ABSTRACT

Title of Dissertation: COMPUTATIONAL CHARACTERIZATION OF INFORMATION PROCESSING IN IMPULSIVITY

Sean Richard Maulhardt, Doctor of Philosophy, 2026

Dissertation directed by: Professor Caroline Charpentier, Department of Psychology

Impulsivity has been theorized as a multidimensional construct, necessitating focused research into determining how these dimensions manifest together and separately. This work aims to uncover how information processing, particularly in learning and decision-making, may be

susceptible to rapid and unplanned responding. While reinforcement learning (RL) has traditionally sat outside the umbrella of impulsivity, sequentially chaining prediction errors could be vulnerable to any assortment of impulse-control valuation and exploratory influence. We propose that impulse-control RL, such as in probabilistic and delayed learning, may have relationships with other impulsivity dimensions, such as temporal discounting (TD) or cognitive control (CC). Participants (N = 210) completed five online cognitive tasks (RL: two-armed probabilistic and delayed bandits, CC: Go/No-Go and flanker, and TD: delay discounting task). The two-armed delayed bandit was designed to experimentally isolate the effect of delay learning and required an explicit estimation of time delay. We simultaneously fit a hierarchical framework with the five drift diffusion models that estimated covariance parameters within and across tasks. Within this framework, a Rescorla-Wagner and hyperbolic discounting model scaled their respective value differences into evidence accumulators for probabilistic bandit and discounting tasks, respectively. We find multiple points of support with some unexpected results for our hypotheses. First, probabilistic and delayed reward learning showed positive relationships in exploratory behavior toward optimal RL policies. Consistency in RL learning behaviors also appeared to relate to TD and CC. Relationships with self-report metrics were predominantly weak, but impulsivity measured by self-report showed a possible influence on stay-rate of the delayed bandit. Taken together, our findings suggest that impulsivity might share related mechanisms of information processing across a battery of tasks.

Keywords: impulsivity, learning, discounting, reward, delay, computational modeling

COMPUTATIONAL CHARACTERIZATION OF INFORMATION PROCESSING IN IMPULSIVITY

by

Sean Richard Maulhardt

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2026

Advisory Committee:
Professor Caroline Charpentier, Chair
Professor Michael Dougherty
Professor Naomi Feldman
Professor Paul Hanges
Professor Matthew Roesch

© Copyright by

Sean Richard Maulhardt

2026

Acknowledgments

I would like to extend my deepest thanks to the faculty at the University of Maryland, especially my advisor Dr. Caroline Charpentier, for her mentorship, consistency, and guidance. I am not sure I would have made it here without her expert support and understanding across such a wide range of techniques and scientific understanding. I am truly blessed to somehow become her first doctoral student. I would also like to thank those on my committee for their assistance and facilitation of various scientific ideas, frameworks, and approaches. Thank you to Michael Dougherty who was a consistent and guiding mentor in brainstorming ideas, having a laugh, and being an incredible leader in the work we do. To Matthew Roesch for his generosity and insights into information quality, which has given me confidence in measuring all sorts of populations. Thanks to Paul Hanges for providing expert understanding of statistical inference and deepening my understanding of correlational approaches in statistics. And to Naomi Feldman, who, given your talents, was an incredible inspiration and guide in helping me better understand advanced modeling approaches.

I would also like to extend additional thanks to Alec Solway for his incredible kindness and initial guidance in computational modeling and Bayesian approaches. Special thanks to Archilline (Archie) Tablada, who has been so understanding, swift, and helpful. Thank you to Yuqing Lee for your persistence, guidance, and mentorship. And finally, a special thanks to all my lab mates, Valentin, Chaebin, Yukta, Xioadi (Sara), and Deanna, who are incredibly talented, understanding, and driven.

I would also like to extend my deepest thanks to my friends and family, who have been immensely important throughout my life. To my parents, Lynn and Marsha – I love you to the world and back, and I am so blessed to have your teaching and wisdom in my life. To my older

siblings, Pamela, Brian, and Terrance: you have been such wonderful role models, and your ease in life's difficulty gave me a gift of laughter. To all my wonderful nephews, I hope I can become a better role model for you, and I am truly blessed to have such wonderful mini-me's who have become way cooler than me. To my friends back home, I miss you immensely and am so blessed we have maintained such a strong relationship, I hope we will be reunited for amazing times ahead. To my friends here, I cannot believe I have had the honor of meeting such a kind and generous group of people, and I am endlessly grateful for your presence in my life.

Dedication

I would like to dedicate my dissertation to my late grandparents, Ruth, Richard, Rose, and Michael, who gave me the gift of learning.

Table of Contents

Acknowledgments....ii

Dedication....iv

Table of Contents....v

List of Tables....viii

List of Figures....ix

List of Abbreviations....x

Introduction....1

Chapter 1: Behavioral Impulsivity....2

Choice Impulsivity....3

Action Impulsivity....6

Learning Impulsivity....9

Aim 1: Information Processing in Behavioral Impulsivity....12

Methods....15

Participants....15

Materials....15

Analysis....20

OLS and Mixed-Effects Regressions....20

Individual Task Models....21

Hierarchical Bayesian Model (HBM)....24

Results....26
Inhibitory Control (Go/No-Go)....28
Attentional Control (Flanker)....31
Temporal Discounting (Delay Discounting Task)....33
Probabilistic Learning (Probabilistic Bandit)....36
Delayed Learning (Delayed Bandit)....39
Task Correlations (Shared and Unique Variance in Impulsivity Dimensions)....42
Chapter 1 Summary....47
Chapter 2: Personality Impulsivity....48
Impulsivity Questionnaires....49
Self-Report and Behavioral Impulsivity....52
Clinical Impulsivity....55
Aim 2: Information Processing in Personality Impulsivity....56
Methods....56
Participants....56
Materials and Procedure....57
Analysis....58
Results....60
Confirmatory Factor Analysis....60
Bivariate Correlations....62
Exploratory Analysis of BIS-11 and Delayed Bandit....63
HBM and BIS-11 Total Factor....66

Chapter 2 Summary 66

General Discussion 67

Appendix A. Supplemental Information 76

Appendix B. Supplemental Plots 79

References 83

List of Tables

Table 1. Choice behavior and RT for each experiment by various conditions
Table 2. Go/No-Go hierarchical drift diffusion untransformed parameters
Table 3. Flanker hierarchical drift diffusion natural-scale parameters
Table 4. Delay Discounting hierarchical hyperbolic drift diffusion natural-scale parameters
Table 5. Probabilistic Bandit hierarchical Rescorla-Wagner drift diffusion natural-scale parameters
Table 6. Delayed bandit hierarchical drift diffusion natural-scale parameters
Table 7. Model fit indices before and after factor corrections
Table 8. Fixed interaction effects from individual multilevel logistic regressions

List of Figures

Figure 1. Impulsivity Task Battery Diagrams
Figure 2. Go/No-Go Performance
Figure 3. Flanker Performance
Figure 4. Delay Discounting Performance
Figure 5. Probabilistic Bandit Learning
Figure 6. Delayed Bandit Learning
Figure 7. Correlations between behavioral measures
Figure 8. Correlations between self-report factors
Figure 9. Correlations between behavioral and self-report measurements
Figure 10. Delayed bandit stay behavior as a function of BIS-11 factor scores interacting with unrewarded trials
Figure 11. BIS-11 total score effects on HBM parameters

List of Abbreviations

ADHD = Attention-Deficit/Hyperactivity Disorder
AUC = Area Under the Curve
BIS-11 = Barratt Impulsiveness Scale, Version 11
BSSS = Brief Sensation Seeking Scale
CC = Cognitive Control
CFA = Confirmatory Factor Analysis
CI = Confidence Interval
CrI = Credible Interval
DASS-21 = Depression, Anxiety, and Stress Scale, Short Form
DB = Delayed Bandit
DDM = Drift Diffusion Model
DSM = Diagnostic and Statistical Manual of Mental Disorders
GMR = Geometric Mean Ratio
HBM = Hierarchical Bayesian Model
IUS-12 = Intolerance of Uncertainty Scale, Short Form
MCQ-SF = Marijuana Craving Questionnaire, Short Form
NDT = Nondecision Time
OCD = Obsessive-Compulsive Disorder
OCDS = Obsessive Compulsive Drinking Scale
OLS = Ordinary Least Squares
OR = Odds Ratio
PB = Probabilistic Bandit
RL = Reinforcement Learning
RT = Reaction Time
SDT = Signal Detection Theory
SSM = Sequential Sampling Model
TD = Temporal Discounting
UPPS-P = (Negative) Urgency, Premeditation, Perseverance, Sensation Seeking, Positive Urgency

Introduction

First published in 1952, the Diagnostic and Statistical Manual of Mental Disorders (DSM) began advancing the classification of psychopathology (American Psychiatric Association, 2022). However, the advent of the DSM has not been without shortcomings, where classification has suffered from comorbidity between diagnoses and the heterogeneity within (Carragher et al., 2015; Cuthbert & Insel, 2013; Insel et al., 2010; Lilienfeld, 2014). Given these various issues, new initiatives have aimed at advancing classification and proclaiming solutions to these shortcomings (Insel et al., 2010; Kotov et al., 2017; Montague et al., 2012). These new paradigms further advocate for a holistic approach, spanning an array of disciplines, intended for better understanding, treatment, and prediction. For example, newer approaches emphasize dimensional ratings of symptom severity rather than binary judgments, providing clinicians with more nuanced assessment tools (Krueger et al., 2007). Complementing these efforts, new statistical frameworks and psychological paradigms have begun to shift towards the identification of transdiagnostic markers, such as temporal discounting (TD) (Bickel et al., 2012), cognitive control (CC) (McTeague et al., 2016), or reinforcement learning (RL) (Gillan et al., 2016). While TD has been considered the choice dimension of impulsivity, CC is the broad process of regulating thoughts and actions, commonly defined as the action dimension of impulsivity. Given these impulsivity dimensions, the present work proposes and further defines the overlap between impulsivity and RL.

History has left impulsivity with a broad boundary in measurement, and the present work seeks greater clarification. Multiple forms of measurement have been used to characterize an individual's degree of impulsivity, from behavioral tasks to self-report. While TD and CC fall under an umbrella of behavioral impulsivity, the self-report measures are constructed to capture

latent factors of personality. This work estimates behavioral and personality impulsivity measures and clarifies their associations with RL, which does not traditionally fall under the umbrella of impulsivity (Dalley & Robbins, 2017) but remains theoretically important (Poulton & Hester, 2020). The history of defining impulsivity is longstanding: William James (1890) regarded the very nature of consciousness as impulsive. Thoughts, feelings, and associations often burst into conscious awareness from internal ideas or external stimuli (Moors & De Houwer, 2006). Examining information processing in RL's relationship to impulsivity dimensions may reveal shared latent mechanisms. To tease these latent mechanisms apart, a computational framework is constructed to answer, "How do impulse-control vulnerabilities in behavior and personality manifest when learning in uncertain and delayed environments?"

This work broadly aims to computationally characterize shared and unique dimensions involving information processing, impulsivity, and impulse-related psychopathology. The methodology employs various computational models to characterize the effect of TD in a hyperbolic model, the accumulation of CC evidence in drift diffusion modeling (DDM), and credit assignment in RL. The following gaps are of focus: (1) the relationship between probabilistic and delayed learning in RL, (2) relationships between RL, TD, and CC, and (3) relationships across impulsivity behavioral tasks and self-report metrics. Specifically, this research measures behavioral impulsivity through the lens of computational modeling (Chapter 1) and subsequently attempts to tie these metrics to self-report questionnaires (Chapter 2) which revolve around impulsivity-related symptomatology.

Chapter 1: Behavioral Impulsivity

Impulsivity, a cardinal feature of personality and mental health, is thought to consist of both trait and state-like features characterized by rapid and/or unplanned responses to internal

and external stimuli regardless of negative consequences (*International Society for Research on Impulsivity*, 2019; Strickland & Johnson, 2021). However, the definition of impulsivity is rarely agreed upon and the construct appears to have a variety of relationships to behavior, neural circuitry, and cognitive function (Evenden, 1999; Moeller et al., 2001). For example, consider a person wanting to buy a coffee. One could acquire this beverage at home, premade, or with an artisan brew. If one were to plan for the grocery store and make one's coffee at home, one could trade time for value by avoiding an additional premium cost. Other times one could trade value for time, where a premade coffee at a convenience store would provide an immediate beverage but at a higher cost. In another scenario one might favor exploration of options, such as the uncertainty and novelty of an artisan brew, and over time, begin to value this alternative. The dilemma of trading value for time occurs in everyday behavior and is susceptible to multiple abnormalities in planning, inhibition, or learning. In some cases, one plans to purchase coffee beans and prepares coffee at home. In other cases, one struggles to inhibit one's immediate purchase and consumption from a convenience store. Or perhaps an uncertain and novel artisan recipe heightens one's valuation of the coffee. Rather than focusing on emergent behavior (consuming coffee), individuals' behaviors may reflect combinations of long-term planning (choice), short-term inhibition (action), and over-time adaptation (learning).

Choice Impulsivity

The famous marshmallow task (Mischel et al., 1972; Mischel & Ebbesen, 1970) sought to understand whether children would forgo one marshmallow in favor of two marshmallows if they waited 15 minutes. The results showed a large degree of variance in children's observed strategies, where they attempted to direct their attention away from immediate reward through cognitive or overt changes in behavior (Mischel, 1974; Mischel & Ebbesen, 1970). Walter

Mischel used the term ‘delay of gratification’ to describe an individual’s deferral of immediate reward for larger-later rewards (Mischel, 1974; Mischel et al., 1972; Mischel & Ebbesen, 1970). Despite weak correlations with self-reported measures of self-control (Mischel, 2013), performance at age four predicted future cognitive and social competencies in adolescence (Mischel et al., 1989). Impressively, task performance in preschoolers predicted greater performance on an emotion-based Go/No-Go and predicted associations with neural circuitry for motivational processes 40 years later (Casey et al., 2011). However, this paradigm produced biased measurement estimates as a function of censoring waiting time and the shared variance of achievement-adjacent covariates (Falk et al., 2020).

New experimental paradigms of impulsive choice, using an economic normative framework (Frederick & Loewenstein, 2002), have attempted to address certain methodological limitations, including one-shot, censored times, and collinear covariates, to better describe participant choice behavior. To characterize individual behavior in the deferment of reward, intertemporal choice paradigms are constructed with repeated choice dilemmas between smaller-sooner and larger-later rewards (Hamilton et al., 2015; Kirby, 1997). One benefit of repeated intertemporal choices is that they identify an individual’s delay of gratification, reveal participant trade-offs, and capture patterns inconsistent with economic models, such as preference reversals or choice inconsistencies (Green & Myerson, 2004; Kable & Glimcher, 2010). Intertemporal choice paradigms have been implemented with hypothetical scenarios (Baumann & Odum, 2012; Meyers & Koehler, 2021), real monetary compensation (Madden et al., 2003), experiential feedback (Horan et al., 2017; Reynolds & Schiffbauer, 2004), and waiting tasks (Kvam et al., 2021; Xu et al., 2020). Despite the large variety in task methods, intertemporal choice tasks are often collapsed into latent measures and used for predicting real-world behavior (Huang et al.,

2024; Sharma et al., 2014), heightening the need for more ecological and person-centric psychological tasks (Bailey et al., 2021; Botvinick et al., 2015). While ecological validity in task methods is crucial, one should consider the model often constructed alongside intertemporal choice tasks.

One common metric of impulsivity is temporal discounting (TD) or the estimation of an individual's rate of devalues the larger-later option (commonly denoted *k*). Specifically, a low discounting rate leads to higher valuation for deferment of reward. The mechanisms underlying discounting rates are often interpreted as the tradeoff of value and time to plan for future reward (Redish et al., 2008; Shenhav et al., 2017; Solway et al., 2017). To estimate discounting, one can integrate the area under the curve (Myerson et al., 2001), utilize models and optimization (McKerchar et al., 2009), or employ Monte Carlo methods (Vincent, 2016). These early methods of approximation found closed-form solutions to the hyperbolic model, but later methods leverage open-form solutions due to innovations in computing power. These advancements allow greater precision and flexibility in modifying the original hyperbolic model.

These TD parameter estimates have revealed steeper discounting in various mental health outcomes including attention-deficit/hyperactivity disorder (ADHD) (Bitsakou et al., 2009; Jackson & MacKillop, 2016), substance use disorders (Bickel et al., 2012; Romeu et al., 2020), and depressive disorders (Amlung et al., 2019). Furthermore, mental disorders centered around restriction, like those of anorexia nervosa, show shallower discounting rates unlike other eating disorders such as binge-eating (Amlung et al., 2019; Maxwell et al., 2020). The TD parameter appears to capture individual differences in temporal processing of future reward (Franken et al., 2008; Redish et al., 2008) and can serve as a transdiagnostic marker of mental health outcomes (Bickel et al., 2014; Keidel et al., 2024; Lempert et al., 2019). In summary, the study of choice

impulsivity has shown a trajectory of advancements across economic and psychological disciplines but requires further systematic investigation to be a robust transdiagnostic marker and fully establish utility to mental health outcomes (Keidel et al., 2024; Levitt et al., 2023).

Action Impulsivity

Choice behavior constitutes more than just preference or the planned balance between trading value and time; deficiencies may arise in the performance of an individual distinct from that of their preference (Broos et al., 2012; MacKillop et al., 2016). This 'action' dimension of impulsivity has been commonly measured through a participant's habitual or prepotent response that they successfully inhibit (Broos et al., 2012; MacKillop et al., 2016; Strickland & Johnson, 2021). The mechanisms involving the successful inhibition have been thought of as a dual-system of cognitive control (CC) (Bargh & Chartrand, 1999; Kahneman, 2011; Metcalfe & Mischel, 1999) or in top-down conflict monitoring (Botvinick et al., 2001). These dual-system paradigms have spanned behavioral economics (Tversky & Kahneman, 1974), reinforcement learning (Daw et al., 2011; Gläscher et al., 2010), and social cognition (Bargh & Chartrand, 1999). Furthermore, dual-system paradigms typically entail an 'automatic' system and a 'controlled' system which interact to guide behavior (Bargh & Chartrand, 1999; Moors & De Houwer, 2006). These systems interact in competitive (Daw et al., 2006, 2011), cooperative (Gershman et al., 2014; Moran et al., 2019), or mixed (Gershman & Daw, 2017) manners that challenge the features traditionally ascribed to each system (e.g. efficiency, flexibility, consciousness), which remain debated (Evans & Stanovich, 2013; Melnikoff & Bargh, 2018).

Competitive task environments are often used to measure some form of CC, like inhibitory control in the Go/No-Go (Luria, 1973) or attentional control in the flanker task (Eriksen & Eriksen, 1974). To measure inhibitory control, the Go/No-Go task was developed

through interference in visuomotor processing (Gomez et al., 2007; Ratcliff et al., 2018) where participants would respond to 'Go' cues and withhold a response to infrequent 'No-Go' cues. To measure attentional control, the flanker task (Eriksen & Eriksen, 1974) introduced interference in visual processing by presenting participants with a central target in congruent ('>>>') and incongruent ('><>') surrounding flankers. Similar to delay discounting, inhibitory control has been shown to have transdiagnostic features as it links to decreased evidence sampling with ADHD (Huang-Pollock et al., 2017; Weigard et al., 2016), increased error-rates in substance use (Ames et al., 2014; Finn et al., 1999), and varying effects of inhibition in externalizing behaviors (Hall et al., 2021; Hoyniak & Petersen, 2019; Letkiewicz et al., 2024). Likewise, the flanker task has also shown mixed relationships to mental health outcomes, such as increased reaction time (RT) to negative faces in anxiety (Chen et al., 2016), increased response accuracy and patience in depression (Dillon et al., 2015), and decreased evidence threshold to respond when associated with higher trait antagonism (Hall et al., 2021).

However, given that impulsive behavior can look naturally random (noise), estimated individual differences may only partially reflect trait-level differences. Across different task designs variance might arise from inattention to instructions (Feher Da Silva & Hare, 2020), manner of task presentation (C. L. Smith & Hantula, 2008), or measure-related noise (Ahn et al., 2020) rather than representing true underlying trait impulsivity. This is particularly relevant for interference tasks (Stroop task, Go/No-Go, and Flanker) as they appear to share accounted variance (Hall et al., 2021), further cautioning interpretation of impulsivity-related individual differences (MacKillop et al., 2016). One such solution has been to use a process model that accounts for Gaussian noise in both choice and RT through a diffusion process (Gomez et al.,

2007; Ratcliff, 1978), thereby separating latent decision processes from routine behavioral variability.

