values

Collapsing across diagnostic group, a linear regression predicting retention from right and left SLF FA covarying for age, sex, and scanner was non-significant ($F(7,78)=0.806$, $p=.525$). Post-hoc linear regression examining whether right and left SLF FA predict retention of action values in SZ alone was non-significant ($F(4,38)=0.9$, $p=0.474$). Post-hoc linear regression examining whether right and left SLF_FA predict retention of action values in BP alone was also non-significant ($F(4,38)=0.9$, $p=0.474$).

Figure 17. Group Differences in Right Accumbofrontal Tract Integrity. * p <0.05



SZ and BP: AF Integrity and Value Updating after Positive Feedback

Collapsing across diagnostic group, a linear regression predicting value updating after positive feedback from right and left AF F covarying for age, sex, and scanner was non-significant ($F(4,78)=0.691$, $p=0.600$). Post-hoc linear regression examining whether right and left AF FA predict value updating after positive feedback in SZ was non-significant ($F(4,38)=.296$, $p=0.879$). Post-hoc linear regression examining whether right and left AF FA predict value updating after positive feedback in BP was non-significant ($F(4,35)=.595$, $p=0.669$).

Exploratory Analyses

Medication Effects

Medication history as indexed by chlorpromazine equivalents (CPZ) was not significantly correlated with integrity in right SLF, left SLF, right AF, or left AF. As a result, medication was not controlled for or evaluated as a confound in any additional analyses.

Trait and Symptom Correlations

I conducted non-parametric correlations (Spearman's rho) to evaluate if FA values in the right and left SLF and right and left AF were associated with symptom and trait measures within groups.

In HC, FA values were not correlated with any trait measures of impulsivity, bipolar traits, or anhedonia. In SZ, FA values were not correlated with negative symptoms. In BP, FA values were not correlated with mania severity.

Updating After Negative Feedback

In light of some evidence suggesting that OFC damage(Wheeler & Fellows, 2008) leads to deficits in negative feedback learning specifically, I evaluated the relationship between AF FA and updating after negative feedback. A linear regression predicting value updating after negative feedback from right and left AF FA and group, covarying for age, sex, and scanner was non-significant.

CONCLUSION

Schizophrenia participants did show decreased integrity relative to controls in the right accumbofrontal tract, but contrary to my hypotheses, white matter integrity in the accumbofrontal tract and the superior longitudinal fasciculus did not moderate computational reinforcement learning parameters in schizophrenia and bipolar disorder. When developing this study and its hypotheses, the rationale for examining the superior longitudinal fasciculus in relation to retention was based on literature suggesting that retention may be mediated by working memory performance. However, as detailed in Chapter 3, I found no significant relationship between retention and working memory; other cognitive processes mediated by other circuitry may better account for variability in retention in schizophrenia and bipolar disorder.

Despite findings in clinical research suggesting that connectivity between the NAcc and OFC is associated with reward hypersensitivity, preclinical research suggests discrete functions from different OFC projections that support different reinforcement learning processes. NAcc-OFC circuitry has been implicated in positive symptomatology and over-learning of associations between irrelevant stimuli, rather than in negative symptomatology and use of reward feedback to guide future behaviors (Powers, Mathys, & Corlett, 2017). I initially chose to focus on the accumbofrontal tract based on existing findings in the clinical literature, but the link between the accumbofrontal tract and value updating after positive feedback is more tenuous in light of preclinical evidence. Other brain structures like the hypothalamus and its projections to the ventral tegmental area may be mechanistically more relevant (Nieh et al., 2016), but there are a number of challenges in attempting to image this circuit in the human brain (Billot et al., 2020). Additionally, value updating after positive feedback is not the same as reward hypersensitivity and responsiveness to reward, and may be more cognitively mediated and thus reliant on prefrontal circuits as opposed to NAcc-OFC circuitry.

While planned analyses were based prior literature, OFC circuitry has been of particular scientific interest in recent years with evolving information about its function and role in reinforcement learning emerging rapidly (Stalnaker, Cooch, & Schoenbaum, 2015). Likewise, while working memory has been implicated in reinforcement learning, the types of parameters fit and paradigms used vary considerably which may be why I was not able to replicate findings with this specific form of modeling.