Interference tasks usually occur in quick decisions (under two seconds) where researchers have advanced both theoretical understanding and measurement. These decisions are thought to occur in a probabilistic fashion, where adjusting an individual's criterion shifts their response in relation to the underlying environment's signal-to-noise ratio. Signal Detection Theory (SDT) explains the trade-off between two-alternative forced choices when one choice (left or right) is associated with the underlying truth (correct or incorrect) in a probabilistic judgment (Green & Swets, 1966). Typically, SDT uses two criteria (or multiple if you expand the choice space) along a joint Gaussian distribution where a participant responds when evidence exceeds that criterion ("Go" or "Left" vs. "No-Go" or "Right"). When distributions overlap, discrimination becomes more difficult, and an agent may adopt a more conservative criterion when presented with increased noise. However, SDT has only been able to quantify participants' choices rather than illuminate the decision processes of evidence gathering. One way to capture the evidence in the decision process is to consider both RT and choice in a speed-accuracy trade-off (Ratcliff, 1978). While these assumptions regarding evidence gathering are crucial, the incorporation of RT through sequential sampling models (SSM) can allow us to make inferences about a participant's decision mechanisms.

SSMs have become a staple model for speed-accuracy measures of two-alternative forced choice tasks. For a review of the different implementations of SSM, see Ratcliff et al. (2016). One such SSM, the drift diffusion model (DDM), uses a competing evidence accumulator to capture the processes involved in accuracy and RT across continuous time. The speed-accuracy tradeoff entails the problem of error rates increasing as decision time decreases. DDM has

mostly been used for motion discrimination for relatively quick decisions, rather than decisions over long time intervals. Despite this, DDM has been used in a variety of experiments to measure various choice constructs, such as memory (Ratcliff, 1978), perception (Solway, Lin, & Vinaik, 2021), learning (Peters & D'Esposito, 2020), confidence (Ratcliff & Starns, 2013), or planning (Lei & Solway, 2022) as the choice rule (or activation function). Due to the stochastic nature of sensory information, the accumulation of information is assumed to drift towards one of the bounds with Gaussian noise across the passage of time. Upon reaching the boundary, a decision is executed. Although various expansions of the model exist (Boehm et al., 2018), the base model has four necessary parameters. The drift rate corresponds to the velocity of information encoding or slope of the random walk, reflecting evidence accumulation. The decision threshold is the cutoff when a decision is made and illustrates the amount of evidence needed to declare a decision. The starting point is a response bias towards one of the two boundaries, representing a relative asymmetry of response evidence needed for either boundary. Finally, the nondecision time represents the processing components that are unrelated to evidence accumulation, such as in sensory detection. These models and their parameters have become highly relevant to clinical group differences, such as a fairly stable transdiagnostic reduction of drift rate across behavioral tasks (Letkiewicz et al., 2024; Sripada & Weigard, 2021), with more inconsistencies in boundary separation (Hall et al., 2021; Huang-Pollock et al., 2017; Solway, Lin, & Vinaik, 2021).

Learning Impulsivity

Learning can also be thought of as two distinct controllers of behavior, which are often mapped onto habit-based (automatic) and goal-directed (controlled) systems. The habit-based system is most clearly demonstrated in Thorndike's (1911) "law of effect", which illustrates how animals increase approach behavior when reinforced and decrease when punished. These

behaviors require experience to create associative learning (Pavlov, 1927), but once established, they are quick, efficient, and effortless to deploy. In contrast, the goal-directed system can be traced to Tolman (1948) theory of cognitive maps. The cognitive map constructs a representation of the environment through latent learning and appears to compete, coordinate, and interact with the habit system (Gershman et al., 2014; Lei & Solway, 2022; Moran et al., 2019). This system is more controlled, deliberate, and sensitive to changes in the environment (commonly referred to as state-transitions). Both systems help an organism navigate a complex world by performing routine behaviors and overriding such routine when dynamic shifts in the environment are detected.

Reinforcement Learning (RL) is a modern mathematical framework built upon animal models of association learning that utilizes recursion to explain learning through trial-and-error behavior (Sutton & Barto, 2018). Recursive learning is dependent on the foundational Rescorla-Wagner model, which characterizes how expectation errors of reinforcement drive learning in simple conditioning (Rescorla & Wagner, 1972). This recursive framework has led researchers to consider how reward prediction errors might play a role in the interplay between brain and behavior (Niv, 2009). Substantiating this relationship, reward prediction errors have been largely tied to phasic changes in dopaminergic neurons (Kobayashi & Schultz, 2008; Schultz et al., 1997) and have led to the theory of dopamine as a reward prediction error. This would imply that one of the functions of dopamine is to help us learn and gather rewards in a variety of complex environments. Researchers usually probe learning in environments using (reward or stimulus) reversals or devaluation to assess how ‘sticky’ a behavior is. Specifically, the prefrontal cortex encodes relevant and irrelevant environment features (Asaad et al., 2017) and helps to overcome reversals and devaluations (Frank & Claus, 2006; Wilson et al., 2014). This has led researchers

to consider a goal-and-habit system through careful crafting of an experimental task (two-step task) that maps with two classes of RL models (Daw et al., 2011). Whereas one class, known as model-free RL, uses the chaining of reinforcement through the recency of past sequences, model-based RL uses state prediction errors to approximate an environment's state-transitions (Gläscher et al., 2010). While this framework allows an approximation to how an agent might transition through complex and competitive environments, new studies have attempted to integrate multiple cognitive mechanisms in unique ways.

To link impulsivity metrics, such as TD and CC, and RL, one must consider the connection between models and experimental designs. One of the most common approaches is to measure different tasks and their associated models and identify if there is shared variance in theoretically similar parameters (Kvam et al., 2021; Solway et al., 2017). Often these different tasks and models lead to poor associations, such as in planning (Solway et al., 2017) and discounting (Kvam et al., 2021; Romeu et al., 2020). However, independent task batteries have found shared structures in learning and other impulsivity metrics (Sharma et al., 2014) and a relationship between RL directed exploration and TD (Sadeghiyeh et al., 2020). Another way is to use a unified model in an experimental design that connects differing mechanisms, such as in a CC and RL framework (Burghoorn et al., 2025; Guitart-Masip et al., 2012), or in a working memory and RL frame (Collins et al., 2014). Lastly, one can fit participant-level models and cluster participants by their dominant model-fit, wherein non-learning clusters showed stronger associations with impulsivity compared to other clusters (Cockburn et al., 2024).

While methods for connecting behavioral metrics are varied, options for substantiating the relationship between the DSM, impulsivity, and learning remain limited. Researchers can examine descriptive differences between clinical vs control groups (Kanen et al., 2019; Marzuki

et al., 2021), or towards self-reported measures of psychopathology (Dubois & Hauser, 2022; Gillan et al., 2016; M. S. Lim et al., 2015; Skvortsova & Hauser, 2022; R. Smith et al., 2022). Although many of these aspects of impulsivity are theoretically relevant (Dalley & Robbins, 2017), these parameters represent distinct mechanisms of learning and often show mixed results across the spectrum of clinical disorders. RL-related impulsivity might manifest in several ways: abnormal or noisy updating of stimulus values (M. S. Lim et al., 2015; Skvortsova & Hauser, 2022), excessive exploration regardless of value function (Cockburn et al., 2024; Dubois & Hauser, 2022; Gillan et al., 2016; Lorains et al., 2014), or asymmetric processing of gains and losses (Franken et al., 2008; Lorains et al., 2014; Zou et al., 2022). In contrast, impulsivity may also be poorly characterized by RL models (Cockburn et al., 2024; Zou et al., 2022). Each of these potential learning abnormalities may reflect temporal processing dysfunction when sampling and integrating new information, and advancements in computational modeling have made these latent mechanisms increasingly accessible to measurement. The present study addresses some of these inconsistencies while simultaneously measuring latent mechanisms of three possible dimensions of impulsivity (TD, CC, and RL) within a single sample.

Aim 1: Information Processing in Behavioral Impulsivity

Given the dimensional characterization of impulsivity, the current study would simultaneously investigate information processing mechanisms underlying choice impulsivity (TD), action impulsivity (CC), and reward learning (RL). To achieve this, we implemented computational modeling in a unified hierarchical Bayesian framework to capture both unique and shared relationships of these processes (detailed below). To unify these separate dimensions, we modeled each independently and examined their relationship in a joint covariance matrix between model parameters. The same participants (N=210) completed a battery of five tasks,

including well-established Go/No-Go and flanker tasks (CC), delayed discounting (TD), probabilistic two-armed bandit (RL), as well as a novel two-armed delayed task designed to assess delay learning. In this task, participants must learn the delay with which each arm delivers a reward, thereby exacerbating a problem of credit assignment (Minsky, 1961). Such a task may reveal impulsive individual differences, such as the rate of switching between arms, which increases the complexity of credit assignment (Gershman, 2020; Pike-Burke et al., 2018), as well as the tendency to accumulate rewards delivered at shorter delays rather than prioritizing delay learning.

We preregistered specific hypotheses (https://osf.io/3tq67/) focusing on cross-task parameter correlations. These hypotheses generally assume that higher behavior impulsivity (greater control errors or preference for immediacy) would be associated with reduced ability in learning of reward contingencies. During analysis, we refined these hypotheses for two reasons. First, we decided to incorporate DDM across all tasks to jointly model choice and reaction time (RT). Second, we restricted delayed bandit analyses to the pseudo-optimal stay strategy (instead of following reward strategy), as no reward effects on behavior were observed. Consequently, reward effect correlations with discounting and probabilistic bandit were not pursued. Additionally, the pseudo-optimal stay strategy of the delayed bandit was assumed, but still needed to be validated as a hypothesis (Pike-Burke et al., 2017, 2018). Our specific hypotheses preserve the direction and relational structure of the preregistered predictions; only parameter operationalizations were modified (Note: this was partially expected and described in preregistration, "Parameters are subject to change based on…").

The first preregistered hypothesis to the relationship between RL and TD with a predicted inverted relationship between bandits' drift rates and discounting rate (**H1**). This parametric

relationship reflects that those with higher degrees of impulsivity, or preference for immediacy, would associate with a reduction in ability to learn reward contingencies. The second hypothesis deals with correlations between probabilistic RL and CC, predicting positive relationships between probabilistic learning and CC drift rates (**H2**). The parametric relationship assumes that impulsivity leads to worse cognitive control and would correlate with a reduction in ability to learn probabilistic reward — either by updating recency of reward more quickly or excessive exploration despite learned values. The third hypothesis deals with the relationship between learning probability and delay information with a predicted positive relationship between bandits' drift rates and predicted negative relationship between probabilistic learning rate and delayed drift rate (**H3**). This parametric relationship predicts that participants who better at learning, whether stay-on-choice, quicker updating, or greater evidence accumulation, would be shared across contingency environments. To validate the delayed bandit, we tested whether participants who repeat their previous choice (vs. switched) would lead to overall better delay estimation and better delay estimation for the shorter-delayed arm (**H1.V**).

There were also some intuitive exploratory predictions where similar behavioral and task-modalities will correlate, such as Go/No-Go and flanker (CC), and probabilistic and delayed bandits (RL) (**E1**). We also predicted correlations between the same computational parameters across tasks, such as drift rates or inverse temperatures, boundary separations, and starting-point biases (**E2**). Finally, classical manipulations of behavioral tasks will replicate known effects of lower accuracy and higher RT across Go/No-Go (No-Go and reversal), flanker (incongruent), and probabilistic bandit (reversal) (**E3**). Specifically, each manipulation condition (No-Go, incongruent, and reversal) should produce lower accuracy and higher RT across the three

experiments. While these hypotheses will have specific effects, the analysis will be comprehensive and remind the reader of the preregistered and exploratory hypotheses.

Methods

Participants

A total of 222 participants were recruited from Prolific Academic (https://prolific.com) for a median time of 70 minutes (IQR = 26.28). Prolific inclusion criteria included: fluency in English, ages over 18, and no color blindness. All participants were paid $13.35 for 80 minutes, along with a bonus based on accuracy and RT (M_{bonus} = \$1.81, SD_{bonus} = \$0.14). Among the 222 participants, four were removed and rejected for early termination, seven had repeated the study, but only three were removed on an individual basis of having seen more than half of the experiment in the first session. Twelve participants showed noisy response patterns in at least one of the experiments (half of trials under .2 seconds and below 75% accuracy), and five of these 12 were removed for violating this contingency in two or more experiments. We removed trials with RTs in a certain limit, such that anything below 200ms or above 4 seconds were removed from analysis (5.22% of trials across all experiments) with Go/No-Go (1.96%), Flanker (4.12%), Probabilistic Bandit (7.17%), Delay Discounting (3.83%), and Delayed Bandit (8.69%). Removing certain RTs helps to stabilize model fitting from aberrant response, and we carried this throughout the rest of the analyses. The resulting total sample (N = 210) contained 141 males and 69 females with a moderate age range (M_{age} = 32.49, SD_{age} = 10.49, Range = [18, 72]). This study was approved by the University of Maryland College Park IRB (ID: 2129967-2).

Materials

The impulsivity task was built with PsychoPy3 and the self-report surveys were administered through Qualtrics (Pierce et al., 2022; Qualtrics, 2005). Participants were first

asked to sign a consent form via Qualtrics, then redirected to the main experiment hosted on Pavlovia, and finally redirected back to Qualtrics for the self-report surveys. Upon completion, participants were given a code to be used on the Prolific database to confirm their completion and receive payment.

Impulsivity Task Battery

Participants were given overarching instructions on the tasks that they would complete alongside individual instructions at the beginning of each task. After the instructions, each task had a demo period with a small number of practice trials. A total point counter was added to the bottom right to incentivize bonus payments. The task order was randomized across participants. The most common sequence was experienced by five participants, but most sequences were unique or shared by only two or three participants, illustrating effective randomization across the sample. Any task with two decisions used the 'Q' letter for left choices and 'P' letter for right choices, whereas 'space' was used for the Go/No-Go task. Despite individually randomized sequences, we present the five tasks as follows (Figure 1).

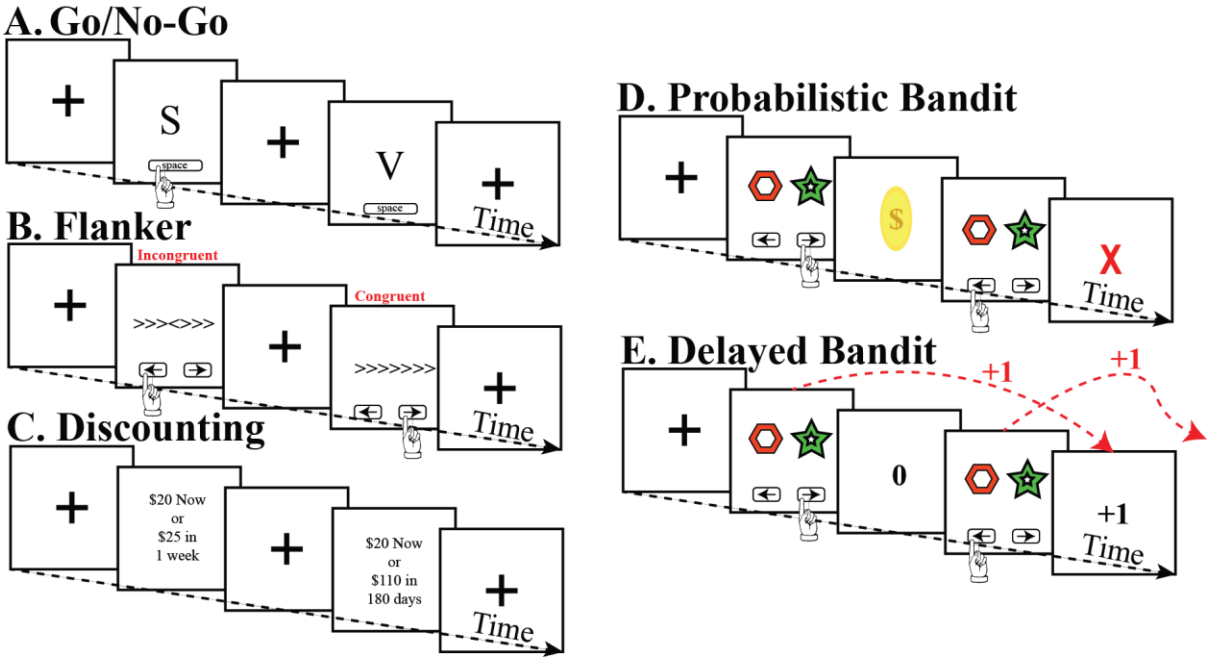


Figure 1. Impulsivity Task Battery Diagrams. (A). Go/No-Go Task where participants press 'space' on Go trials ('S') and without space on No Go ('V') trials. **(B).** Flanker task has participants select the correct direction of the target for incongruent ('<') and congruent ('>'). **(C).** Delay Discounting Task presents two-alternative forced choice for an immediate ('$20 Now') or delayed ('$25 in 1 week') option with varying magnitude and delay ('$110 in 180 days'). **(D).** Probabilistic Bandit Task with two-alternative forced choices with two objects (red hexagon and green star) containing differing probabilities (70%/30%) of receiving a gold coin or red 'X'. **(E).** Delayed Bandit Task with two-alternative forced choices with two objects (red hexagon and green star) containing fixed rewards (+1) that are carried forward into future trials by differing amounts (green star carries reward 1 trial into the future), rewards both are aggregated (sum on trial) and anonymous (no cue to differentiate which object to assign credit).

Go/No-Go Task (Figure 1A)

Participants were asked to perform a letter-based Go/No-Go task (MacKillop et al., 2016) with a single reversal (Finn et al., 1999). After completion of the practice session, the task was separated into two sessions with 58 trials per session, wherein each session featured a different target letter. Each trial showed the letter for 200 milliseconds (ms) and then disappeared; participants could answer anytime but had a maximum of 700 ms in total (Šašinka et al., 2023). At the end of the trial a green or red 'X' was displayed to indicate a successful or unsuccessful trial, respectively. Each session consisted of four blocks presented in randomized order with a new target letter per block. Six distinct letters (i.e, 'S', 'A', 'O', 'V', 'T', 'N') were selected for their visual distinctiveness and neutral connotations, with 'V' serving as the target in the first session and 'N' in the second. For each session, four blocks contained structured Go/No-Go sequences, each with 14 trials (10 Go, 4 No-Go) systematically varying the number of consecutive Go trials before a No-Go trial. Each of those four blocks incorporated all possible sequences: single, double, triple, and quadruple Go trials before No-Go stimuli. We added two remaining single No-Go trials randomly between blocks. The design ensured that while two No-Go trials could occur consecutively (when a No-Go at the end of one block was followed by a No-Go at the

beginning of another), no more than two consecutive No-Go trials would ever appear. Within each block, the trial sequence remained fixed.

Flanker Task (Figure 1B)

Participants were instructed to perform a flanker task (Banca et al., 2015; Eriksen & Eriksen, 1974). After being given several practice trials, participants completed the task over two sessions of 59 trials per session. On each trial, a set of seven keyboard arrows (e.g., '>>><>>>') were displayed with the middle arrow being the target (<). These stimuli were displayed for 200 ms and then disappeared; participants were allowed to answer anytime within a maximum time of 700 ms in total. At the end of the trial a green or red 'X' was displayed to indicate a successful or unsuccessful trial, respectively. Trials included three conditions: Congruent (e.g., '<<<<<<<'), where flanking arrows matched the target; Incongruent (e.g., '>>><>>>'), where flanking arrows opposed the target; and Neutral (e.g., "---<---"), where flankers were neutral symbols. Target direction (left/right) was counterbalanced across trials. Due to a coding error, one incongruent trial was dropped, resulting in 20 congruent trials, 20 neutral trials, and 19 incongruent trials per session.

Delay Discounting Task (Figure 1C)

Participants were asked to select their preferred option between an immediate and a delayed reward (Kable & Glimcher, 2007, 2010), and were informed that either option would equally increase their total bonus points. They were told that the options were hypothetical, but they should pick based on if they were experienced. After being given several practice trials, they started the main task which asked them to select the option they preferred. The immediate option was always fixed at $20 and the delayed option varied by delay (1, 10, 20, 50, 90, and 180 days) and magnitude ($21 to $110 in increments of $5). Thus, there were a total of 114 trials, all

with a unique combination of delay and magnitude. Participants had 15 seconds to select one of the options on each trial.

Probabilistic Bandit Task (Figure 1D)

Participants were instructed to select the stimulus they believed would have a higher probability of reward. After performing several practice trials, participants went through 5 blocks of a 30-trial (150 total) two-armed probabilistic bandit task with a single reversal randomly selected without replacement between the 10th and 20th trial mark. One arm provided a gold coin (worth 1 point) for 70% of the time while the other gave a reward only 30% of the time; whereupon a reversal within the block swapped the probabilities of reward for the two options. After a selection, a feedback stage showed their reward of the gold coin and an increase in their total points, or a red 'X' with no increase in bonus points. Participants had 2 seconds to select one of the two options on each trial, otherwise they were shown feedback of a red 'X' with a 'TOO SLOW' message.

Delayed Bandit Task (Figure 1E)

The goal of this task was to assess participants' strategy and ability to learn the delay associated with a cue. On each trial they made a choice between two stimuli (the two arms of the bandit), each stimulus being associated with a delay of 1 to 9 trials. This means that when an arm was chosen, its reward (always 1 point) would be delivered several trials later depending on the delay. Feedback was delivered without a discriminating cue (anonymous). Participants were instructed to learn the delay of each arm in the task for 12 trials. Then, at the end of the round, they were asked to provide an estimate of the trial-delay of each arm and instructed that correct estimates would be rewarded by 10 points for each arm. They were given a single practice session of 12 trials which ended with them providing an explicit delay estimate for both arms

with the question ‘Provide the Trial Delay’ with two empty boxes in each position. They played a total of 9 blocks with 12 trials per block and had the following trial delays for each arm: (1,9), (2,7), (4,5), (2,3), (5,7), (3,6), (2,4), (3,5), (5,6) with a randomized presentation of blocks and randomized left or right position. The feedback was aggregated, so for example, if a person on the first trial selected an arm with a 3-trial delay and then on the next trial selected the other arm with a 2-trial delay, then on the fourth trial they would receive 2 points (1 point for each arm, aggregated). Participants were given 15 seconds to decide but had unlimited time for the explicit recall questions.

Analysis

To understand the effects of the individual tasks and overarching relationship between tasks we implemented a stepwise analysis framework. First, raw data from each task was inspected for choice and RT. Second, we applied both frequentist ordinary least squares (OLS) and mixed-effects regressions to each individual task and predicted their marginal group means. OLS regressions were additionally reported to aid interpretability for readers across disciplines, and divergence in regression approaches would further illustrate the impact of violating trial-level independence assumptions. Third, we applied a unified hierarchical Bayesian framework to understand both individual tasks and the covariance structure between them. Fourth, we utilized point estimates in our Bayesian models to reproduce the key effects from the frequentist mixed-effects models. We explain each section in detail below.

OLS and *Mixed-Effects Regressions*

Data from each task was inspected by calculating choice and RT across experimental conditions. Next, we looked towards each individual task and used both OLS and mixed-effects regressions to identify differences in the behavioral metrics as a function of experimental

conditions (equation 1). While mixed-effects models account for the dependency among trials within participants, OLS on participant-level summaries provide estimates under the simplifying assumption of trial independence (participant-level). For brevity on the OLS regressions, we report only p-values and effect sizes, but all statistics can be found in the appendix (**Participant-Level OLS Regressions** in Appendix A). Because we were interested in predicted choice probabilities and RTs, we used the 'emmeans' package to calculate marginal means based on each experiment's conditions (**Mixed-Effects Regressions** in Appendix A).

$$\text{metric} \sim \text{regressors} + (1 \mid \text{participant}) \quad (1)$$

Individual Task Models

Go/No-Go task (ng). The modeling here will replicate and validate typical behavioral signatures (**E3**) of reversal and condition (Go and No-Go). A drift diffusion model (DDM) with stimulus-coding was used to model upper-bound Go and lower-bound No-Go responses (**Drift Diffusion Modeling, Priors, and Transformation** in Appendix A). We implemented our statistical architecture in Stan, which provides the 'wiener' function to implement the DDM (Stan Development Team, 2024; Vandekerckhove et al., 2011). While we consider stimulus coding (i.e., Go and No-Go responses as boundaries) to directly translate into DDM parameters in terms of response across tasks, correct No-Go and incorrect Go responses lack observable RTs and thus cannot be directly modeled using standard DDM likelihoods. However, one can approximate the probability of observation using the following formula (Gondan et al., 2014; Letkiewicz et al., 2024; Wiecki et al., 2013):

$$p = \frac{exp\,(-2 * \upsilon_{ng} * \alpha_{ng} * \beta_{ng}) - 1}{exp\,(-2 * \upsilon_{ng} * \alpha_{ng}) - 1} \quad (2)$$

Here, the probability of the Go/No-Go (ng) hitting the upper boundary p uses the drift rate (υ_{ng}), boundary separation (α_{ng}), and starting point bias (β_{ng}), as the asymptotic first-passage

probability, integrating over all possible RTs (i.e., assuming an infinite response window). However, new approximations might assume the likelihood as a survival function using the response deadline (700 ms) accounting for right-hand censoring of RT (Henrich & Klauer, 2026). Either of these will be approximately equivalent when the drift rate is above 3 (**S1** in Appendix B). Finally, to account for the effects of reversal and No-Go trials on DDM parameters, we included reversal and No-Go coefficients for drift rate, boundary separation, and starting bias. Reversal coefficients reflect the deviation in parameters after versus before a reversal, while No-Go coefficients reflect the deviation on parameters on No-Go versus Go trials[1].

Flanker task. The modeling here will replicate and validate typical behavioral signatures (**E3**) of condition (congruent vs. incongruent). We also used a single DDM with accuracy coding across the three experimental conditions; that is, correct responses as the upper boundary and incorrect responses as the lower boundary. Due to the high similarity between neutral and congruent conditions we decided to represent these two conditions in the base parameter. We coded a regressor for the incongruent condition for deviating from the base drift rate, boundary separation, and starting bias, which allowed us to understand how incongruency affected cognitive processing of choice and RT. We justify the combination of neutral and congruent in the results below.

Delay Discounting task (DD). We fit the standard hyperbolic function (**Hyperbolic Drift Diffusion Model** in Appendix A) to calculate the value of the delayed option based on an hierarchical discounting parameter (k) (Horan et al., 2017; Kirby, 1997; Laibson, 1997).

$$V = \frac{A}{(1 + kD)} \tag{5}$$

[1] Note. All deviation parameters are transformed into whole values for visualization.

Here the value of the delayed option is calculated as the amount of the wager (A) discounted based on the delay (D). To provide some basic normalization to our equation, we divided the delayed amount by the immediate amount ($20) so that all value differences were calculated based on comparison to '1'. We estimated a hierarchical inverse temperature parameter that scaled value differences (V_D - V_I) linearly into a drift rate, which was then incorporated in a DDM to fit choices and RTs (Peters & D'Esposito, 2020; Rodriguez et al., 2015). We used stimulus coding so that the upper boundary was the selection of the delayed option as there were no correct or incorrect responses as to measure a person's preferences.

Probabilistic Bandit task (PB). The modeling here will replicate and validate typical behavioral signatures (**E3**) of reward probability reversal. We fit a Rescorla-Wagner recursive model, which generates actions from value estimates that were constructed from prediction errors (Rescorla & Wagner, 1972). Value calculations for each arm were initialized at 0.5 and updated independently in each block of 30 trials, with prediction errors updating the value of the chosen action:

$$\delta_t = r_t - v_t(a_t) \quad (3)$$

$$v_{t+1}(a_t) \leftarrow v_t(a_t) + \alpha\delta_t \quad (4)$$

At the current time point (t), this calculates the prediction error (δ) between the actual reward of 0 or 1 (r), and the current expected value (v). Actions (a) are then updated with the prediction error and a hierarchical learning rate parameter (α). The value difference between the two arms ($V_1 - V_2$) is then linearly scaled with a hierarchical inverse temperature parameter to be used as a drift rate. Utilizing this value-informed drift rate allows us to use the DDM as the choice rule with typical parameters of boundary separation, starting point bias, and non-decision time (Lei & Solway, 2022; Pedersen et al., 2017). We use accuracy coding so that the upper boundary

represents selecting the higher probability option. Additionally, we used reversal coefficients for the deviation from the initial inverse temperature, boundary separation, and starting point bias where upon a reversal the other stimulus becomes the upper boundary (inverting the value difference).

Delayed Bandit (DB). The modeling here will validate the stay/switch behavioral signature of better delay estimation (**H1.V**). We fit the delayed bandit with a single DDM using stimulus-coding with the upper boundary representing stay on current arm and the lower boundary representing switch to the other arm. Given the stay or switch nature of the bounds, we excluded the first trial from analysis. We reasoned that both the RT and choice to stay on the current arm would be important for estimation of the delay, and that this perseverance might be implicated in impulsivity. Our initial direction was from Pike-Burke et al. (2017, 2018) who proposes a solution to when delayed feedback is both aggregated (sums from previous trials) and anonymous (lacking cues to indicate which arm gave a reward) that relies on the agent's perseverance. That is, the credit assignment problem can be avoided if the agent selects one arm until feedback is received and then switches to the other arm for delay estimation. Given that we incentivized participants to estimate each arm's delay, we should assume that low switch rates would lead to better delay arm estimation. To validate this assumption, we also display the relationship between stay-switch behavior and the explicit delay estimation to justify the previous model (see Results).

Hierarchical Bayesian Model (HBM)

We also developed a framework for a Hierarchical Bayesian model (HBM), optimized under the programming language Stan (Stan Development Team, 2024), which provides robust

integration methods. This simultaneously fits each task separately into the models above with an explicit estimate of the correlation matrix between and across tasks.

$$\mu_i \sim N(0, 10), \qquad \sigma_i \sim N^+(0, 10), \tag{6}$$

$$L_i^\rho \sim LKJCorr_{Cholesky(2)}, \qquad Z_{i,j} \sim N(0, 1),$$

$$\theta_{i,j} = \mu_i + diag(\sigma_i) \cdot L_i^\rho \cdot Z_{i,j}$$

Here the generative model assumes that each task parameter (i) is sampled across all participants (j) from a Gaussian distribution with group-level mean (μ) and associated standard deviations (σ). These group-level parameters are joined by a matrix for individual deviations (*Z*), and a Cholesky factor for a correlation matrix (L^ρ)[2]. These are combined into the covariance matrix as the diagonal (i.e., diag) and the Cholesky factor. Finally, we use all these components in a linear combination to create a full matrix of task parameters for each participant (θ), representing each model's parameters (e.g, drift rate, boundary separation, discounting rate, etc.). As such, we use this hierarchy to identify the contribution of each parameter using a non-centered method to reduce optimization time (Stan Development Team, 2024). Note that the non-centered method is equivalent to the centered method, $\theta_{i,j} \sim N(\mu_i, \sum)$, with the covariance matrix ($\sum$). Converting the covariance matrix into a correlation matrix allows us to test many of our hypotheses (**H1**, **H2**, **H3**, **E1**, and **E2**).

Next, to perform posterior predictive checks, we used the point-estimates of the HBM to reconstruct trial-by-trial behavior of the participants in our mixed-effects models (**HBM** in Appendix A). We regenerated choice behavior if any RTs were under 200ms (Go/No-Go = 22.64%, all others < .001%), due to censored RTs producing large uncertainty in non-decision

[2] The Cholesky Factor improves efficiency and can be transformed back into a covariance or correlation matrix.

time. In total, we depict all tasks effects in five separate figures with both the mixed-effects regressions, posterior predictive checks, and distributions of our HBM model parameters.

Lastly, metrics of model parameter recovery and reliability were approximated. Using this simulated choice behavior to refit the HBM model led to an overall high correlation of recovery (r = .854) with recovery scatterplots for each parameter (**S2** in Appendix B). We reasoned that using a range of values to identify recovery would be too difficult due to the restricted limits of the DDM producing implausible values. For an estimate of reliability, we split the data for each task into even and odd trials. That is using the same model as described above, we fit our total model to two-difference splits of even and odd trials. This allowed us to get a split-half reliability for each parameter (**S3** in Appendix B).

Results

We display descriptives of the data from our five tasks with specific focus on the experimental conditions or metric of interest (Table 1). That is, for Go/No-Go, Flanker, and Probabilistic Bandit, we calculated the accuracy and RT for each condition. For Delay Discounting, we show the preference for delayed reward and RT for the different time delays (in days). Finally, for Delayed Bandit, we examine the absolute (denoted: '| |') delay deviation as compared to the objective trial arm-delay and the RT per number of switches in a block.

Table 1. Choice behavior and RT for each experiment by various conditions.

Task	$Condition_1$	$Condition_2$	M_{choice}	SD_{choice}	M_{RT}	SD_{RT}
Go/No-Go	Go	Initial	.95	.23	.52	.37
	No-Go	Initial	.84	.37	$.39^{!}$	$.1^{!}$
	Go	Reversal	.92	.27	.57	.43
	No-Go	Reversal	.77	.42	$.42^{!}$	$.11^{!}$
Flanker	Congruent		.95	.22	.39	.1

	Neutral		.95	.23	.44	.11
	Incongruent		.75	.43	.4	.1
DD	1-Day		.93	.25	1.06	.52
	10-20 Days		.65	.48	1.26	.68
	50-90 Days		.43	.5	1.26	.69
	180-Days		.36	.48	1.24	.67
PB	Initial	<5 trials	.68	.47	.52	.24
	Initial	>5 trials	.75	.43	.49	.21
	Reversal	<5 trials	.38	.49	.5	.22
	Reversal	>5 trials	.55	.5	.49	.21
	Switch Rate		$M_{\|error\|}$	$SD_{\|error\|}$	M_{RT}	SD_{RT}
DB	1		1.33	1.47	.75	.32
	2-3		1.89	1.95	.76	.38
	4-5		2.18	1.99	.74	.39
	6-7		2.48	2.13	.7	.3

Note. For Go/No-Go, mean RT (M_{RT}) for No-Go trial (!) corresponds to incorrect Go responses only, as correct No-Go responses are censored and do not have a RT. For Probabilistic bandit (PB), we partitioned the first five initial trials and five trials after reversal (<5 trials) to get an intuition on learning behavior compared to subsequent trials in each period (>5 trials). Delay Discounting (DD), we aggregated delayed preferences of 10 or 20 Days (10-20 Days) and 50 or 90 Days (50-90 Days). This was to reduce the size of the table, but the pattern holds without aggregation. For Delayed Bandit (DB), we calculated explicit memory recall error (0 = perfect recall) as the absolute (denoted: '| |') difference from the objective trial delay (for each arm), separately for different numbers of switches (per participant and block).

Given the initial descriptives, we can see that on the Go/No-Go task, participants appeared to be more accurate when performing initial and Go trials as compared to reversal and No-Go trials, whereas RT is quicker for No-Go trials (incorrect Go responses) compared to (correct) Go responses. In the flanker task, congruent and neutral trials show similar accuracy and RT, whereas for incongruent trials accuracy decreases but RT is about the same. In the probabilistic bandit task, we can see that over the course of learning (> 5 trials) accuracy

increases (compared to <5 trials) and the reversal decreases overall accuracy (compared to initial), but RTs appear stable. In the delay discounting task, the propensity to select the immediate option appears to increase as delays get longer, and RT only appears quicker in the 1-day condition. Finally, data from the delayed bandit task appears to match the delay estimation heuristic, as a single switch leads to better performance on predicting the arm's trial delay, compared to more switches. Overall, the effects of the conditions in the experiments matched our initial intuition for choice behavior, whereas RT shows a less consistent pattern across manipulations. However, these initial descriptives cannot make claims on whether these differences were meaningful or simply due to chance. Thus, we proceeded to analyze the data from each task using regression and HBM analyses.

Inhibitory Control (Go/No-Go)

To assess if there was a constant effect of the manipulations across trials, we aggregated accuracy and RT for each participant. We then used two OLS regressions with an interaction term (No-Go * reversal conditions) for both accuracy and RT. Accuracy residuals showed a large negative skew, which OLS regression is reasonably robust against. There was a significant interaction between no-go and reversal in predicting accuracy, $p = .036$, $\eta^2_p = .005$. There was no significant effect of reversal, $p = .116$. For no-go blocks, there was a significant effect, $p < .001$, $\eta^2_p = .06$. For RT, residuals were acceptable for normality. Predicting RT by the No-Go and reversal interaction, there was not a significant effect, $p = .13$. For reversal, there was a significant effect, $p = .04$, $\eta^2_p = .005$. For No-Go trials, there was also a significant effect, $p < .001$, $\eta^2_p = .03$. In the next paragraph, we find differences to the mixed-effects models and provide interpretation in reference to the OLS findings.

Next, we fit a mixed-effects model to the trial-by-trial data (equation 1), predicting accurate choice (0 = incorrect, 1 = correct) and RT from an interaction between trial type (No-Go vs Go) and reversal (reversal vs initial). For accurate choice (Figure 2A), the interaction was

significant with a reduction in odds-ratios (OR), $\chi^2(1) = 33.36$, $p < .001$, $OR_{nogo*reversal} = .925$, 95% CI [.77, 1.106][3], such that accuracy declined more sharply when No-Go trials occurred in the reversal block. The reversal regressor also had a significant effect on accuracy, with a reduction in odds, $\chi^2(1) = 35.83$, $p < .001$, $OR_{reversal} = .63$, 95% CI [.55, .711], confirming that accuracy decreases after the reversal compared to before. To note, this differs from the participant-level probability, $p = .116$, which may demonstrate that reversal effects are not constant across trials. Finally, the effect of no-go trials was also significant with a large reduction in odds, $\chi^2(1) = 331.15$, $p < .001$, $OR_{nogo} = .266$, 95% CI [.232, .305], indicating lower accuracy on No-Go than Go trials. Choice behavior was replicated through these manipulations.

For RTs (Figure 2B), the interaction was significant with an increase in geometric mean ratio (GMR), $\chi^2(1) = 148.68$, $p < .001$, $GMR_{nogo*reversal} = 1.034$, 95% CI [1.013, 1.055], suggesting participants' RT decreased more for No-Go trials than Go trials after reversal. Once again, this differed from participant-level effects, $p = .13$, which may demonstrate that participants were not constant on No-Go trials during reversal blocks. Reversal had a significant effect on RTs, showing an increase in ratio, $\chi^2(1) = 55.09$, $p < .001$, $GMR_{reversal} = 1.031$, 95% CI [1.026, 1.037], indicating participants slowed down after a reversal. Finally, the effect of no-go trials was also significant with a large reduction in ratio, $\chi^2(1) = 13579.34$, $p < .001$, $GMR_{nogo} = .913$, 95% CI [.899, .928], confirming the observed pattern of faster RTs when incorrect Go responses are made on No-Go trials. Overall RT manipulation effects were replicated.

Using point estimates from the HBM for each participant, data were simulated and then fit to the same equation from the multilevel models to assess predictive strength. Behavioral recovery was strong for all eight conditions (accuracy, rt; Go, No-Go; initial, reversal), with

[3] Note that p-values are calculated using the likelihood ratio test from the 'mixed' function in the 'afex' package (Luke, 2017), and thus, df-correct p-values might not reflect overlap in 95% CI's.

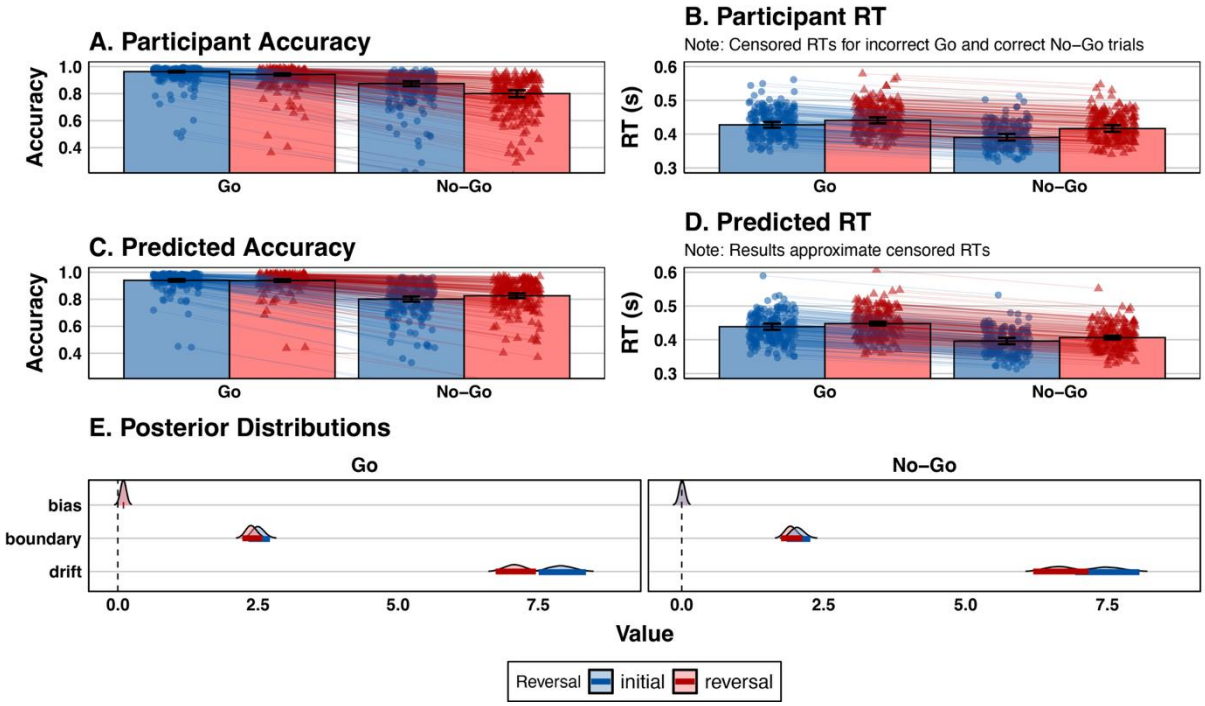
participant-level correlations between simulated and observed data ranging from r = .687 to .871, all p's < .001 (**S4** in Appendix B). Posterior distributions (group means) of the model parameters are displayed in Figure 2E, along with specific numerical means, and 95% Credible Intervals (CrI) (Table 2). Using the average of initial and reversal, parameters showed notable differences between Go and No-Go trials with boundary and bias reaching significance, (Go: drift rate = 7.499, boundary = 2.44, bias = .1; No-Go: drift rate = 7.07, boundary = 1.98, bias = .007). No-Go trials showed a non-significant reduction in drift rate and a significant reduction in boundary separation illustrating increased errors and faster RT. No-Go trials were associated with a significantly higher negative bias, illustrating less evidence was needed to withhold a response. Drift rate also showed an attenuation of effects in Go and No-Go trials when in the reversal block as the credible intervals overlapped. Overall, participants adapted their bias and boundary when presented with No-Go trials rather than modifying their rate of evidence accumulation. While the core parameters show hypothesized results, the asymmetry in starting-point bias was surprising and may illustrate a decrease in necessary evidence to withhold response.