Limitations and Future Directions

The acquisition of diffusion data in the CNP presented some limitations for the planned analyses. CNP data was collected in a single phase-encoding direction. To address

susceptibility-induced distortions from echoplanar imaging, diffusion images were normalized to an anatomical scan without distortions. More advanced methods of addressing susceptibility-induced distortion involve collecting reverse phase-encoding scans or collecting a field map; however, these methods increase overall scan time which can be undesirable when working with clinical populations. Previous research has shown that group analyses can be affected by the polarity of phase-encoding direction, as can FA estimates (Kennis, Van Rooij, Kahn, Geuze, & Leemans, 2016). Due to its location, the accumbofrontal tract is particularly prone to susceptibility artifacts. While scan sequences can be optimized for imaging the accumbofrontal tract, the CNP diffusion scan sequence was not optimized in this way.

Future analyses will focus on whole brain analyses, which would involve entering the computational parameters into a GLM to see which areas across the TBSS skeleton correlate with the parameters. Additionally, as I endeavored to link computational mechanisms with possible neural mechanisms, future directions also include identifying possible functional neural correlates of retention and value updating. The CNP LA5C study included resting state functional magnetic resonance imaging (rs-fMRI), as well as task-based fMRI including task-switching and a scanner-modified version of the BART.

CHAPTER 5

General Discussion

This series of studies compliments a larger body of work investigating decision-making and reinforcement learning deficits in schizophrenia and psychosis and identifying possible mechanistic links to negative symptomatology. They indicate that there is considerable variability within patient groups, and that some aspects of reinforcement learning may be unimpaired in schizophrenia and may even be shared with other psychopathology like bipolar disorder. While individuals with psychosis are seemingly acquiring initial associations between actions and outcomes in a manner similar to healthy controls, there appear to be differences in how they use feedback to update action values and modify future choice behaviors. This difference appears to be moderated by the severity of their negative symptomatology, such that increased negative symptom severity is associated with poor optimization of reward when making decisions under ambiguity, and with greater weighting of negative feedback and lesser weighting of positive feedback when reward contingencies shift. While previous studies have implicated working memory, corticostriatal circuitry, and frontoparietal circuitry in reinforcement learning, I was not able to replicate these findings.

Early versus Chronic Schizophrenia

In using a dimensional approach, I examined both early psychosis in Study 1, and chronic illness in Studies 2 and 3. Though I used different paradigms reliant on reinforcement learning in Study 1 versus Studies 2 and 3, overall findings suggest that across early and chronic phases of illness, schizophrenia is not associated with gross impairments relative to controls and within-group variability in task performance is similarly

accounted for negative symptomatology. There has been limited work explicitly linking the precise nature of reinforcement learning deficits and deficits in reinforcement learning-reliant behaviors across the trajectory of illness, though there has been substantial research specifically within early psychosis and chronic schizophrenia related to other domains of functioning. This work is an early step in identifying how reinforcement learning mechanisms of negative symptomatology manifest similarly or different across phase of illness. Deepening understanding of the role of phase of illness has implications for the timing of interventions and for identifying when in illness trajectory may be optimal for interventions.

Limitations of Dimensional Approaches

In this dissertation, I adopted a largely dimensional approach to psychopathology through use of a psychosis spectrum sample in Study 1, investigation of transdiagnostic reinforcement learning deficits and possible contributions of shared cognitive deficits between schizophrenia and bipolar in Study 2, and exploration of how white matter integrity may be a meaningful predictor of reinforcement learning parameters across diagnostic groups in Study 3. Furthermore, I also sought to identify if lower-level symptoms in healthy controls show similar relationships with reinforcement learning parameters. Dimensional approaches confer many advantages, including that they account for the considerable overlap seen in clinical presentation and in conventional treatments for severe mental illness. This approach enabled me to identify within-group contributors of variability; continued research on moderators of performance and variability in outcome is an essential aspect of the growing field of precision medicine.