Table 2. Go/No-Go hierarchical drift diffusion untransformed parameters (Figure 1E). Means and 95% credible intervals displayed.

Parameter	Reversal	Go/No-Go	Mean	95% CrI
Drift rate	initial	Go	7.91	[7.51, 8.35]
		No-Go	7.48	[6.93, 8.06]
	reversal	Go	7.09	[6.74, 7.46]
		No-Go	6.66	[6.19, 7.17]
Boundary	initial	Go	2.51	[2.33, 2.71]
		No-Go	2.04	[1.85, 2.26]
	reversal	Go	2.38	[2.22, 2.57]
		No-Go	1.92	[1.75, 2.13]
Bias	initial	Go	.1	[.09, .11]
		No-Go	.007	[.005, .009]
	reversal	Go	.1	[.09, .11]
		No-Go	.007	[.006, .009]
NDT	-	-	.14	[.13, .16]

Note. NDT = nondecision time which was not fit with condition-based regressors or to the covariance matrix. CrI = credible interval.

Figure 2. Go/No-Go Performance. (A). Marginal accuracy means with participant-level means (dots – jittered) and lines from initial to reversal. **(B).** Marginal RT means for correct Go and incorrect No-Go trials with participant-level means (dots – jittered) and lines from initial to reversal. Marginal means in (A) and (B) were obtained from mixed-effects regressions predicting accuracy (A) and RTs (B) from an interaction between trial type and reversal. **(C).** Predicted accuracy effects using simulation data from Bayesian parameter points estimates. **(D).** Predicted RT effects using simulation data from Bayesian parameter points estimates. **(E).** Posterior distributions of group-level natural-scale means from all Bayesian hierarchical DDM parameters excluding nondecision time. Darker segment lines for each parameter represent 95% credible interval.



Attentional Control (Flanker)

To see if there was a constant effect of the manipulations across trials, we aggregated accuracy and RT for each participant. We then used two OLS regressions with a single regressor term (incongruent) for both choice and RT. For choice behavior, residuals showed a large negative skew. Predicting choice by incongruency showed a significant effect, $p < .001$, $\eta^2_p = .25$. For RT, residuals were acceptable for normality. Predicting RT by incongruency also showed a significant effect, $p < .001$, $\eta^2_p = .15$. We find no substantial differences to the mixed-effects models and provide interpretations in the next paragraph.

Next, we examined participants' trial-by-trial choice and RT (equation 1) with an incongruency regressor. For choice behavior (Figure 3A), we show that the effect of incongruency significantly reduced odds-ratios, $\chi^2(1) = 2583.75$, $p < .001$, $OR_{incongruency} = .096$, 95% CI [.086, .11], showing an overall reduction in accuracy. For RT (Figure 3B), we show that incongruency had a significant effect with an increase in geometric mean ratio, $\chi^2(1) = 2968.96$, $p < .001$, $GMR_{incongruency} = 1.121$, 95% CI [1.116, 1.125], where participants took longer in incongruent trials. Both effects were consistent with participant-level effects and may illustrate consistency of effects across trials. This provides support for replication of known conditional effects in both choice and RT.

Using point estimates from the HBM for each participant, we simulated data and fitted the same equation from the multilevel models to assess predictive strength. Behavioral recovery was strong, with participant-level correlations between simulated and observed data ranging from: $r = .88$ to .98, all p's $< .001$ (**S5** in Appendix B). We display both the graphical representation of the group-means (Figure 3E) along with specific numerical means, and 95% Credible Intervals (Table 3). Incongruent trials showed significant reductions in drift rate (3.48) without a compensatory boundary increase (congruent = .88, incongruent = .87) suggesting participants did not adopt a more conservative threshold, despite the increased difficulty. Additionally, starting-point bias was significantly negative across both trial types, indicating a general bias toward the incorrect boundary. The difference between congruent and incongruent conditions bias was non-significant, (congruent = .31, incongruent = .28). These results show overall hypothesized effects with the reduction in accuracy and increase to RT.

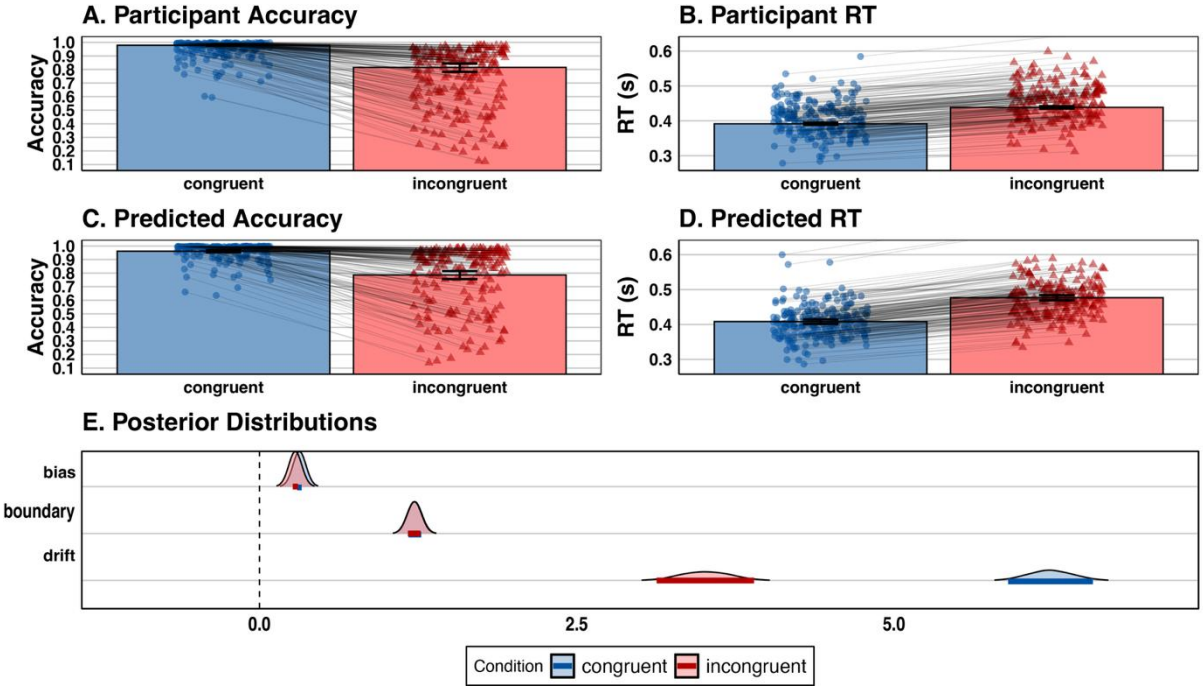
Table 3. Flanker hierarchical drift diffusion natural-scale parameters (Figure 3E). Means and 95% credible intervals displayed.

Parameter	Condition	Mean	95% CrI
Drift rate	congruent	6.23	[5.9, 6.57]
	incongruent	3.52	[3.13, 3.9]
Boundary	congruent	1.22	[1.17, 1.28]

	incongruent	1.22	[1.17, 1.27]
Bias	congruent	.31	[.297, .332]
	incongruent	.28	[.262, .303]
NDT	-	.26	[.25, .27]

Note. NDT = nondecision time which was not fit with regressors or to the covariance matrix. CrI = credible interval.

Figure 3. Flanker Performance. (A). Marginal accuracy means with participant-level means (dots – jittered) and lines from congruent to incongruent trials. **(B).** Marginal RT means with participant-level means (dots – jittered) and lines from congruent to incongruent trials. Marginal means in (A) and (B) were obtained from mixed-effects regressions predicting accuracy (A) and RTs (B) from trial type. **(C).** Predicted accuracy effects using simulation data from Bayesian parameter points estimates. **(D).** Predicted RT effects using simulation data from Bayesian parameter points estimates. **(E).** Posterior distributions of group-level natural-scale means from all Bayesian hierarchical DDM parameters excluding nondecision time. Note. Darker segment lines for each parameter represent 95% credible interval.



Temporal Discounting (Delay Discounting Task)

We used participant-level regression with delay as an orthogonal polynomial contrast to assess a constant effect of delay and magnitude across trials. Thus, two OLS regressions were used to predict both choice and RT. For choice behavior, residuals showed a departure from normality, which is predictable given the typical non-linear effect of delay. Predicting choice by delay showed a significant linear effect, $p < .001$, and a significant quadratic effect, $p < .001$,

with an overall large effect size, $\eta^2_p = .25$, all other effects were not significant. For RT, residuals were acceptable for normality. Predicting RT by delay showed a significant linear effect, $p < .001$, and significant quadratic effect, $p < .001$, with an overall small effect size, $\eta^2_p = .03$, all other effects were not significant. For both metrics, the participant-level quadratic effect demonstrates nonlinearity in choice and RT. We present the mixed-effects models below, which show no differences to the OLS regressions.

Next, we looked towards trial-by-trial choice and RT (equation 1) with both a fixed magnitude and delay regressor along with a random intercept of participant. For choice behavior (Figure 4A), we show that delay significantly decreased odds-ratios (OR) of choosing the delayed option, $\chi^2(5) = 6958.38$, $p < .001$, $OR_{linear} = .006$, 95% CI [.005, .007], $OR_{quadratic} = 5.602$, 95% CI [4.93, 6.37], and a significant increase for magnitude, $\chi^2(1) = 2249.15$, $p < .001$, $OR_{magnitude} = 2.37$, 95% CI [2.28, 2.46]. All delay effects were significant for the other polynomials, showing that preference for the delayed option had a non-monotonic decline in selection. For RT (Figure 4B), we show that delay was significant with an increase in ratio, $\chi^2(5) = 588.39$, $p < .001$, $GMR_{linear} = 1.12$, 95% CI [1.09, 1.14], $GMR_{quadratic} = .87$, 95% CI [.87, .88], and a significant decrease for magnitude, $\chi^2(1) = 158.06$, $p < .001$, $GMR_{magnitude} = .98$, 95% CI [.976, .983], all effects other than polynomial-5 were significant, demonstrating that RT also had a non-monotonic effect. Both metrics showed agreement on participant-level effects of linearity and rate of change, but trial-level revealed more variability regarding the differing delays.

Using point estimates from the HBM for each participant, we simulated data which was refit in the multi-level model equation to assess predictive strength. Behavioral recovery was strong, with participant-level correlations between simulated and observed data ranging from $r = .84$ to .96, all p's $< .001$, and choice behavior for 1-day delay: $r = .61$ (**S6** in Appendix B). We display both the graphical representation of the group-means (Figure 4E) along with specific numerical means, and 95% Credible Intervals (Table 4). The low discount rate indicates

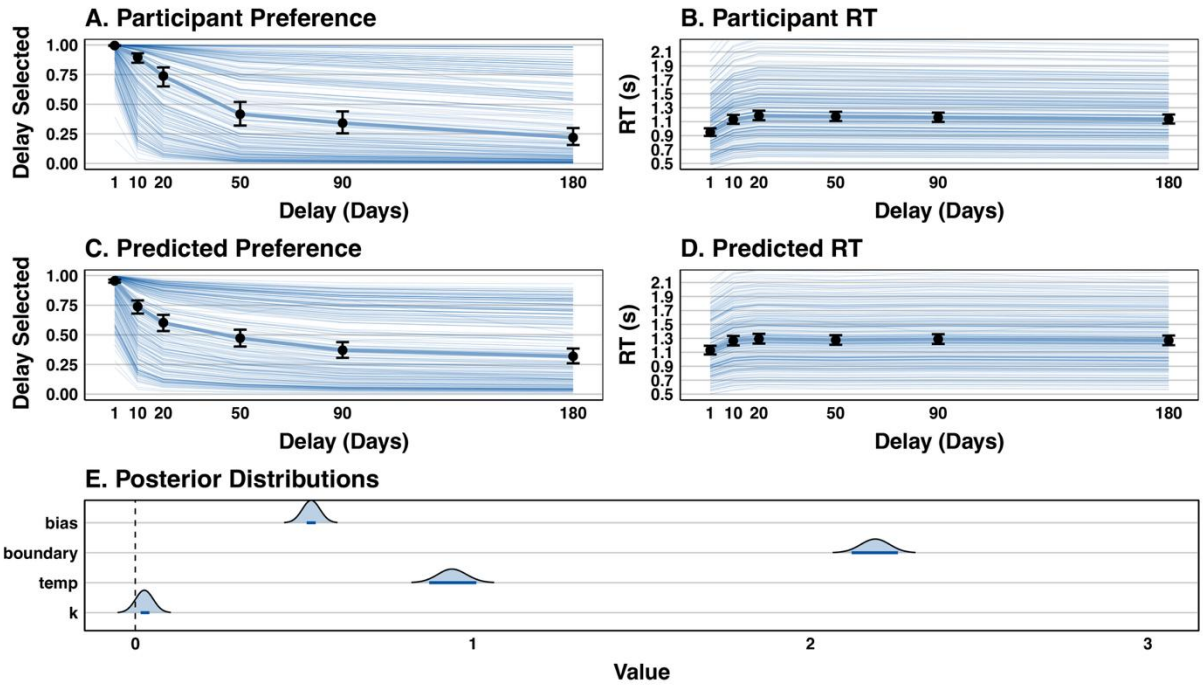
generally patient preferences where we display the 95% credible interval transformed on the delayed value of $100 using the mean value of k: 95% 1-day CrI [$96.01, $98.43], 95% 10-days CrI [$70.65, $86.24], 95% 20-days CrI [$54.61, $75.81], 95% 50-days CrI [$32.49, $55.63], 95% 90-days CrI [$21.1, $41.05], and 95% 180-days CrI [$11.79, $25.83]. Noteworthy, the k parameter had a large degree of variance in the group-distribution reflecting a wide range of k values. The inverse temperature parameter (temp = .94) indicated choice consistency, such that calculated value differences from k largely determined preference of choice. Participants maintained a moderate decision boundary (2.19), reflecting consistent response caution across choices. Finally, the starting point bias (.52) was significantly greater than .5 and participants required more evidence to select the immediate option.

Table 4. Delay Discounting hierarchical hyperbolic drift diffusion natural-scale parameters (Figure 4E). Means and 95% credible intervals displayed.

Parameter	Mean	95% CrI
k	.027	[.016, .042]
temp	.94	[.87, 1.09]
Boundary	2.19	[2.12, 2.26]
Bias	.52	[.508, .534]
NDT	.48	[.445, .517]

Note. NDT = nondecision time which was not fit to the overall covariance matrix. CrI = credible interval.

Figure 4. Delay Discounting Performance. (A). Marginal means for whether participants selected the option (preference) with participant individual lines for each delay. **(B).** Marginal RT means with participant-level means (dots – jittered) and lines for each delay. Marginal means in (A) and (B) were obtained from mixed-effects regressions predicting preference (A) and RTs (B) from trial type. **(C).** Predicted preference effects using simulation data from Bayesian parameter points estimates. **(D).** Predicted RT effects using simulation data from Bayesian parameter points estimates. **(E).** Posterior distributions of group-level natural-scale means from all Bayesian hierarchical DDM parameters excluding nondecision time. Darker segment lines for each parameter represent 95% credible interval (CrI).



Probabilistic Learning (Probabilistic Bandit)

To see if there was a constant effect of the manipulations on the trial-level, we aggregated accuracy and RT for each participant. We then used two OLS regressions with an interaction between reversal (0 = initial, 1 = post-reversal) and trial timing for both choice and RT. Trial timing was a dummy-coded variable with zero representing trials below the median for that specific block and one otherwise. This allowed us to maintain participant-level aggregation while also capturing early and late behavior from the learning task. For choice behavior, residuals followed a normal curve. Predicting choice showed a significant interaction, $p = .002$, $\eta^2_p = .01$. A significant main effect of reversal, $p < .001$, $\eta^2_p = .15$, and a main effect of timing, $p < .001$, $\eta^2_p = .03$. For RT, residuals showed a deviation from normality with a positive skew indicating possible violations. Predicting RT showed a significant interaction, $p = .004$, $\eta^2_p = .01$. A significant main effect of reversal, $p < .001$, $\eta^2_p = .02$, and a main effect of timing, $p < .001$, $\eta^2_p = .03$. We find consistency in significant differences to the mixed-effects models below.

Next, we looked towards the trial-by-trial choice and RT using equation 1 with both a fixed dichotomous variable for reversal with an interaction for a fixed effect of the current block trial. For choice behavior (Figure 5A), we show a significant interaction between reversal and current block trial, $\chi^2(1) = 29.5$, $p < .001$, $OR_{trial*reversal} = 1.18$, 95% CI [1.12, 1.24], identifying that participants learned at a faster rate after a reversal. There was a significant main effect of reversal, $\chi^2(1) = 1174.71$, $p < .001$, $OR_{reversal} = .51$, 95% CI [.48, .53], which indicates that accuracy went down after reversal. Additionally, there was a significant effect of trial, $\chi^2(1) = 207.42$, $p < .001$, $OR_{trial} = 1.28$, 95% CI [1.21, 1.35], once again showing an effect of learning over time. For RT (Figure 5B), we show a significant interaction between reversal and current block trial, $\chi^2(1) = 260.73$, $p < .001$, $GMR_{trial*reversal} = 1.09$, 95% CI [1.078, 1.098], indicating that after reversal participants took longer as trials increased. There was a significant main effect of reversal, $\chi^2(1) = 98.87$, $p < .001$, $GMR_{reversal} = .947$, 95% CI [.939, .955], highlighting that RT decreased after a reversal. Additionally, there was a significant effect of trial, $\chi^2(1) = 227.45$, $p < .001$, $GMR_{trial} = .907$, 95% CI [.897, .917], which demonstrates that RT went down over the course of the block. Both trial-level and participant-level showed the same effects, demonstrating possible consistency in participants' choice behavior. While this provides support for the reversal effects on reducing accuracy, RT showed the opposite of the hypothesized prediction with quicker decisions after reversal.

Using point estimates from the HBM for each participant, we simulated data and fitted the same multilevel models to assess predictive strength. Behavioral recovery was strong, with participant-level correlations between simulated and observed data, $r = .73$ to .95, all p's $< .001$ (**S7** in Appendix). We display both the graphical representation of the group-means (Figure 5E) along with specific numerical means and 95% Credible Intervals. First, we see a substantial individual variability in learning rate with a high mean update rate (learning rate = .52). To note, parameter recovery on learning rate was low-to-moderate (**S2** in Appendix B) and oftentimes is

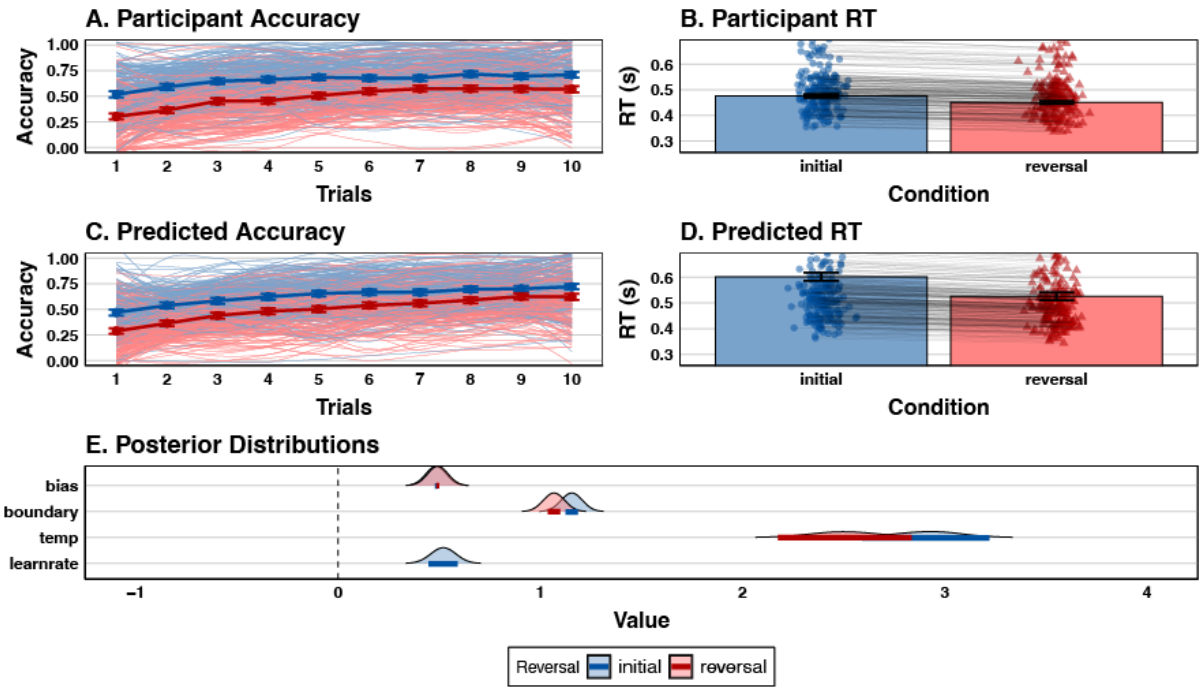
not a reliable parameter between contexts (Eckstein et al., 2022), limiting the reliability on interpretation. We can see that most of the parameters showed minimal to no difference across the initial and reversal conditions. However, boundary separation showed a reduction when moving from initial (1.16) to the reversal block (1.07), which illustrated a counterintuitive notion that participants required less evidence for the reversal block. Starting point bias showed no meaningful effect in either condition. The HBM results were consistent with previous analysis on reduction in accuracy through reducing inverse temperature, but boundary separation was also reduced, indicative of quicker decisions needing less evidence against our hypothesized slower RT.

Table 5. Probabilistic Bandit hierarchical Rescorla-Wagner drift diffusion natural-scale parameters (Figure 5E). Means and 95% credible intervals displayed.

Parameter	Condition	Mean	95% CrI
Temp	Initial	2.91	[2.59, 3.22]
	Reversal	2.5	[2.17, 2.83]
Boundary	Initial	1.16	[1.12, 1.18]
	Reversal	1.07	[1.04, 1.1]
Bias	Initial	.48	[.479, .49]
	Reversal	.49	[.486, .5]
Learning Rate	-	.52	[.45, .59]
NDT	-	.23	[.22, .24]

Note. NDT = nondecision time which was not fit to the overall covariance matrix. Learning rate did not differentiate for reversal. CrI = credible interval.

Figure 5. Probabilistic Bandit Learning. (A). Marginal accuracy means with participant-level means (dots – jittered) and lines from initial to reversal blocks. **(B).** Marginal RT means with participant-level means (dots – jittered) and lines from initial to reversal blocks. Marginal means in (A) and (B) were obtained from mixed-effects regressions predicting accuracy (A) and RTs (B) from block type. **(C).** Predicted accuracy effects using simulation data from Bayesian parameter points estimates. **(D).** Predicted RT effects using simulation data from Bayesian parameter points estimates. **(E).** Posterior distributions of natural-scale group-level means from all Bayesian hierarchical DDM parameters excluding nondecision time. Note. For (E), Darker segment lines for each parameter represent 95% credible interval (CrI).



Delayed Learning (Delayed Bandit)

Our analysis of the delayed bandit follows a slightly different approach to justify our use of stay-switch behavior as the principal metric for impulsivity. We performed participant and block-level aggregation and perform OLS regressions to test for interaction of two dummy-coded variables - switch rate within block and arm (quicker or slower delay) – to predict deviation from delay estimation, calculated as the absolute estimation error from the true arm delay (0 is perfect recall), as well as to predict RTs. For delay estimation, residuals had a slightly positive curve but once again OLS regression is robust to these violations. There was a positive interaction between switch rate and arm in delay estimation error, $p < .001$, $\eta^2_p = .02$, where increased switch rate was more detrimental to the slower arm's delay estimation. There was a main effect of switch rate, $p = .003$, $\eta^2_p = .003$, such that more switches during a block led to larger delay estimation error. Lastly there was also a significant main effect of arm, $p = .002$, $\eta^2_p = .002$, showing that the slower arm resulted in greater delay estimation error. For RT, residuals showed a deviation from normality with a positive skew. There was only a significant main effect of switch, $p =$

.002, η^2_p = .002, such that more switches were associated with faster RTs. All other effects were not significant, $p > .05$.

Given that delay estimation only occurred once within a block, we examined each participant's block-level estimation error as a function of trial-level switch interacting with long-arm with a gaussian link (equation 1). There was a significant interaction between switch rate and arm on estimation error (Figure 6A), $F(1, 41629.71) = 12.02$, $p < .001$, $b_{switch*arm} = 1.124$, 95% CI [1.052, 1.201], such that for the slower arm switch rate dramatically increased delay estimation error. There was a significant main effect of switch, $F(1, 41549.63) = 6.91$, $p = .009$, $b_{switch} = .986$, 95% CI [.94, 1.033], which showed that switch rate only appears to affect the slower arm delay estimation. Additionally, there was a significant effect of arm, $F(1, 214.34) = 148.35$, $p < .001$, $b_{arm} = 2.678$, 95% CI [2.285, 3.138], indicating as expected that delay estimation error was larger for longer delay. While this does provide support for the validating hypothesis (**H1.V**), the shorter arm delay estimation was largely unaffected by switch rate. For RT (Figure 6B), we investigated to see if there was a quadratic relationship with block trial as a multilevel log link regression (equation 1). There was a significant linear effect of trial, $F(1, 4941.78) = 767.92$, $p < .001$, $b_{linear} = .87$, 95% CI [.848, .884], and a significant quadratic effect, $F(1, 20557.34) = 571.75$, $p < .001$, $b_{quadratic} = 1.118$, 95% CI [1.108, 1.129]. Effects show that participants appeared to be more patient at the beginning and the end of a block.

Using point estimates from the HBM for each participant, we simulated participants' trial-level behavioral data. Using this simulated data, we attempted to predict delay estimation error using the trial-level simulated data and block-level best-fitting multilevel model. Although this only approximates the delay estimation mechanism, we could not find a well-fitting delay estimation mechanism to include in the HBM. Behavioral recovery was strong, with participant-level correlations between simulated and observed data, $r = .952$ and .975, all p's $< .001$. We display both the graphical representation of the group-means (Figure 6E) along with specific

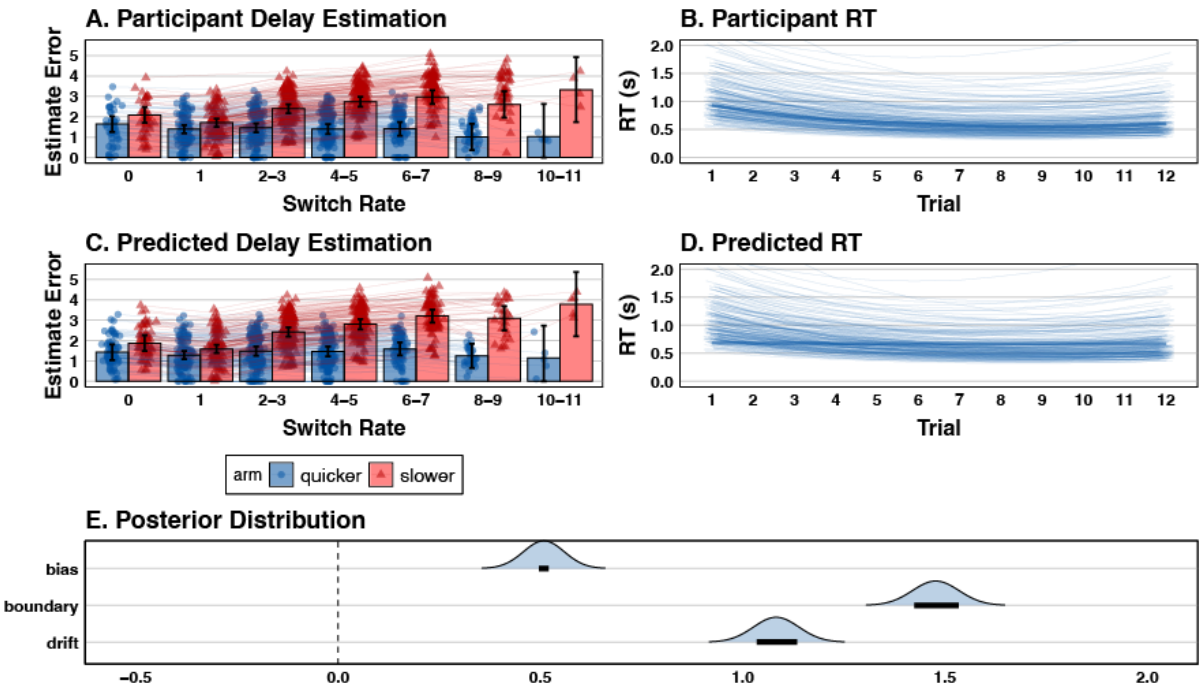
numerical means and 95% Credible Intervals (Table 6). The positive drift rate shows that on average participants were more likely to sample evidence to stay on their current arm. Starting-point bias did show a slight asymmetry favoring stay, though the effect was not significant and small.

Table 6. Delayed bandit hierarchical drift diffusion natural-scale parameters (Figure 6E). Means and 95% credible intervals displayed.

Parameter	Mean	95% CrI
Drift rate	1.08	[1.03, 1.13]
Boundary	1.48	[1.42, 1.53]
Bias	.508	[.497, .52]
NDT	.214	[.203, .225]

Note. NDT = nondecision time which was not fit to the overall covariance matrix. CrI = credible interval.

Figure 6. Delayed Bandit Learning. (A). Marginal mean delay estimation error with participant-level means (dots – jittered) and lines from participants' four block estimates. **(B).** Marginal means RTs with participant-level lines across block trials. Marginal means in (A) and (B) were obtained from mixed-effects regressions predicting delay estimation error (A) and RTs (B) from switch rate and arm, and block trial, respectively. **(C).** Predicted delay estimation error effects using simulation data from Bayesian parameter points estimates fed through best-fitting mixed-effects model from (A). **(D).** Predicted RT effects using simulation data from Bayesian parameter points estimates and real RT. **(E).** Posterior distributions of natural-scale group-level means from all Bayesian hierarchical DDM parameters excluding nondecision time. Note. For (E), Darker segment lines for each parameter represent 95% credible interval (CrI).



Task Correlations (Shared and Unique Variance in Impulsivity Dimensions)

To give greater insight into the relationships amongst the various experiments we performed several analyses. First, we correlated choice behavior of the five experiments (Figure 7A) and RT (Figure 7B), both of which were first averaged for each participant then correlated across participants between tasks. Then, we calculated typical metrics of impulsivity in these tasks and placed them into a correlation matrix (Figure 7C). Finally, the posterior population correlations were modelled in a covariance matrix within our hierarchical Bayesian modelling framework (Figure 7D). To deal with multiple comparisons we display a Bonferroni-corrected p-value above and non-corrected below the diagonal (Figure 7A-C), whereas for the HBM we use a covariance prior that leaned towards the null hypothesis ($\rho = 0$) for Figure 7D (**Drift Diffusion Modeling, Priors, and Transformation** in Appendix A).

Associations between Go/No-Go and flanker reveal shared quality of evidence accumulation

We find relationships in accuracy, RT, drift rate, and boundary for Go/No-Go and flanker task (**E1** and **E2**). Specifically, the ability to inhibit motor response (Go/No-Go choice accuracy) and attentional control (Flanker choice accuracy) were positively correlated, $r(208) = .392$, $p < .001$, 95% CI [.216, .544], as were the RTs in those decisions across participants, $r(208) = .31$, $p < .001$, 95% CI [.125, .474] (Figure 7A). This was also reflected in typical impulsivity-like metrics like Go/No-Go false alarm rate and inaccuracy in Flanker incongruent trials, $r(208) = .391$, $p < .001$, 95% CI [.199, .554] (Figure 7B). Additionally, we examined difference metrics for Go/No-Go (Go – No-Go) and flanker (congruent – incongruent), there was a positive correlation between the two metrics which did not survive correction for multiple comparisons, $r(208) = .2$, $p = .19$, 95% CI [-.026, .409], all other correlations between difference metrics were not significant.

In the modelling parameters (Figure 7D), the participant-level correlations aligned with the posterior correlations between the two tasks in drift rate, $r = .37$, $p < .001$, 95% CrI [.26, .48], and boundary separation, $r = .26$, $p < .001$, 95% CrI [.11, .41], but not bias, $p = .11$ (Figure 7C). The regressors for No-Go and reversal – which capture the deviation in parameter values between No-Go and Go trials, and after versus before the reversals, respectively – exhibited mixed associations with flanker incongruent regressors, deviation between incongruent and congruent trials (Figure 7D). Given the low recovery rate of the regressors deviations (**S2** in Appendix B), one should be mindful of interpretation on deviation effects as these could be artifacts of within-task correlations (within black borders in Figure 7D). Specifically, we can see that on average participants' performance generalized across the two tasks in both accuracy (drift rate) and RT (drift rate and boundary separation).

Positive associations between probabilistic and delayed bandit display similar behaviors for optimal policy

The two bandit tasks (probabilistic bandit and delayed bandit), despite tapping into two distinct forms of credit assignment, showed several relationships across participants. Stay behavior in the delayed bandit was positively related to stay behavior in the probabilistic bandit, $r(208) = .44$, $p < .001$, 95% CI [.26, .6] (Figure 7C), which reflects the predicted hypothesized direction (**H3**) with participants who were mostly consistent with staying or exploiting in the two contextually different bandits. In terms of model parameters, this pattern emerged in a positive relationship between the probabilistic bandit inverse temperature and the delayed bandit drift rate, $r = .27$, $p < .001$, 95% CrI [.13, .42], as well as between boundary separation, $r = .29$, $p < .001$, 95% CrI [.16, .41], and bias, $r = .24$, $p = .02$, 95% CrI [.03, .45] between the two tasks (Figure 7D). These positive relationships in computational parameters appear to be consistent in information processing as to the relationship of stay behavior and accuracy.

Additionally, switching behavior in both bandits predicted delay estimation in the delayed bandit. Specifically, staying in the delayed bandit led to better delay estimation (i.e. less delay estimation error), $r(208) = -.49$, $p < .001$, 95% CI [-.635, -.316], consistent with Figure 6A, along with delay estimation on stay behavior in the probabilistic bandit, $r(208) = -.23$, $p = .016$, 95% CI [-.419, -.024]. Furthermore, accuracy in the probabilistic bandit was related to stay behavior in the delayed bandit, $r(208) = .21$, $p = .002$, 95% CI [.078, .337], and stay behavior in the probabilistic bandit, $r(208) = .44$, $p < .001$, 95% CI [.32, .54]. Accuracy (higher probability or lower delay arm selected) did not correlate overall between bandits, $p = .565$, but participant-level RT did, $r(208) = .383$, $p < .001$, 95% CI [.205, .536] (Figure 7A-B). In summary, bivariate correlates between delayed bandit and probabilistic bandit showed support for some of our

hypothesized relationship; however, the lack of any correlates with learning rate in the probabilistic bandit did not confirm our hypothesis.

Attentional control shows stronger associations to delayed learning than inhibitory control

Comparing variables across cognitive domains (cognitive control versus reward learning), we find the strongest associations (i) between Flanker task accuracy and delayed bandit stay rate, (ii) between probabilistic bandit and both inhibitory tasks for RT, and (iii) in drift rate/inverse temperature between delayed bandit and Flanker and between probabilistic bandit and both inhibitory tasks. Stay rate in the delayed bandit task was positively correlated with accuracy in the Flanker task, $r(208) = .27$, $p < .001$, 95% CI [.077, .436] (Figure 7A), but not with Go/No-Go accuracy, $r(208) = .18$, $p = .081$, 95% CI [-.011, .362]. Probabilistic bandit accuracy was not correlated with either inhibitory task accuracy for Go/No-Go, $p = 1$, and Flanker, $r(208) = .19$, $p = .068$, 95% CI [-.01, .37]. Mean RTs were significantly correlated between the probabilistic bandit task and both Go/No-Go, $r(208) = .287$, $p < .001$, 95% CI [.099, .454], and Flanker, $r(208) = .361$, $p < .001$, 95% CI [.181, .518] (Figure 7B). However, RT correlations between the delayed bandit task and both inhibitory tasks did not survive correction for multiple comparisons, Flanker, $r(208) = .174$, $p = .14$, 95% CI [-.031, .345], Go/No-Go, $p = 1$. Accuracy on Flanker incongruent trials was correlated with delayed bandit stay, $r(208) = .239$, $p = .01$, 95% CI [.033, .426], and positively correlated with probabilistic bandit switch rate, $r(208) = .294$, $p = .01$, 95% CI [.092, .473] (Figure 7C). However, associations between Go/No-Go false alarm rate and stay/switch rate from either bandit task were not significant, both p's = 1. Specifically, delayed bandit drift rate correlated with Flanker drift rate, $r = .199$, $p = .002$, 95% CrI [.066, .325], while probabilistic bandit inverse temperature correlated with both Flanker drift

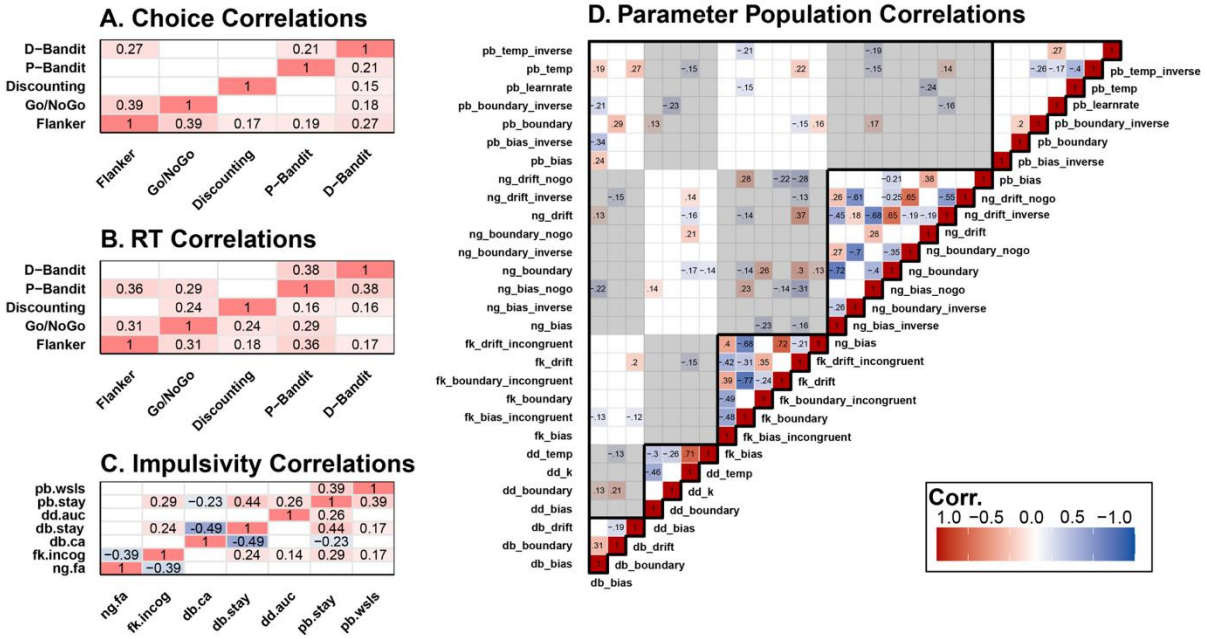
rate, r = .222, p = .002, 95% CrI [.073, .36], and Go/No-Go drift rate, r = .144, p = .029, 95% CrI [-.005, .287]. Bias or boundary separation posterior parameters were not correlated between inhibitory tasks and bandit tasks. This did not provide much support for our hypothesized relationship between parameters, but there was a relationship between both bandits' drift and inverse temperature and flanker whereas the relation to Go/No-Go was weaker (**H2**).