However, this dissertation also highlights possible challenges of using a dimensional approach. Namely, while risk-taking and reinforcement learning constructs have tremendous relevance in terms of social and role functioning, defining and describing syndromes by these transdiagnostic mechanisms rather than diagnostic classifications introduces significant challenges in communicating amongst and between researchers, clinicians and patients. It is also tremendously challenging to identify patient profiles based on computational parameters. In other words, at some point, it becomes essential to identify people on some level to facilitate treatment and to group them by whichever primary mechanism drives their symptomatology.

While this dissertation does identify sources of variability within schizophrenia and psychosis that relate to clearly defined reinforcement learning mechanisms and decision-making constructs, it also revealed some shortcomings of using a dimensional approach particularly in chronic samples with less variability in presentation overall. While there were no differences between schizophrenia and bipolar disorder on any metric across any of the three studies presented here, I was also not able to fully evaluate the overlap in distributions between diagnostic groups, or account for any cognitive or neural mediators that play a causal role in the pathophysiology of either diagnosis.

Broader Implications and Future Directions

The work presented in this dissertation has many implications for areas for further study. Future investigations may seek to explicitly probe continuity in deficits across phase of illness, potentially by modeling the same task across phase of illness and examining within-person changes in parameters and associated correlates. Furthermore, it may be particularly elucidating to see if deficits or alterations in reinforcement learning are more reflective of current and active symptomatology, rather than trait-like risk for particular

disorders. For example, future studies may probe mood state (mania, euthymia, depression) to identify state-related changes that further clarify overlap in symptomatology in schizophrenia and bipolar disorders. Lastly, future studies should continue to test and build a unified framework identifying the role of reinforcement learning deficits in the pathophysiology of positive symptoms, negative symptoms, and cognitive deficits seen in psychotic illness.

REFERENCES

Addington, J., Leriger, E., & Addington, D. (2003). Symptom outcome 1 year after admission to an early psychosis program. *The Canadian Journal of Psychiatry, 48*(3), 204-207.

Alloy, L. B., Olino, T., Freed, R. D., & Nusslock, R. (2016). Role of reward sensitivity and processing in major depressive and bipolar spectrum disorders. *Behavior therapy, 47*(5), 600-621.

Andreasen, N. C. (1984). *Scale for the Assessment of Positive Symptons:(SAPS)*: University of Iowa.

Andreasen, N. C. (1989). The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *The British Journal of Psychiatry, 155*(S7), 49-52.

Badner, J., & Gershon, E. (2002). Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Molecular psychiatry, 7*(4), 405-411.

Barber, A. D., Lindquist, M. A., DeRosse, P., & Karlsgodt, K. H. (2018). Dynamic functional connectivity states reflecting psychotic-like experiences. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(5), 443-453.

Barch, D. M., & Dowd, E. C. (2010). Goal representations and motivational drive in schizophrenia: the role of prefrontal–striatal interactions. *Schizophrenia bulletin, 36*(5), 919-934.

Barch, D. M., Treadway, M. T., & Schoen, N. (2014). Effort, anhedonia, and function in schizophrenia: reduced effort allocation predicts amotivation and functional impairment. *Journal of abnormal psychology, 123*(2), 387.

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature neuroscience, 7*(4), 404-410.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1295.

Bilder, R. M., Goldman, R. S., Robinson, D., Reiter, G., Bell, L., Bates, J. A., . . . Woerner, M. G. (2000). Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. *American Journal of Psychiatry, 157*(4), 549-559.

Billot, B., Bocchetta, M., Todd, E., Dalca, A. V., Rohrer, J. D., & Iglesias, J. E. (2020). Automated segmentation of the hypothalamus and associated subunits in brain MRI. *Neuroimage, 223*, 117287.

Boka, E., Del Pozzo, J., Goetz, D., Remsen, B., Walsh-Messinger, J., Getz, M., . . . Malaspina, D. (2020). Ineffective risk-reward learning in schizophrenia. *Psychiatry Research, 293*, 113370.

Brown, E. C., Hack, S. M., Gold, J. M., Carpenter Jr, W. T., Fischer, B. A., Prentice, K. P., & Waltz, J. A. (2015). Integrating frequency and magnitude information in decision-making in schizophrenia: An account of patient performance on the Iowa Gambling Task. *Journal of psychiatric research, 66*, 16-23.