Delay discounting shows some associations to other dimensions

Preference for the delayed reward was correlated with performance on flanker, $r(208)$ = .17, p = .11, 95% CI [-.02, .36] but did not survive Bonferroni-correction along with other tasks' choice behavior (all p's > .11, Figure 7A). The delay discounting RTs were correlated with Go/No-Go RTs, $r(208)$ = .242, p = .004, 95% CI [.051, .415], all other correlations with RTs in other tasks did not survive multiple comparisons, all p's > .108 (Figure 7B). In terms of impulsivity-specific metrics, delay discounting, measured as the area under the curve (AUC), represents a metric of discounting using indifference points and integration where higher values indicate less discounting (Myerson et al., 2001). AUC was positively correlated with probabilistic bandit stay rate, $r(208)$ = .259, p < .001, 95% CI [.054, .443], suggesting that the more patient a participant is (preferring larger-later option) the more likely to stay on decision in a probabilistic environment. Finally, the model's discounting parameter (k) showed relationships to Go/No-Go drift rate, r = -.158, p = .006, 95% CrI [-.283, -.033], flanker drift rate, r = -.154, p = .008, 95% CrI [-.281, -.027], and probabilistic bandit inverse temperature, r = -.145, p = .028, 95% CrI [-.29, .005]. That is, if a participant selects the larger-later option in a delay discounting framework, then they are more likely to have higher quality of information gathering in probabilistic learning and cognitive control. While we find ample support for our hypothesized relationships, the delay discounting boundary and delayed bandit boundary was significant, r =

.21, p = .008, 95% CrI [.079, .344], but not between drift rates or discounting rate against our hypothesis (**H1**).

Figure 7. Correlations between behavioral measures. Cross-task correlations for data (A-C) with Bonferroni-corrected for multiple comparisons and below diagonal are non-corrected. **(A).** Correlations between mean participant-level accuracy (P-Bandit = probabilistic bandit, Go/No-Go, flanker), preference for the delayed option (discounting = delay discounting), and stay rate (D-Bandit = delayed bandit). **(B).** Correlation between mean participant-level RT (RT) across tasks. **(C).** Correlations between mean participant-level impulsivity metrics. **(D).** Posterior correlates between hierarchical Bayesian model parameters, extracted from the model's covariance matrix, both within- (above thick black line) and across tasks.



Note. Blank correlations were not significant (p > .05). For **(D)**, added black lines to distinguish within-task correlations from between-task correlations with grey overlay to help distinguish between-task parameters. Tasks are abbreviated as follows: pb = probabilistic bandit, db = delayed bandit, dd = delay discounting, fk = flanker, ng = go/no-go. Metrics include: wsls = win-stay lose-switch rate, switch = switch rate, auc = discounting area under the curve, ca = delay estimation, incog = accuracy on incongruent trials, fa = false alarm rate.

Chapter 1 Summary

Impulsivity is one such construct that is difficult to collapse and remains best described as a multidimensional construct (Dalley & Robbins, 2017; Evenden, 1999; Moeller et al., 2001; Okan & Hallquist, 2025; Strickland & Johnson, 2021). To better characterize the

multidimensional nature of impulsivity, the methodology utilized a Bayesian hierarchical framework to estimate computational models and their associated parameters. We found support for multiple hypotheses along with many unexpected findings. First, the implementation of a novel delayed bandit task showed that choice consistency led to better delay estimation (**H1.V**). Specifically, this was driven by participants' better estimation of the slower arm as a function of stay behavior. Second, we found correlations between probabilistic and delayed RL, such that more consistent behavior in both bandits led to greater reward maximization and delay estimation (**H3**). Response consistency in delayed learning was related to higher accuracy in the probabilistic learning. Third, we found evidence that discounting rate was related to evidence accumulation for CC and probabilistic RL but not for delayed RL (**H1**). Fourth, RL and CC did have weak-to-mixed correlations, but reward learning was more strongly associated to attentional control than to motor control (**H2**). The correlations across RL and CC appeared to be in relation to RT, but upon closer inspection, the rate at which a participant would scale their learned values was related to the vigor of evidence accumulation. In our exploration we find multiple instances of correlations between task-modalities (**E1**) and similar computational parameters (**E2**). Lastly, we replicated classic manipulations of reversing contingencies and inducing interference in participants' information processing (**E3**). Given the amount of information, we cover each of these hypotheses along with our exploratory search in the general discussion chapter.

Chapter 2: Personality Impulsivity

Impulsivity is comorbid with a variety of diagnoses, occurring under diagnoses of ADHD, bipolar disorder, personality disorders, and anxiety disorders to name a few (American Psychiatric Association, 2022). In measurement, trait impulsivity has been measured through a diverse assortment of paradigms, which spans psychomotor efficiency (Barratt, 1959; Stanford et

al., 2009), functional adaptivity (Dickman, 1990; Eysenck & Eysenck, 1978), and the big-five factor model of personality (Cyders et al., 2007; Whiteside & Lynam, 2001) paradigms. Specifically, psychomotor efficiency focuses on the efficiency of cognitive processes coordinating with motor movements to produce fluid and adaptive behavior. Overlapping with functional adaptivity, which divides impulsivity into functional and dysfunctional features that describe a participant's speed of processing in relation to the behaviors at hand. Once again, other measures of impulsivity have been yoked to the big-five factor model of personality, where extraversion, openness to experience, and neuroticism might uncover various features of an impulsive personality.

While each paradigm focuses on different attributes, the overall relationship of impulsivity to mental health outcomes is multidimensional and complex. Due to the sheer vastness of measuring apparatuses, our review will focus on Barratt's impulsivity measure (BIS-11), which is arguably the most common self-report instrument for research in measuring personality and behavioral impulsivity (Stanford et al., 2009). As such, we will review some brief history, the BIS-11, and its relation to other personality questionnaires, behavioral tasks, and clinical disorders.

Impulsivity Questionnaires

Historically, impulsivity has a long and broad background referenced from times of the Greeks to more modern western views on the 'will' (Fernández & Bravo, 2003). Overall, many of these early philosophical theories stemmed from disease-models of psychopathology which have normatively loaded the term of impulsivity to equate to problematic behavior. In psychology, the difficulty in defining the term reinforces the issues of the past, where operationalizing of impulsivity has found wide divergence (Evenden, 1999). As such,

impulsivity is thought to be a latent variable that is constructed from several independent factors with little agreement to the factors between studies (Evenden, 1999; Moeller et al., 2001). Certainly the broadness of defining and measuring impulsivity has left many with questions of what might be specific and precise to the construct (Strickland & Johnson, 2021). However, one should be mindful that most psychological constructs, such as intelligence or psychopathology, have taken generations to create stable, measurable, and predictive traits of impulsivity (Huang et al., 2024). One such example of the overtime improvement in impulsivity is found in the BIS-11.

The BIS-11 was originally designed to measure aspects of anxiety and impulsivity in relation to psychomotor learning efficiency (Barratt, 1959). Specifically, Barratt used two questionnaires of impulsivity and anxiety with a conflict learning task and attempted to tease anxiety from impulsivity in relation to RTs and error-rates. After several editions (Stanford et al., 2009), the measure has been carefully curated to produce a multidimensional composite of impulsivity while decorrelating similarly looking anxiety behaviors that would show in psychomotor learning. The questionnaire is composed of a two-level hierarchy with the first-order factors as the lower level and second-order factors as the higher level. There are two first-order factors for each of the three second-order factors with the second-order of attentional (first-order factors: attention and cognitive instability), motor (motor and perseverance), and nonplanning (self-control and cognitive complexity). The second-order factor of attentional measures the ability to focus (attention) and racing of thoughts (cognitive instability). The second-order factor of motor measures acting without thinking (motor) and behavior despite negative outcomes (perseverance). The last second-order factor of nonplanning measures enjoyment of mental challenge (cognitive complexity) and ability to plan and resist temptation

(self-control). As such, the BIS-11 factor structure has found relationships to mental health outcomes, other impulsiveness personality questionnaires, and real-world behaviors in a variety of studies (for a review, see Stanford et al., 2009).

While colloquial use of impulsivity has become synonymous with problematic behaviors (Bickel et al., 2012), rapid information processing and acting speed are crucial for advantageous behaviors (Bargh & Chartrand, 1999). That is, if the speed of action exceeds the time necessary to process information, then problematic behaviors may arise. Some of these questionnaires have attempted to capture the adaptability of rapid information processing, such as in the Dickman Impulsivity Inventory (Dickman, 1990). This inventory has two factors with a functional factor measuring how participants who engage in rapid information processing when applied to simple or routine behaviors but can become a dysfunctional factor when applied to more complex behaviors. For example, cutting your processing time short when deliberating over dinner plans is more adaptive than which job to choose. Interestingly, the BIS-11 factors have found relationships to the dysfunctional factor (Huang et al., 2024; Miller et al., 2004; Mobini et al., 2007) with some weak-and-mixed results between the BIS-11 and the functional factor (Miller et al., 2004; Mobini et al., 2007). While the BIS-11 appears to capture more aspects of dysfunctional impulsivity, the mixed results might indicate that a variety of personality attributes interact significantly with impulsivity.

To measure personality, researchers have often turned to the five-factor model of personality (McCrae & Costa, 1987), which has become especially popular for the development of scales measuring mental health outcomes. Another interesting application has been in using this personality framework to construct an impulsivity questionnaire. Whiteside & Lynam (2001) used the five-factor model of personality and a variety of impulsivity questionnaires to identify a

total of four factors, negative urgency, premeditation, perseverance, and sensation seeking (UPPS). With the addition of positive urgency (Cyders et al., 2007), the structure leaves five-factors in total (UPPS-P). Thus, positive and negative urgency relate to an internal drive to act after a positive or negative experience, premeditation is the forethought used before an action, perseverance is maintaining behavior despite punishment, and sensation seeking reflects engagement in behaviors that provide a rush of pleasure. Across the two studies, the BIS-11 attention and motor factors related to premeditation and negative urgency, whereas nonplanning related to premeditation and perseverance. While nonplanning appears most stable in relation to premeditation and perseverance, motor and attention appear to oscillate in relation to the two urgency factors, premeditation, and perseverance (Huang et al., 2024; MacKillop et al., 2016). Consequently, between-study heterogeneity of UPPS-P and BIS-11 estimates are well documented (Berg et al., 2015; Nuske et al., 2026).

Self-Report and Behavioral Impulsivity

The typical relationship between constructs of differing measurement apparatuses (correlational vs experimental) has been poor (Hedge et al., 2018; Mischel, 1974; Rouder & Haaf, 2019). This wouldn't be surprising considering the tenuous relationship of the abstractions researchers make of hypothetical measures, constructs, and apparatuses. Self-reported questionnaires often target latent constructs, reducing item-level measurement error through dimensional reduction methods. Whereas, behavioral tasks emphasize causal inference through experimental manipulation, where measurement error is mitigated via trial repetition and task design rather than participant sampling alone. The two measurement types have only identified a signal that is weak-or-mixed across studies of planning (Solway et al., 2017), waiting and discounting (Kvam et al., 2021), self-control (Duckworth & Kern, 2011), meta-analyses of

impulsivity (Cyders & Coskunpinar, 2011; Sharma et al., 2014), and task-specific reliabilities (Enkavi et al., 2019). We further consider that these relationships capture a large degree of individual heterogeneity in both objective-state and subjective-trait based measures of impulsivity and learning.

The relationship between delay discounting and the BIS-11 has long been considered from the perspective of separated dimensions of impulsivity (Broos et al., 2012; MacKillop et al., 2016). For example, the second-order factor of nonplanning relates to the rate of discounting (MacKillop et al., 2016; Mobini et al., 2007), but others have found that discounting relates to every BIS-11 factor (Cosenza & Nigro, 2015) and to none of the factors (Baumann & Odum, 2012; Broos et al., 2012; Huang et al., 2024). To note, most of these studies find a faint signal (r = .15~) with some exceptions and often rely on large sample sizes to declare significance. One meta-analysis (MacKillop et al., 2011) on differences in TD between a criterion group displaying addictive behavior and a control showed a moderate effect size on subclinical (d = .45) and clinical (d = .61) groups. Despite the overall effects, discounting rate was largely dependent on which task was administered as across all 64 studies the effect size was low (d = .15). These mixed findings are often noted in discounting (Yeh et al., 2021), which has led to additional arguments about the importance of discounting as an individual measure (Bailey et al., 2021). Stated more plainly, due to the large diversity in measurements and task specifications, discounting can lead to a diversity of results that don't clearly map with one another. One such proposed solution (Bailey et al., 2021) is to increase the size of effect by making TD tasks more context-dependent on a particular item of habit.

While TD might map more to response preparation, CC appears to recruit similar and different processes for response initiation (Dalley et al., 2011). Inhibitory control measured in a

Go/No-Go task was found to inversely relate to the first-order cognitive complexity (Reynolds et al., 2006) along with its second-order factor of nonplanning (Skippen et al., 2019). Whereas other studies have found no relationship between the BIS-11 and inhibitory control in the Go/No-Go (MacKillop et al., 2016; Sharma et al., 2014; Wilbertz et al., 2014). Furthermore, these weak-to-mixed relationships are often seen between differing CC tasks and impulsivity, even when using advanced modeling techniques (Hall et al., 2021), or in meta-analyses (Sharma et al., 2014). One note of interest was that among the factors created across the behavioral tasks, the Go/No-Go was the only one to load on both inattention and inhibition factors. While certain CC tasks might tap into more decision mechanisms, learning over the course of time might be susceptible to personality dimensions of impulsivity.

Among the RL literature, we find BIS-11 is commonly used in a variety of learning tasks. In one study, the BIS-11 only revealed weak negative correlations with the later blocks of the Iowa Gambling Task, where one must learn the reward value of four decks (Vasconcelos et al., 2014). In another case, BIS-11 was related to increased propensity to switch following three repeated unrewarded outcomes (Zou et al., 2022). Skvortsova & Hauser (2022) found that both impulsivity and compulsivity were related to an increase in exploration, but upon adding a free parameter of noise to the learning rate, motor impulsivity was significantly related to learning noise. Using the BIS-11 to capture low-and-high impulsivity groups found a complex interaction between the groups, model-based behavior, and fluid intelligence (Reiter et al., 2016). In a gambling task, the nonplanning factor related to a reduction in learning rate and asymmetry in probability weighting (M. S. Lim et al., 2015), but was found to increase learning rates in ADHD participants (Kim & Seo, 2021). The BIS-11 also appears to primarily load onto an anxious-depression and compulsive-intrusive factor where the compulsive-intrusive factor is related to

maladaptive weighting of option values (Gillan et al., 2016; Lee et al., 2023). In another interesting modeling approach, high BIS-11 scores were associated with a cluster of participants that didn't engage any RL model (neither model-free nor model-based) and low BIS-11 scores were associated with increased reliance on model-free algorithms over that of model-based (Cockburn et al., 2024).

Clinical Impulsivity

Along with behavioral tasks, the BIS-11 has found to dissociate control groups from a plethora of clinical samples (Stanford et al., 2009), such as gambling disorders (Cosenza & Nigro, 2015; M. S. Lim et al., 2015; Nuske et al., 2026), substance use disorders (Haines et al., 2020; Kanen et al., 2019; T. V. Lim et al., 2025), impulse-related behaviors (Huang et al., 2024; Sharma et al., 2014), ADHD (Kim & Seo, 2021; Sripada & Weigard, 2021), and food addiction (Maxwell et al., 2020). While BIS-11 is reliable in capturing impulse-control behaviors in a variety of mental health outcomes, the issue of precision and specificity remains. There is still significant work to be done to translate the multidimensional construct of impulsivity into a transdiagnostic factor of mental health (Dalley & Robbins, 2017; Robbins et al., 2012; Strickland & Johnson, 2021). Furthermore, given both self-report and behavioral impulsivity mapping with clinical samples, finding a way to combine the two may lead to greater predictive validity.

One such theory has argued that impulsivity sits on similar dimensions as compulsivity (Robbins et al., 2012). This was reflected previously due to the BIS-11 mapping onto compulsivity and anxiety factors (Gillan et al., 2016; Lee et al., 2023), which may suggest that impulse-driven behavior leads to compulsive habit formation and anxious thoughts. In obsessive-compulsive disorder (OCD), while there is neurocognitive overlap with drug dependence, OCD lacks the additional propensity of impulsivity seen in drug addiction. Furthermore,

neurochemicals and circuitry in basal ganglia appear to contribute to the various dimensions of impulsivity though this is beyond the scope of this review (see Dalley & Robbins, 2017). At a broad level, modern theories have attempted to integrate both self-report and behavioral measures of impulsivity under transdiagnostic neuromodulation frameworks (Zhong et al., 2025).

Aim 2: Information Processing in Personality Impulsivity

In addition to our previous study, the current analysis extends information processing mechanisms into participants' self-report behavior. The same participants who completed the five behavioral tasks (Go/No-Go, flanker, delay discounting, two-armed probabilistic and delayed bandits) also completed six questionnaires revolving around impulsivity, addiction, and mental health symptoms along with a time estimation task and Raven's Progressive Matrices. As such, we justify our factors in a confirmatory factor analysis (CFA), perform post-hoc overviews using linear regression and amend our HBM with the BIS-11. We preregistered an overall general hypothesis (https://osf.io/3tq67/) which focuses on the relationship of our candidate parameters from each task (drift rates, inverse temperature, and discounting rate) and self-reported metrics (**H4**).

Methods

Participants

Using the same participants from our previous study (see Chapter 1 methods for details), we were able to retain a total of 215 participants to complete end-of-study surveys. Seven participants were removed due to not having data from behavioral tasks (due to missing data or not passing exclusion criteria). The resulting total sample (N = 208) contained 140 males and 68 females with a moderate age range (M_{age} = 32.47, SD_{age} = 10.52, Range = [18, 72]). The end-of-

study survey was brief (M_{time} = 13.24, SD_{time} = 6.67, Range = [4.25, 54.45]). This study was approved by the University of Maryland College Park IRB (ID: 2129967-2).

Materials and Procedure

After completion of the five cognitive tasks outlined in Chapter 1, participants were asked to fill out a random order of six surveys to the best of their abilities. First, we used the 4-point Likert scale (1 = rarely/never, 4 = almost always/always), Barratt Impulsivity Scale (BIS-11) which contains six first-order factors, including attention and cognitive instability (second-order: attention), motor and perseverance (motor), and self-control and cognitive complexity (nonplanning) (Barratt, 1983; Stanford et al., 2009). 30 items were administered that include questions like "I plan tasks carefully" or "I am a steady thinker". Next, we measured sensation seeking in the 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), Brief Sensation Seeking Scale (BSSS) which taps into willingness to engage in risky experiences (Hoyle et al., 2002). There was a total of 8 items which give questions like "I like wild parties" and "I would love to explore strange places". Third, the 5-point Likert scale (each item has unique labels), Obsessive Compulsive Drinking Scale (OCDS) measured a person's thinking and response behavior to consumption of alcohol (Anton et al., 1995). There were a total of 13 items covering ideas, thoughts, and distressful thinking about drinking, for example "How successful are you in stopping or diverting these thoughts when you're not drinking?" Fourth, the 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), Marijuana Craving Questionnaire - Short Form (MCQ-SF) assesses individuals' level of craving for cannabis (Heishman et al., 2009). Example questions include: "I could easily limit how much marijuana I smoked right now" or "If I smoked marijuana right now, I would feel less tense". Fifth, we used the 4-point Likert scale (0 = did not apply to me at all, 3 = applied to me very much, or most of the time), Depression,

Anxiety, and Stress Scale - Short Form (DASS-21) which contains a total of 21 questions with factors anxiety, stress, and depression (Henry & Crawford, 2005). The items are worded such as "I felt that I had nothing to look forward to" or "I felt scared without any good reason". Finally, we administered the 5-point Likert scale (1 = not at all characteristic of me, 5 = entirely characteristic of me), Intolerance of Uncertainty Scale Short Form (IUS-12) which measured a person's threshold to uncertain events with two subscales as prospective anxiety and inhibitory anxiety. This 12-item questionnaire contains items like "Uncertainty keeps me from living a full life" or "One should always look ahead so as to avoid surprises".

After these questionnaires were presented in random order, we presented the short 9-item Raven's Progressive Matrices with randomly selecting one of two different versions (Bilker et al., 2012). The 9-items (18 in total for two versions) are taken from the larger inventory of 60 items which produce similar psychometric properties while only using a quarter of the time of the 60-item inventory. This provided a proxy for intelligence to use as a covariate. Lastly, we administered three time-estimation tasks (intervals of 4, 8, and 12 seconds) that were presented in random order. The instructions for the time estimation tasks (Moreira et al., 2016; Wittmann & Paulus, 2008) required the participant to provide an estimate of the time passed on their screen before the 'next' button came up. Participants who have quicker internal clocks may display impulsivity due to processing time.

Analysis

To test our preregistered hypotheses, we fit our HBM with the BIS-11 total factor to measure the unique influence on our computational models. This was done to prevent fitting the three second-order factors separately (or any other potential factors) as the model fitting process is computationally taxing with limitations on the number of parameters and collinearity among

factors. We provide post-hoc estimates of the influence of other self-reports outside the model-fitting processes. This analysis focuses on a subset of the preregistered hypotheses, which is the relationship between computational parameters from the various impulsivity tasks and self-report metrics with a focus on the BIS-11.