Cannon, T. D., & Keller, M. C. (2006). Endophenotypes in the genetic analyses of mental disorders. *Annu. Rev. Clin. Psychol., 2*, 267-290.

Carter, C., Robertson, L., Nordahl, T., Chaderjian, M., Kraft, L., & O'Shora-Celaya, L. (1996). Spatial working memory deficits and their relationship to negative symptoms in unmedicated schizophrenia patients. *Biological psychiatry, 40*(9), 930-932.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of abnormal psychology, 85*(4), 374.

Chase, H. W., Nusslock, R., Almeida, J. R., Forbes, E. E., LaBarbara, E. J., & Phillips, M. L. (2013). Dissociable patterns of abnormal frontal cortical activation during anticipation of an uncertain reward or loss in bipolar versus major depression. *Bipolar Disorders, 15*(8), 839-854.

Cheng, G. L., Tang, J. C., Li, F. W., Lau, E. Y., & Lee, T. M. (2012). Schizophrenia and risk-taking: impaired reward but preserved punishment processing. *Schizophrenia Research, 136*(1-3), 122-127.

Chung, Y. S., & Barch, D. M. (2016). Frontal-striatum dysfunction during reward processing: Relationships to amotivation in schizophrenia. *Journal of abnormal psychology, 125*(3), 453.

Claes, L., Vertommen, H., & Braspenning, N. (2000). Psychometric properties of the Dickman impulsivity inventory. *Personality and individual differences, 29*(1), 27-35.

Cloutier, M., Aigbogun, M. S., Guerin, A., Nitulescu, R., Ramanakumar, A. V., Kamat, S. A., . . . Henderson, C. (2016). The economic burden of schizophrenia in the United States in 2013. *The Journal of clinical psychiatry, 77*(6), 764-771.

Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *Journal of Neuroscience, 34*(41), 13747-13756.

Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience, 35*(7), 1024-1035.

Conley, R. R., Ascher-Svanum, H., Zhu, B., Faries, D. E., & Kinon, B. J. (2007). The burden of depressive symptoms in the long-term treatment of patients with schizophrenia. *Schizophrenia Research, 90*(1-3), 186-197.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature reviews neuroscience, 13*(9), 636.

Culbreth, A. J., Gold, J. M., Cools, R., & Barch, D. M. (2015). Impaired activation in cognitive control regions predicts reversal learning in schizophrenia. *Schizophrenia bulletin, 42*(2), 484-493.

Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M., & Barch, D. M. (2016). Reduced model-based decision-making in schizophrenia. *Journal of abnormal psychology, 125*(6), 777.

Culbreth, A. J., Westbrook, A., Xu, Z., Barch, D. M., & Waltz, J. A. (2016). Intact ventral striatal prediction error signaling in medicated schizophrenia patients. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(5), 474-483.

Damme, K. S., Young, C. B., & Nusslock, R. (2017). Elevated nucleus accumbens structural connectivity associated with proneness to hypomania: a reward hypersensitivity perspective. *Social cognitive and affective neuroscience, 12*(6), 928-936.

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.

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral science, 19*(1), 1-15.

Deserno, L., Boehme, R., Mathys, C., Katthagen, T., Kaminski, J., Stephan, K. E., . . . Schlagenhauf, F. (2020). Volatility estimates increase choice switching and relate to prefrontal activity in schizophrenia. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(2), 173-183.

Deserno, L., Schlagenhauf, F., & Heinz, A. (2016). Striatal dopamine, reward, and decision making in schizophrenia. *Dialogues in clinical neuroscience, 18*(1), 77.

Deserno, L., Sterzer, P., Wüstenberg, T., Heinz, A., & Schlagenhauf, F. (2012). Reduced prefrontal-parietal effective connectivity and working memory deficits in schizophrenia. *Journal of Neuroscience, 32*(1), 12-20.

Dowd, E. C., Frank, M. J., Collins, A., Gold, J. M., & Barch, D. M. (2016). Probabilistic reinforcement learning in patients with schizophrenia: relationships to anhedonia and avolition. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(5), 460-473.

Dunayevich, E., & Keck Jr, P. E. (2000). Prevalence and description of psychotic features in bipolar mania. *Current psychiatry reports, 2*(4), 286-290.

Eckblad, M., & Chapman, L. J. (1986). Development and validation of a scale for hypomanic personality. *Journal of abnormal psychology, 95*(3), 214.

Ellison-Wright, I., & Bullmore, E. (2010). Anatomy of bipolar disorder and schizophrenia: a meta-analysis. *Schizophrenia Research, 117*(1), 1-12.

Fervaha, G., Foussias, G., Agid, O., & Remington, G. (2014). Impact of primary negative symptoms on functional outcomes in schizophrenia. *European Psychiatry, 29*(7), 449-455.

First, M. B., & Gibbon, M. (2004). The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II).

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological review, 113*(2), 300.

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science, 306*(5703), 1940-1943.

Friston, K. (2023). Computational psychiatry: from synapses to sentience. *Molecular psychiatry, 28*(1), 256-268.

Fusar-Poli, P., & Meyer-Lindenberg, A. (2013). Striatal presynaptic dopamine in schizophrenia, part II: meta-analysis of [18F/11C]-DOPA PET studies. *Schizophrenia bulletin, 39*(1), 33-42.

Fusar-Poli, P., Papanastasiou, E., Stahl, D., Rocchetti, M., Carpenter, W., Shergill, S., & McGuire, P. (2014). Treatments of Negative Symptoms in Schizophrenia: Meta-

Analysis of 168 Randomized Placebo-Controlled Trials. *Schizophrenia bulletin, 41*(4), 892-899. doi:10.1093/schbul/sbu170

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience, 26*(25), 6885-6892.

Glahn, D. C., Kim, J., Cohen, M. S., Poutanen, V.-P., Therman, S., Bava, S., . . . Lönnqvist, J. (2002). Maintenance and manipulation in spatial working memory: dissociations in the prefrontal cortex. *Neuroimage, 17*(1), 201-213.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron, 66*(4), 585-595.

Gold, J. M., Waltz, J. A., Matveeva, T. M., Kasanova, Z., Strauss, G. P., Herbener, E. S., . . . Frank, M. J. (2012). Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Archives of general psychiatry, 69*(2), 129-138.

Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: a deficit in the representation of value. *Schizophrenia bulletin, 34*(5), 835-847.

Good, K. P., Rabinowitz, J., Whitehorn, D., Harvey, P. D., DeSmedt, G., & Kopala, L. C. (2004). The relationship of neuropsychological test performance with the PANSS in antipsychotic naıve, first-episode psychosis patients. *Schizophrenia Research, 68*(1), 11-19.

Groman, S. M., Keistler, C., Keip, A. J., Hammarlund, E., DiLeone, R. J., Pittenger, C., . . . Taylor, J. R. (2019). Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. *Neuron*.

Groman, S. M., Rich, K. M., Smith, N. J., Lee, D., & Taylor, J. R. (2018). Chronic exposure to methamphetamine disrupts reinforcement-based decision making in rats. *Neuropsychopharmacology, 43*(4), 770-780.

Hager, O. M., Kirschner, M., Bischof, M., Hartmann-Riemer, M. N., Kluge, A., Seifritz, E., . . . Kaiser, S. (2015). Reward-dependent modulation of working memory is associated with negative symptoms in schizophrenia. *Schizophrenia Research, 168*(1-2), 238-244.

Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *Neuroimage, 104*, 347-354.

Hegarty, C. E., Jolles, D. D., Mennigen, E., Jalbrzikowski, M., Bearden, C. E., & Karlsgodt, K. H. (2019). Disruptions in white matter maturation and mediation of cognitive development in youths on the psychosis spectrum. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4*(5), 423-433.

Heydebrand, G., Weiser, M., Rabinowitz, J., Hoff, A. L., DeLisi, L. E., & Csernansky, J. G. (2004). Correlates of cognitive deficits in first episode schizophrenia. *Schizophrenia Research, 68*(1), 1-9.

Hindley, G., Bahrami, S., Steen, N. E., O'Connell, K. S., Frei, O., Shadrin, A., . . . Dale, A. M. (2021). Characterising the shared genetic determinants of bipolar disorder, schizophrenia and risk-taking. *Translational Psychiatry, 11*(1), 466.