Confirmatory Factor Analysis (CFA). Given our sample size and short-form questionnaire battery, we started with confirming the theoretical factor structure to the sample factor structure. The package 'lavaan' was used to perform the 'cfa' function using the 'MLR' estimator, which provide robust standard errors and scaled test statistics under non-normality. For each model, we examined the loadings and overall model fit and adjusted four of the questionnaires, BIS-11, BSSS, MCQ-SF, and IUS-12. These adjustments were performed for use in subsequent analyses, providing estimates that account for indicator-level measurement error. Any adjustments to the models were due to questions either having weak loadings ($\lambda < .3$), or Heywood cases ($\lambda => 1$ or $\rho => 1$, or an error variance of $\theta < 0$), and negative variance ($\theta < 0$). The adjusted factors were then used in the rest of the analysis models. For weak loadings we removed those items and for Heywood cases we collapsed the factors. We report McDonald's omega (ω) over Cronbach's alpha, which provides weighted factor loadings rather than assumed equal loadings (Hayes & Coutts, 2020; McDonald, 2013). The interpretation of McDonald's omega is the same for Cronbach's alpha for internal consistency; we report this metric before and after correction.

Bivariate Correlations. While using point estimates from a Bayesian model violates traditional assumptions on parameter independence, the focus was to give an approximate overview of one-to-one relationships. Correlations between accuracy, RT, and computational parameters were calculated with respect to self-report factors. Typically, multiple comparison

corrections are used to reduce type I error, but none of the correlates survived this process. Due to the exploratory nature of this correlational analysis, we additionally explore potential relationships between the primary impulsivity questionnaire, the BIS-11, and behavioral task effects.

Bayesian Model. We use our HBM (equation 6) with the following additions for each computational parameter.

$$\beta_{i,j} \sim N(0, 1)$$

$$\theta_{i,j} = \theta_{i,j} + \beta_{i,j} * factor \quad (7)$$

Here, we have all model parameters (θ) for each parameter (i) and each participant (j), where each parameter is a linear combination of itself and a participant-level factor score, weighted by a free parameter (β). In this case, we have one deviation for each computational parameter, with beta representing the influence of the total factor of BIS-11 combining questions from the subscales of nonplanning, motor, attention. This was predominantly done due to the high multicollinearity between the scales (Figure 7) which creates an unidentifiability issue and artifacts in the estimates. To be explicit, the beta parameter for the total score was not fit hierarchically and thus there is no pooling to the group mean to avoid a double shrinkage effect from the CFA.

Results

Confirmatory Factor Analysis

We display the CFA model-fit results (Table 7) and justify the reasoning behind the corrected forms. This was done to prevent inadmissible solutions that make fit indices uninterpretable. For the BIS-11, our correction was based on removing motor items: Q3 and Q4 (λ_{q3} = .109, λ_{q4} = .04) and nonplanning items: Q27 and Q29 (λ_{q27} = .207, λ_{q29} = .189). We see

that internal reliability for motor ($\omega_{uncorrected}$ = .665, $\omega_{corrected}$ = .803), nonplanning ($\omega_{uncorrected}$ = .752, $\omega_{corrected}$ = .794), and attention ($\omega_{uncorrected}$ = .639, $\omega_{corrected}$ = .648) was improved. To enter the factor into our HBM (equation 7), we combined all questions into a total score (excluding Q3, Q4, Q27, and Q29). For the BSSS, there was a Heywood case with boredom and thrill factors ($\rho_{bored\sim thrill}$ = 1.011), so we collapsed all questions into a single factor. The internal reliability of the single factor was good ($\omega_{corrected}$ = .829), whereas the original factors showed questionable-to-poor reliability for experience ($\omega_{uncorrected}$ = .666), boredom ($\omega_{uncorrected}$ = .412), thrill ($\omega_{uncorrected}$ = .609), and disinhibition ($\omega_{uncorrected}$ = .692). For the MCQ-SF, Q2 from the compulsivity factor showed negative loadings with Q7 and Q10 (λ_{q7} = -.507, λ_{q10} = -.851). Q2 was removed due to a researcher's error in improper inclusion of the reversal language from the original 47-item scale, as the short-form did not contain any reversals. After removal, the MCQ-SF showed a Heywood case with emotionality and expectancy factors ($\rho_{emotion\sim expect}$ = 1.061), so again all questions were collapsed into a single factor. We see excellent reliability for the single factor ($\omega_{corrected}$ = .963), whereas compulsivity ($\omega_{uncorrected}$ = .289), emotionality ($\omega_{uncorrected}$ = .887), expectancy ($\omega_{uncorrected}$ = .869), and purposefulness ($\omega_{uncorrected}$ = .91) had a mix of excellent-to-unacceptable. Lastly, the IUS-12 contained a single Heywood case between prospective and inhibitory factors ($\rho_{prosp\sim inhibit}$ = 1.01), and we collapsed all questions into one factor. Reliability once again improved with the single factor ($\omega_{corrected}$ = .91), as prospective ($\omega_{uncorrected}$ = .868) and inhibitory ($\omega_{uncorrected}$ = .773) were acceptable-to-good. For the OCDS and DASS-21, all loadings exclusion criteria were met and showed robust model properties. We display the correlation matrix between all self-report factors after corrections (Figure 8).

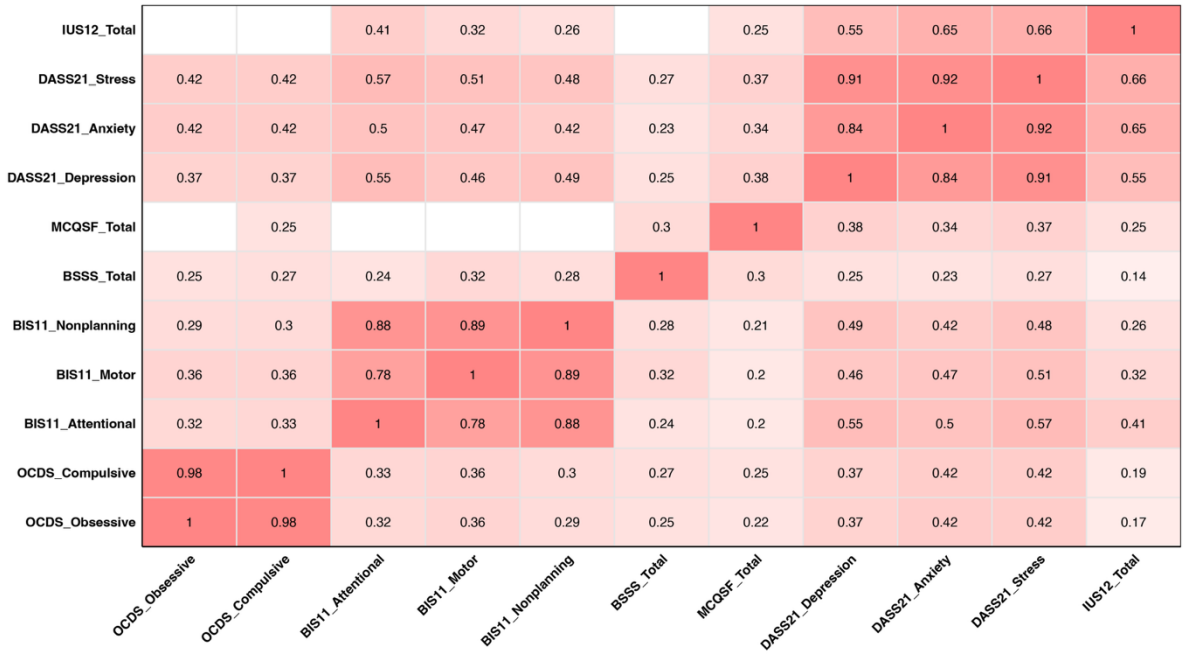
Table 7. Model fit indices before and after factor corrections.

Uncorrected	**CFI**	**TLI**	**RMSEA**	**SRMR**
BIS-11	.709	.676	.087	.093
BSSS	.951	.902	.084	.043

OCDS	.953	.953	.054	.052
MCQ-SF	.876	.829	.111	.083
DASS-21	.946	.94	.049	.051
IUS-12	.815	.769	.145	.093
Corrected	**CFI**	**TLI**	**RMSEA**	**SRMR**
BIS-11	.767	.723	.097	.09
BSSS	.913	.917	.094	.057
MCQ-SF	.816	.815	.139	.077
IUS-12	.858	.822	.132	.079

Note. CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Squared Error Approximation, SRMR = Standardized Root Mean Square Residual. Generally, CFI and TLI are expected above .9 and RMSEA and SRMR should be below .05

Figure 8. Correlations between self-report factors. After corrections, within and between factor correlations are displayed with blanks indicating nonsignificant correlates ($p > .05$). Above diagonal are Bonferroni-corrected p-value correlations and below diagonal are uncorrected.

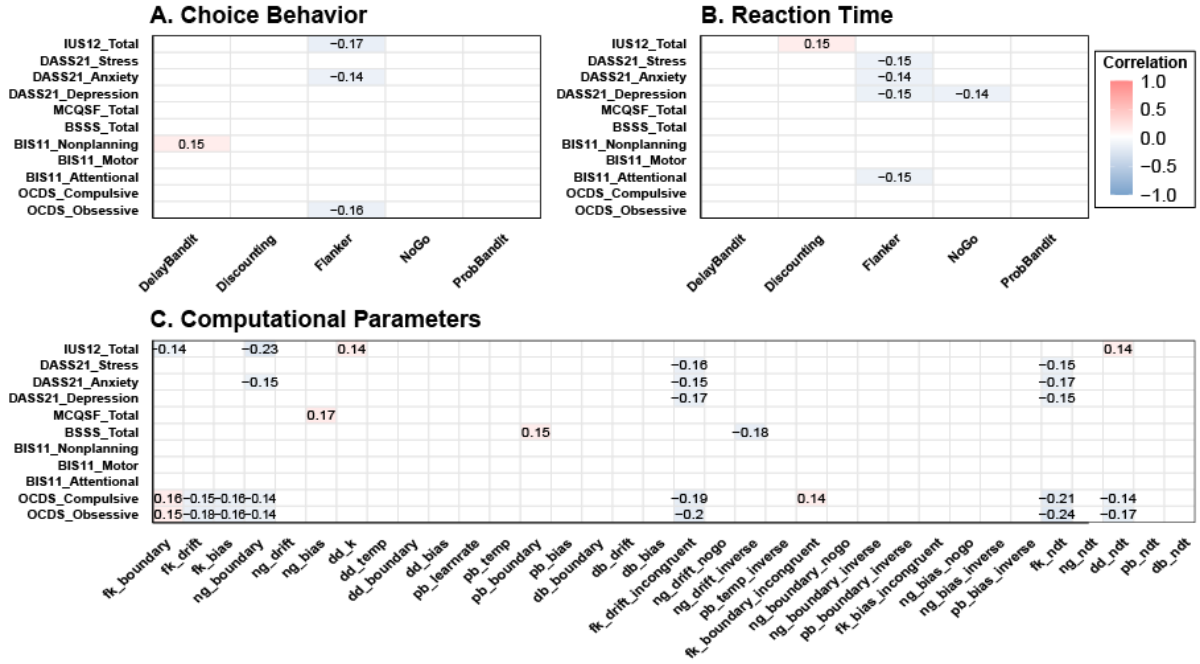


Bivariate Correlations

After clarifying the structure of our questionnaires, we entered them into multiple bivariate correlational analyses with choice behavior (Figure 9A), RT (Figure 9B), and model parameters (Figure 9C). None of the correlations survived corrections for multiple comparisons ($p > .05$), illustrating that associations between response modalities were faint and overall inconclusive. However, considering the lack of meaningful relationships, the signal that did

appear strongest was for the flanker across correlational structures, especially that of incongruent-deviate regressors. As per our measure of impulsivity, the BIS-11, there was one positive association between nonplanning and the stay rate in the delayed bandit, $r(206) = .15$, $p = .03$, 95% CI [.02, .28], along with another interesting relationship between intolerance of uncertainty and RT in delay discounting, $r(206) = .15$, $p = .03$, 95% CI [.02, .28].

Figure 9. Correlations between behavioral and self-report measurements. (A). Correlations between choice behavior in tasks and participants scores on questionnaires. **(B).** Correlations between RT in tasks and participants scores on questionnaires. **(C).** Correlations between point estimates from HBM for all behavioral tasks and questionnaire scores. Note. Blank correlations were not significant ($p > .05$) and show Pearson's correlation coefficient.



Exploratory Analysis of BIS-11 and Delayed Bandit

Given the weak but significant relationship of nonplanning with the delayed bandit, we performed a follow-up analysis as to identify whether the novel implementation uniquely identified a dimension of impulsivity. Due to stay rates decreasing the complexity of the problem, the positive relationship between nonplanning and delayed bandit stay rate was counterintuitive. Specifically, one might assume that nonplanning impulsivity would lead to an

increase in switch rate, as the stay behavior is a pseudo-optimal heuristic and increases delay estimation precision (see Chapter 1). First, an additional variable was calculated as the count of unrewarded trials until reward ($unrewarded_{streak}$) and this variable was used to subset our total delayed bandit dataset (26.25% of total dataset, participants average trial size: $M_{trials} = 26.1$, $SD_{trials} = 7.67$, total trials: $N_{trials} = 5491$). With this subset, we performed a multilevel logistic regression for each CFA factor score as a function of the z-scored unrewarded streak interacting with the factor score:

$$stay \sim BIS_{score} * unreward_{streak} + (1 + unreward_{streak} \mid participant) \qquad (8)$$

Each of the factor scores from the CFAs are placed into separate multilevel logistic regressions and then we extracted their interaction along with reported values (Table 8). The interaction captures the participants stay rate as the influence of the self-report factor on the count of unrewarded outcomes until they received a reward. The separation of factors was done due to the high degree of multicollinearity and to prevent estimate artifacts from high correlates. Specifically, the BIS-11 factors were displayed as the interaction with unrewarded outcomes effects to produce stay behavior (Figure 10). However, we provide all self-report factors interactions with unrewarded streak in the appendix (**S8** in Appendix B).

Figure 10. Delayed bandit stay behavior as a function of BIS-11 factor scores interacting with unrewarded trials. Note. Interactions were not corrected for multiple comparisons. For each factor, there was an interaction of unrewarded streak and BIS-11 attentional factor, Odds-Ratio (OR) = 3.42, $p = .02$, BIS-11 nonplanning factor, OR = 2.08, $p = .04$, and non-significant interaction with BIS-11 motor factor, OR = 1.39, $p = .31$.

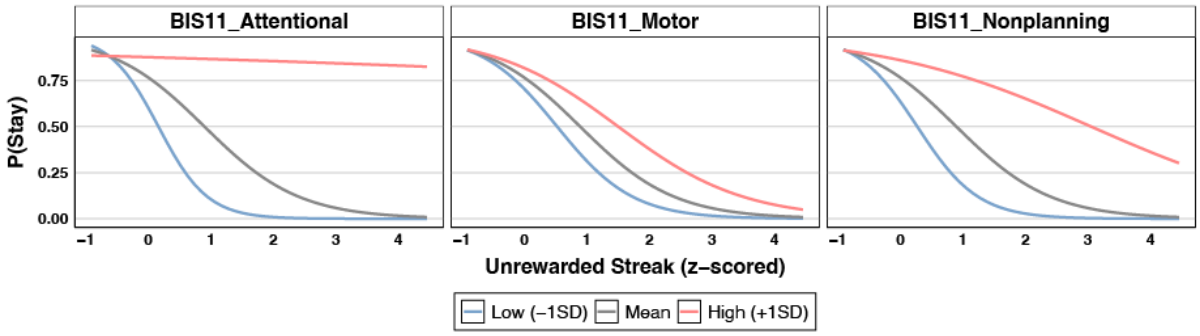


Table 8. Fixed interaction effects (factor * unrewarded streak) from individual multilevel logistic regressions (equation 8).

Interaction	Odds-Ratio	SE	Z-statistic	P-value	95% CI
OCDS-Obsessive	1.19	.33	.53	.6	[.62, 2.27]
OCDS-Compulsive	1.32	.33	.84	.4	[.69, 2.54]
*BIS11-Attentional	3.43	.53	2.35*	.02*	[1.22, 9.6]
BIS11-Motor	1.39	.32	1.03	.31	[.74, 2.6]
*BIS11-Nonplanning	2.08	.35	2.11*	.03*	[1.05, 4.09]
BSSS-Total	1.41	.24	1.42	.16	[.88, 2.25]
MCQSF-Total	1.27	.21	1.15	.25	[.84, 1.92]
DASS21-Depression	1.07	.24	.29	.77	[.67, 1.71]
DASS21-Anxiety	.76	.44	-.64	.52	[.32, 1.78]
DASS21-Stress	1.17	.26	.59	.56	[.7, 1.94]
IUS12-Total	.93	.19	-.46	.65	[.68, 1.27]

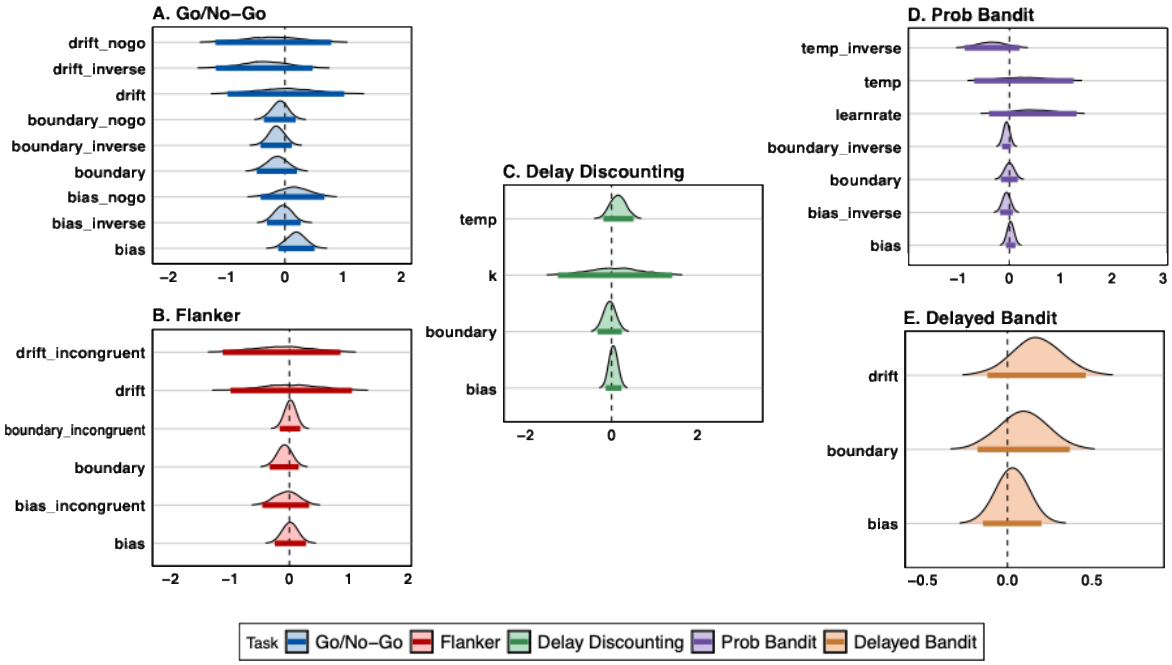
Note. Significant interactions are marked with * and grey-coded lines, $p < .05$. CI = confidence interval.

The exploratory analysis of the early trials in the delayed bandit produced two associations with attentional and nonplanning factors, however the direction of effect would be opposite to what one might expect. That is, higher BIS-11 scores appear to be leading participants to a more optimal strategy for delay estimation. However, one cannot be certain if this influence of BIS-11 factors is due to being more strategic, less sensitive to unrewarded outcomes, or lack of engagement. While our instructions may have included more bias to strategic planning for delay estimation, we performed some additional analyses to clarify that this effect was most likely not due to lack of engagement (**BIS-11 Factors and Lack of Engagement Analysis** in Appendix A).

HBM and BIS-11 Total Factor

Examining the fit of the HBM with the total BIS-11 factor score (after CFA adjustments), we find a large degree of uncertainty in the coefficients from the BIS-11 total factor score and thereby producing no significant results (Figure 11). Some of these parameters are trending towards an average effect of BIS-11 (starting-point bias in Go/No-Go, inverse temperature in delay discounting, and drift rate in the delayed bandit), but one should also be mindful of recovery (**S2** in Appendix B) and reliability (**S3** in Appendix B). That is, given the large number of parameters being estimated (between and within tasks), our sample size may not have sufficient power to precisely detect estimates. As such, the Bayesian framework identified large uncertainty in estimates and thereby finding most effects null but a few possible effects inconclusive.