Ikuta, T., del Arco, A., & Karlsgodt, K. H. (2018). White matter integrity in the fronto-striatal accumbofrontal tract predicts impulsivity. *Brain imaging and behavior, 12*(5), 1524-1528.

Juckel, G., Schlagenhauf, F., Koslowski, M., Wüstenberg, T., Villringer, A., Knutson, B., . . . Heinz, A. (2006). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage, 29*(2), 409-416.

Karlsgodt, K. H., John, M., Ikuta, T., Rigoard, P., Peters, B. D., Derosse, P., . . . Szeszko, P. R. (2015). The accumbofrontal tract: diffusion tensor imaging characterization and developmental change from childhood to adulthood. *Human brain mapping, 36*(12), 4954-4963.

Karlsgodt, K. H., Niendam, T. A., Bearden, C. E., & Cannon, T. D. (2009). White matter integrity and prediction of social and role functioning in subjects at ultra-high risk for psychosis. *Biological psychiatry, 66*(6), 562-569.

Karlsgodt, K. H., van Erp, T. G., Poldrack, R. A., Bearden, C. E., Nuechterlein, K. H., & Cannon, T. D. (2008). Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biological psychiatry, 63*(5), 512-518.

Kennis, M., Van Rooij, S., Kahn, R., Geuze, E., & Leemans, A. (2016). Choosing the polarity of the phase-encoding direction in diffusion MRI: Does it matter for group analysis? *NeuroImage: Clinical, 11*, 539-547.

Keshavan, M. S., Morris, D. W., Sweeney, J. A., Pearlson, G., Thaker, G., Seidman, L. J., . . . Tamminga, C. (2011). A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the Schizo-Bipolar Scale. *Schizophrenia Research, 133*(1-3), 250-254.

Kochunov, P., Hong, L. E., Dennis, E. L., Morey, R. A., Tate, D. F., Wilde, E. A., . . . Favre, P. (2020). ENIGMA-DTI: Translating reproducible white matter deficits into personalized vulnerability metrics in cross-diagnostic psychiatric research. *Human brain mapping*.

Koshiyama, D., Fukunaga, M., Okada, N., Morita, K., Nemoto, K., Usui, K., . . . Kudo, N. (2020). White matter microstructural alterations across four major psychiatric disorders: mega-analysis study in 2937 individuals. *Molecular psychiatry, 25*(4), 883-895.

Koshiyama, D., Fukunaga, M., Okada, N., Morita, K., Nemoto, K., Yamashita, F., . . . Fujimoto, M. (2020). Association between the superior longitudinal fasciculus and perceptual organization and working memory: A diffusion tensor imaging study. *Neuroscience Letters, 738*, 135349.

Lee, D.-K., Lee, H., Park, K., Joh, E., Kim, C.-E., & Ryu, S. (2020). Common gray and white matter abnormalities in schizophrenia and bipolar disorder. *PLoS one, 15*(5), e0232826.

Lee, F. S., Heimer, H., Giedd, J. N., Lein, E. S., Šestan, N., Weinberger, D. R., & Casey, B. (2014). Adolescent mental health—opportunity and obligation. *Science, 346*(6209), 547-549.

Lee, J., & Park, S. (2005). Working memory impairments in schizophrenia: a meta-analysis. *Journal of abnormal psychology, 114*(4), 599.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., . . . Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75.

Loewy, R. L., Pearson, R., Vinogradov, S., Bearden, C. E., & Cannon, T. D. (2011). Psychosis risk screening with the Prodromal Questionnaire—brief version (PQ-B). *Schizophrenia Research, 129*(1), 42-46.

Luk, M., Chang, W., Chong, C., Siu, C., Chan, S., Lee, E., . . . Lo, T. (2019). Altered risky decision making in patients with early non-affective psychosis. *European archives of psychiatry and clinical neuroscience*, 1-9.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature neuroscience, 14*(2), 154.

Mark, W., & Toulopoulou, T. (2016). Psychometric properties of “community assessment of psychic experiences”: review and meta-analyses. *Schizophrenia bulletin, 42*(1), 34-44.

Martin, L. E., & Potts, G. F. (2004). Reward sensitivity in impulsivity. *Neuroreport, 15*(9), 1519-1522.

Millard, S. J., Bearden, C. E., Karlsgodt, K. H., & Sharpe, M. J. (2022). The prediction-error hypothesis of schizophrenia: new data point to circuit-specific changes in dopamine activity. *Neuropsychopharmacology, 47*(3), 628-640.

Morris, R. W., Quail, S., Griffiths, K. R., Green, M. J., & Balleine, B. W. (2015). Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biological psychiatry, 77*(2), 187-195.

Nieh, E. H., Vander Weele, C. M., Matthews, G. A., Presbrey, K. N., Wichmann, R., Leppla, C. A., . . . Tye, K. M. (2016). Inhibitory input from the lateral hypothalamus to the ventral tegmental area disinhibits dopamine neurons and promotes behavioral activation. *Neuron, 90*(6), 1286-1298.

Pantelis, C., Barnes, T., Nelson, H. E., Tanner, S., Weatherley, L., Owen, A. M., & Robbins, T. W. (1997). Frontal-striatal cognitive deficits in patients with chronic schizophrenia. *Brain: a journal of neurology, 120*(10), 1823-1843.

Park, S., & Holzman, P. S. (1992). Schizophrenics show spatial working memory deficits. *Archives of general psychiatry, 49*(12), 975-982.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology, 51*(6), 768-774.

Pearlson, G. D. (2015). Etiologic, phenomenologic, and endophenotypic overlap of schizophrenia and bipolar disorder. *Annual Review of Clinical Psychology, 11*, 251-281.

Pencer, A., Addington, J., & Addington, D. (2005). Outcome of a first episode of psychosis in adolescence: a 2-year follow-up. *Psychiatry Research, 133*(1), 35-43.

Poldrack, R. A., Congdon, E., Triplett, W., Gorgolewski, K., Karlsgodt, K., Mumford, J., . . . Cannon, T. (2016). A phenome-wide examination of neural and cognitive function. *Scientific data, 3*(1), 1-12.

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science, 357*(6351), 596-600.

Radua, J., Schmidt, A., Borgwardt, S., Heinz, A., Schlagenhauf, F., McGuire, P., & Fusar-Poli, P. (2015). Ventral striatal activation during reward processing in psychosis: a neurofunctional meta-analysis. *JAMA psychiatry, 72*(12), 1243-1251.

Reddy, L. F., Horan, W. P., & Green, M. F. (2016). Motivational deficits and negative symptoms in schizophrenia: Concepts and assessments. *Behavioral neuroscience of motivation*, 357-373.

Reddy, L. F., Lee, J., Davis, M. C., Altshuler, L., Glahn, D. C., Miklowitz, D. J., & Green, M. F. (2014). Impulsivity and risk taking in bipolar disorder and schizophrenia. *Neuropsychopharmacology, 39*(2), 456-463.

Rigoard, P., Buffenoir, K., Jaafari, N., Giot, J. P., Houeto, J. L., Mertens, P., . . . Bataille, B. (2011). The accumbofrontal fasciculus in the human brain: a microsurgical anatomical study. *Neurosurgery, 68*(4), 1102-1111.

Sarpal, D. K., Robinson, D. G., Lencz, T., Argyelan, M., Ikuta, T., Karlsgodt, K., . . . Malhotra, A. K. (2015). Antipsychotic treatment and functional connectivity of the striatum in first-episode schizophrenia. *JAMA psychiatry, 72*(1), 5-13.

Satterthwaite, T. D., Kable, J. W., Vandekar, L., Katchmar, N., Bassett, D. S., Baldassano, C. F., . . . Gur, R. C. (2015). Common and dissociable dysfunction of the reward system in bipolar and unipolar depression. *Neuropsychopharmacology, 40*(9), 2258-2268.

Schlagenhauf, F., Huys, Q. J., Deserno, L., Rapp, M. A., Beck, A., Heinze, H.-J., . . . Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage, 89*, 171-180.

Schmack, K., Rothkirch, M., Priller, J., & Sterzer, P. (2017). Enhanced predictive signalling in schizophrenia. *Human brain mapping, 38*(4), 1767-1779.

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron, 91*(6), 1402-1412.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology, 80*(1), 1-27.

Sharpe, M. J., Stalnaker, T., Schuck, N. W., Killcross, S., Schoenbaum, G., & Niv, Y. (2019). An integrated model of action selection: distinct modes of cortical control of striatal decision making. *Annual review of psychology, 70*, 53-76.

Shurman, B., Horan, W. P., & Nuechterlein, K. H. (2005). Schizophrenia patients demonstrate a distinctive pattern of decision-making impairment on the Iowa Gambling Task. *Schizophrenia Research, 72*(2-3), 215-224.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., . . . Matthews, P. M. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage, 31*(4), 1487-1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., . . . Flitney, D. E. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23*, S208-S219.

Spear, L. P. (2013). Adolescent neurodevelopment. *Journal of adolescent health, 52*(2), S7-S13.

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature neuroscience, 18*(5), 620-627.

Strauss, G. P., Frank, M. J., Waltz, J. A., Kasanova, Z., Herbener, E. S., & Gold, J. M. (2011). Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biological psychiatry, 69*(5), 424-431.

Strauss, G. P., Thaler, N. S., Matveeva, T. M., Vogel, S. J., Sutton, G. P., Lee, B. G., & Allen, D. N. (2015). Predicting psychosis across diagnostic boundaries: Behavioral and computational modeling evidence for impaired reinforcement learning in schizophrenia and bipolar disorder with a history of psychosis. *Journal of abnormal psychology, 124*(3), 697.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*: MIT press.

Swainson, R., Rogers, R., Sahakian, B., Summers, B., Polkey, C., & Robbins, T. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia, 38*(5), 596-612.

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS one, 4*(8), e6598.

Valton, V., Romaniuk, L., Steele, J. D., Lawrie, S., & Seriès, P. (2017). Comprehensive review: Computational modelling of schizophrenia. *Neuroscience & Biobehavioral Reviews, 83*, 631-646.

Ventura, J., Nuechterlein, K. H., Subotnik, K., & Gilbert, E. (1995). Symptom dimensions in recentonset schizophrenia: the 24-item expanded BPRS. *Schizophrenia Research, 1*(15), 22.

Wallis, J. D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci., 30*, 31-56.

Waltz, J. A., Frank, M. J., Robinson, B. M., & Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biological psychiatry, 62*(7), 756-764.

Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophrenia Research, 93*(1-3), 296-303.

Waltz, J. A., & Gold, J. M. (2015). Motivational deficits in schizophrenia and the representation of expected value. In *Behavioral neuroscience of motivation* (pp. 375-410): Springer.

Wheeler, E. Z., & Fellows, L. K. (2008). The human ventromedial frontal lobe is critical for learning from negative feedback. *Brain, 131*(5), 1323-1331.

Whitton, A. E., Merchant, J. T., & Lewandowski, K. E. (2020). Dissociable mechanisms underpinning effort-cost decision-making across the psychosis spectrum. *Schizophrenia Research, 224*, 133-140.

Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of general psychiatry, 45*(8), 742-747.

Woods, S. W. (2003). Chlorpromazine equivalent doses for the newer atypical antipsychotics. *The Journal of clinical psychiatry*.

Young, R., Biggs, J., Ziegler, V., & Meyer, D. (2000). Young mania rating scale. *Handbook of psychiatric measures*, 540-542.

Zeng, J., Yan, J., Cao, H., Su, Y., Song, Y., Luo, Y., & Yang, X. (2022). Neural substrates of reward anticipation and outcome in schizophrenia: a meta-analysis of fMRI findings in the monetary incentive delay task. *Translational Psychiatry, 12*(1), 448.

Zeng, J., Yan, J., You, L., Liao, T., Luo, Y., Cheng, B., & Yang, X. (2023). A Meta-Analysis of Neural Correlates of Reward Anticipation in Individuals at Clinical Risk for Schizophrenia. *International Journal of Neuropsychopharmacology, 26*(4), 280-293. doi:10.1093/ijnp/pyad009

Zhang, H., & Yu, T. (2020). Taxonomy of reinforcement learning algorithms. *Deep Reinforcement Learning: Fundamentals, Research and Applications*, 125-133.