Figure 11. BIS-11 total score effects on HBM parameters. (A). Effects on Go/No-Go. **(B).** Flanker. **(C).** Delay discounting **(D).** Probabilistic bandit. **(E).** Delayed bandit. Darker segment lines for each parameter represent 95% credible interval.



Chapter 2 Summary

The multidimensional construct of impulsivity has often been measured in both personality and behavioral paradigms. While both have been successful at predicting real world impulsive behavior (Romeu et al., 2020; Sharma et al., 2014), finding their relationship amongst various constructs has been difficult (Hedge et al., 2018; Rouder & Haaf, 2019). The analysis presented here sought a similar approach using classic metrics of choice and RT, alongside computational parameters. Our overall findings match the literature, where correlations between personality and behavioral approaches appear null or weak (**H4**). Among the weak relationships we do find potential for impulsivity overlap between stay-rate in the delayed bandit (primarily in early trials) and BIS-11, along with potential effects in Go/No-Go and delay discounting. However, these effects were inconclusive due to the high amount of uncertainty when estimating the full structure of the BIS-11's influence on all parameters. As such, a larger sample would be needed to clarify these relationships or to independently measure each task in a separate modeling framework. In conclusion, we provide interpretation and potential areas of further exploration in the general discussion of the multidimensional construct of impulsivity.

General Discussion

Evenden (1999) proposed the multidimensional construct of impulsivity, highlighting a range of possible dimensions across both task and questionnaire paradigms. Indeed, two identified dimensions of impulsivity, choice (TD) and action (CC), have been commonly observed across behavior (MacKillop et al., 2016), personality (Stanford et al., 2009), and cross-species (Broos et al., 2012). These dimensions tap into specific mechanisms of information processing, such as TD between value and time (Green & Myerson, 2004) along with detection and inhibition of prepotent responses (Botvinick et al., 2001). We further attempted to clarify a dimension of impulsivity in RL, although there is some debate as to whether learning task

paradigms should be subsumed into the choice dimension (Broos et al., 2012; Hamilton et al., 2015). To investigate the relationship among impulsivity dimensions (TD and CC) and RL, we ran multiple experimental designs alongside self-report measures, and we used a variety of inference, computational, and correlational modeling approaches. Here, we first explore the hypothesized cross-task relationships between RL and TD, RL and CC, and probabilistic and delayed RL. Second, we look at the personality factors of impulsivity (BIS-11) and their relationship to the behavioral dimensions of impulsivity. Third, we touch on the exploratory analysis, comment on limitations, and shed light into new directions for impulsivity.

When isolating the estimation of delay and probability in RL, we found support for a relationship between probabilistic RL and TD (**H1**). Specifically, the weighing of learned value differences that were scaled into evidence accumulators was weakly related to shallower discounting of hypothetical future reward. While the RL framework handles the discounting of future reward values (Sutton, 1984), investigation into cross-paradigm TD and RL remains sparse. One approach has been to use the two-step task, which decomposes participant behavior into model-free and model-based RL. Model-based RL has been theorized to represent an agent's degree of using estimated environmental contingencies when planning, similar to that of delay discounting, which is a descriptive measure of the ability to simulate future events (Amlung et al., 2019; Redish et al., 2008). However, evidence has shown mixed results: no relationship (Solway et al., 2017), steeper discounting associated with reduced model-based RL weight (Shenhav et al., 2017), and a positive association between evidence accumulation for larger-later reward model-based RL weight (Hunter et al., 2018). These mixed results demonstrate a potential signal between TD and RL paradigms, but once again the strength of association is negligible-to-weak as is usually identified across the literature (Rouder & Haaf, 2019). The

conversation (Hedge et al., 2018; Rouder & Haaf, 2019) has thus shifted towards increasing effect size through improved reliability, integrated cross-task hierarchical modeling approaches, and larger trial and sample sizes. We believe that combining our approach here with greater trial and sample sizes may inspire more confidence in inferring the relationship between theoretical constructs.

The relationship between TD and delayed RL did not substantiate anything meaningful regarding choice behavior. Nevertheless, when both tasks' data were decomposed into DDMs, there did appear to be a weak positive relationship in boundary separation between delayed RL and TD. That is, evidence needed to respond showed a positive relationship between these two constructs. While there has been no study to date to look at the relationship between TD and delayed RL, several studies have found mixed relationships to adjacent time-based constructs. Baumann and Odum (2012) found a weak relationship between overestimating the interval of time and steeper delay (but not probabilistic) discounting. Another study with a sample of adolescents with ADHD found a weak negative relationship between delay sensitivity (as measured by an increase in RT when presented with a stimulus for a longer period) and TD (Bitsakou et al., 2009). Participants with steeper discounting rates have also been shown to have a preference for gathering less complex information (Shenhav et al., 2017). Of course, these parallels are limited due to differences in operationalization, the restriction of associations to RT, and the delayed bandit instructions that prompted a heuristic strategy for delay estimation.

The strategy (**H1.V**) to overcome the complexity of temporal credit assignment (i.e. the assignment of delayed credit to each arm) was simply to stay on current choice until a reward is delivered (Pike-Burke et al., 2017, 2018). In our task designs, delayed rewards were summed across trials (aggregated) and feedback provided no indication of the which arm produced a

given outcome (anonymous). Additionally, we incentivized participants to estimate each arm's trial delay, thereby biasing their strategy toward using reward for delay estimation rather than purely accumulating reward. Despite the standard RL assumption of reward maximization, the evidence here suggests that humans can strategically their choice policy to minimize computational complexity when accruing future reward (Gershman, 2020; Gershman & Daw, 2017). As defined, computational complexity is the tax on cognitive resources to optimize the shortest path, whereas policy complexity is the degree of behavioral difficulty to implement. Other studies have found that consistency in maintaining choice can augment the assignment of credit in long time horizons (Bruinsma et al., 2024). Minimizing a switching policy results in more informational certainty and participants are more likely to deploy more efficient algorithms (Maulhardt et al., 2026). That is, when an environment provides cues, participants can rely on these for optimizing choice behavior (Akaishi et al., 2016; Gureckis & Love, 2009), but in their absence participants can also modify their choice strategy to reduce informational uncertainty.

While the broader term of regulation encompasses CC, the cognitive architecture remains separate, such as in attentional (flanker) and inhibitory (Go/No-Go) control. We found support for the relationship between RL and CC (**H2**) with an emphasis that attentional held an overall stronger correlation than inhibitory control. First, the drift rate associations across both control tasks matched a similar strength of association in another study (Hall et al., 2021). Second, the stay drift rate from the delayed RL was related to the accuracy drift rate in the attentional but not inhibitory control. This was surprising, considering the fact that the Go/No-Go has been shown to load on an inattention and inhibition factor (Sharma et al., 2014), as compared to other cognitive control tasks that form a distinct decision-making and shifting factor. One possibility may be due to similar two-choice response modalities in the bandits to the flanker, as compared

to the single-choice modality in the Go/No-Go. Third, the inverse temperature in probabilistic bandit was related to both attentional and inhibitory control. This is consistent with cognitive control associating more to increased deployment of model-based RL algorithms (Gershman et al., 2014; Otto et al., 2015). Indeed, separate task associations between these constructs are sparse and the assortment of literature has often combined RL with other cognitive constructs, such as working memory or effort (Collins & Frank, 2012; Niv et al., 2015; Otto et al., 2013). Mixing processes within a single task has been shown to create a bundling of motivation processes, such as between negative punishment and reward (Yee et al., 2022). For example, combining RL with a Go/No-Go paradigm revealed a Pavlovian bias, wherein approach-to-reward tendencies interferes with No-Go responses and avoidance-from-punishment tendencies interferes with Go responses (Burghoorn et al., 2025; Guitart-Masip et al., 2012). Unifying cognitive processes within a single paradigm has been incentivized as to account for more naturalistic multi-process environments (Bailey et al., 2021; Botvinick et al., 2015; Otto et al., 2015), such as a measure of their contrast in a competitive environments or measure of their strength in cooperate environments. Fourth, we find a relationship between TD and CC drift rates which demonstrate that steeper discounters tend to be less accurate in control tasks. One study found support for a factor of drift rate across several CC tasks as a transdiagnostic predictor of mental health (Sripada & Weigard, 2021), and future work may benefit from examining diverse paradigms using the DDM as the choice rule (Ahn et al., 2020; Pedersen et al., 2017; Peters & D'Esposito, 2020).

We also found support for the prediction that response consistency across RL tasks (**H3**). Unexpectedly, response consistency in both bandits also predicted accuracy in delay and probability estimation. Often, weighing of probability and delay on reward show differing effects

(Frederick & Loewenstein, 2002; Green & Myerson, 2004), but RL research has often looked at probability and delay in a unified manner (Daw et al., 2011; Niv, 2009) without isolating the effects of delay. Some studies have attempted to isolate delay in RL, such as through fixed-delay reward presentation (Bruinsma et al., 2024; Maulhardt et al., 2026; Tanaka et al., 2009), structural transitions (Moran et al., 2019; Walsh & Anderson, 2011), and combined-with-probabilistic outcomes (Witkowski et al., 2025). The research presented here shows that participants who engage a heuristic strategy to estimate delay are also on average better at learning probabilistic rewards over time. We hope that future RL studies might employ more complicated, unified, and/or isolated probabilistic and delayed designs.

The final hypothesis was one that related the various behavioral tasks and self-report measures (**H4**), where we find mostly null or inconclusive relationships. However, one counterintuitive finding was the interaction between early-trial stay rate in the delayed bandit and the BIS-11 nonplanning and attentional factors. This finding is also inconsistent with Zou et al. (2022), who found that participants scoring higher on impulsivity increased their switching behavior after three consecutive unrewarded outcomes. Additionally, participants higher in impulsivity were more likely to randomly explore, or explore when it did not provide an information gain for future reward (Dubois & Hauser, 2022). While both studies appear to contrast with our findings, they are partially reconciled by evidence that steeper discounters also prefer less complex information (Shenhav et al., 2017). Others have found no relationship between exploration and BIS-11 self-report measures (Huang et al., 2024). These findings suggest that more impulsive participants may either strategically select or naturally default to this pseudo-optimal strategy when estimating delay.

One should be mindful of the context of delayed bandit and the BIS-11 finding. Specifically, we are examining only a small window of behavior, and the reduced trial count introduces a high potential for bias. Another caveat is that we did incentivize participants to learn the delay of each arm, thereby biasing their strategy toward estimation over pure reward accumulation. Furthermore, heightened learning from negative outcomes (Zou et al., 2022) along with heightened learning from positive outcomes (Poulton & Hester, 2020) presents conflicting evidence for a unified theory of impulse-related reward sensitivity. This is consistent with the inclusion of positive urgency in the UPPS framework (Cyders et al., 2007), suggesting that urgency across both positive and negative valence, with context likely playing a key moderating role (Eckstein et al., 2022; Okan & Hallquist, 2025). Another possible influence might be due to heightened degree of delay aversion (Bitsakou et al., 2009; T. V. Lim et al., 2025), which would incentivize impulsive participants to minimize the complexity of the solution. The list is of potential influences is by no means exhaustive, as participants may also be motivated to reduce the reward uncertainty (Cockburn et al., 2022; Maulhardt et al., 2026; O'Doherty et al., 2021). It is also possible that our questionnaire battery did not capture sufficient variance to distinguish between impulsivity and adjacent personality factors such as compulsivity, which has often been found to load with impulse-based questionnaires (Gillan et al., 2016; Hopkins et al., 2022). Overall, the most productive next steps would be to increase sample size for the HBM, conduct a replication study, or employ cross-validation.

This study did find support for a variety of interesting phenomena, but there are several limitations worth mention. First, as briefly explained earlier, replication has become a primary concern for behavioral tasks as low trial counts often result in lower reliability (Hedge et al., 2018; Rouder & Haaf, 2019). Given the time constraints of our task and overall quick decision-

making, the significant findings were encouraging, but we should be mindful that our split-half reliability measures are likely to overestimate the true reliability of computational model parameters. Second, the preregistered exclusion criteria were conservative, but more robust attentional checks might be of importance to identify lack of engagement in personality or behavioral metrics. Third, the delayed bandit task directed participants toward post-choice delay estimation, and participants largely adopted a heuristic stay strategy. More likely, participants used a variety of strategies. Finding a model that captures more behavior and strategy variance may be important for the future characterization of delay estimation and its relation to impulsivity. Fourth, modeling approaches would greatly benefit by explicitly incorporating a group-level covariance into the estimation of their participant parameters. The parameter coefficient is shrunk towards their associated group mean and may mute important individual variation for personality correlations. Unifying the majority of the effects into a single model allows estimates across tasks (Rouder & Haaf, 2019). For example, one expansion could capture a participants' engagement and fatigue trajectories over the course of the experiment (Pedersen et al., 2017).

We found strong support for the exploratory hypotheses concerning cross-modality associations, cross-parameter correlations, and replication of classic manipulation effects. While most of these were to validate our framework here, there were some interesting commonalities. Discounting rate appeared to relate to all drift and inverse temperature parameters, except for the delayed bandit inverse temperature for stay rate. This is surprising because instead of the inverse-temperature having cross-task correlations, we found discounting rate to be more related to cross-task evidence accumulation. More specifically, hyperbolic scaling of the value of the

competing option, as compared to linear scaling of the value difference at decision time, was weakly related to more accurate evidence accumulation across tasks.

In summary, we utilized an assortment of tasks to find various degrees of overlap while implementing more advanced modeling techniques for better dissociation, reduced error, and heightened sensitivity. Impulsivity is a phenomenon with a broad literature across an array of disciplines and its dimensions have shown promise for being transdiagnostic markers of mental health (Bickel et al., 2012; Gillan et al., 2016; Hoyniak & Petersen, 2019). While our study does not fully define or clarify the multidimensional construct of impulsivity, our methodologies and analyses should pave the way for greater integration across choice, action, learning, and personality dimensions of impulsivity. The observed relationships between RL and impulsivity may illustrate that impulsivity is both multidimensional and warrants consideration in the context of sequential behaviors. That is, impulsivity (rapid and unplanned response) may occur at a specific timepoint when moving towards a goal, and future work will need to clarify how these dimensions link sequentially. For example, impulse behavior may be found when a participant wants a delayed reward (TD) but must inhibit some allure (CC) sequentially over the course of time (RL) to achieve that reward. Impulsivity research that utilizes RL in their task batteries should begin to consider how these dimensions might manifest over longer time windows of sequential behaviors. By linking these computational parameters across behavioral and self-report domains, this work contributes to greater integration of dimensional frameworks (Insel et al., 2010; Kotov et al., 2017; Montague et al., 2012) for the classification and diagnoses of mental health outcomes.

Appendix A. Supplemental Information

Mixed-Effects Regressions.

We tested calculated p-values for these mixed-effects models using the 'mixed' function in the 'afex' package (Luke, 2017), this function reduces inflation of a type I error by performing a likelihood ratio test for generalized mixed effects models.

HBM Details.

Each task made use of both accuracy and RT through a DDM that sequential samples evidence with a probabilistic gaussian accumulator until that accumulation hits a specific decision bound. Although DDM have typically been used with perceptual decision making (Ratcliff et al., 2018; Ratcliff & McKoon, 2008), they have also been used across a spectrum of tasks (Letkiewicz et al., 2024; Peters & D'Esposito, 2020; Solway, Lin, & Kaplan, 2021). Additionally, building the DDM with a hierarchical framework (HDDM) provides robust qualities to our estimates (Ahn et al., 2011; Brown et al., 2020; Wiecki et al., 2013). As such, the DDM has four possible free parameters, which are then fit at both individual and group levels within this hierarchical framework. Drift rate (slope of the random walk), boundary separation (decision threshold), and starting point bias (initial start) fit with participants' parameters being sampled from a group-level mean. These individual parameters per model task were parameterized with a covariance matrix that explicitly models the relationship across and within each model task. Non-decision time was also hierarchical but not included in the covariance matrix to prevent inflated R-hat values. To make claims on our covariance matrix, we first converted our matrix into a standardized correlation matrix. The correlation were then subject to a posterior probability of direction by computing the proportion of posterior samples that was greater or less than zero, $p(r < 0)$ or $p(r > 0)$ (Kruschke & Liddell, 2018). This directional index is strongly correlated with typical frequentist p-values (Makowski et al., 2019). To make claims on evidence we looked towards our posterior means, 95% credit intervals, and probability of directionality for correlations (Kruschke & Liddell, 2018). We used typical evidence strength conventions established for Bayes Factors and their translation to posterior probabilities (Jeffreys, 1998; Kass & Raftery, 1995).

Drift Diffusion Modeling, Priors, and Transformation. Furthermore we decided against fitting any inter-variability parameters, as we wanted to use linear combinations to partition the effects of our experimental designs (Boehm et al., 2018). Due to partial pooling, each participant helps to inform the group-level parameter while shrinking estimates towards their group means allowing fitting procedures to use a lower trial count along with preventing high leverage from outliers (Boehm et al., 2018; Vandekerckhove et al., 2011; Wiecki et al., 2013). For priors, we used uninformative gaussian priors (Boehm et al., 2018; Tran et al., 2021), $N(0, 10)$. Some parameters had no transformation (drift rate), other transformations used an inverse logit for 0-to-1 bounds (starting-point bias and non-decision time) and finally a softplus transformation (boundary and inverse temperature). For a prior on the covariance matrix, we used the lower-triangle Cholesky decomposed correlation distribution, which was set to favor the null-hypothesis ($\eta = 2$).

Posterior Predictive Checks. We used the 'rdiffusion' function from the 'rtdists' package (Singmann et al., 2018) with the point-estimates of the HBM to simulate choice and RT. Due to filtering real data for estimates under 200ms, we also reran the simulation if the RT was below the same threshold. We then ran the simulated dataset through the same generalized multilevel model as participants and displayed the predicted marginal means alongside our participants' fit.

Hyperbolic Drift Diffusion Model.

There are more complicated implementations of the hyperbolic curve that we choose not to implement nor test if the exponential curve was a better fit to our data (Green & Myerson, 2004; McKerchar et al., 2009). Instead our goal was to use this value difference as the driver for our drift rate so that we could us a DDM as our choice-rule (Peters & D'Esposito, 2020; Rodriguez et al., 2015).

Participant-Level OLS Regressions.

Go/No-Go. *Accuracy,* $b_{nogo*reversal}$ = -.045, SE = .021, $t(834)$ = -2.1, p = .036, 95% CI [-.086, -.003], η^2_p = .005. b_{nogo} = -.105, SE = .015, $t(834)$ = -7.00, p < .001, 95% CI [-.134, -.075], η^2_p = .06. *RT,* $b_{reversal}$ = .013, SE = .006, $t(779)$ = 2.055, p = .04, 95% CI [.001<, .026], η^2_p = .005. *RT,* b_{nogo} = -.035, SE = .007, $t(779)$ = -5.26, p < .001, 95% CI [-.049, -.022], η^2_p = .03.

Flanker. *Accuracy*, $b_{incongruency}$ = -.21, SE = .017, $t(418)$ = -11.86, p < .001, 95% CI [-.24, -.17], η^2_p = .25. *RT*, $b_{incongruency}$ = .047, SE = .006, $t(418)$ = 8.52, p < .001, 95% CI [.04, .06], η^2_p = .15.

Delay Discounting. *Choice*, b_{linear} = -.46, SE = .024, $t(1254)$ = -19.45, p < .001, 95% CI [-.51, -.41]. $b_{quadratic}$ = .13, SE = .024, $t(1254)$ = 5.33, p < .001, 95% CI [.08, .17]. *RT*, b_{linear} = .115, SE = .033, $t(1254)$ = 3.43, p < .001, 95% CI [.05, .18]. $b_{quadratic}$ = -.14, SE = .033, $t(1254)$ = -4.2, p < .001, 95% CI [-.21, -.075].

Probabilistic Bandit. *Accuracy*, $b_{timing*reversal}$ = .066, SE = .021, $t(836)$ = 3.05, p = .002, 95% CI [.023, .108], η^2_p = .01. $b_{reversal}$ = -.183, SE = .015, $t(836)$ = -12.06, p < .001, 95% CI [-.213, -.153], η^2_p = .15. b_{timing} = .079, SE = .015, $t(836)$ = 5.2, p < .001, 95% CI [.049, .109], η^2_p = .03. *RT*, $b_{timing*reversal}$ = .049, SE = .017, $t(836)$ = 2.91, p = .004, 95% CI [.016, .081], η^2_p = .01. $b_{reversal}$ = -.055, SE = .012, $t(836)$ = -4.61, p < .001, 95% CI [-.078, -.031], η^2_p = .02. b_{timing} = -.059, SE = .012, $t(836)$ = -4.95, p < .001, 95% CI [-.082, -.035], η^2_p = .03.

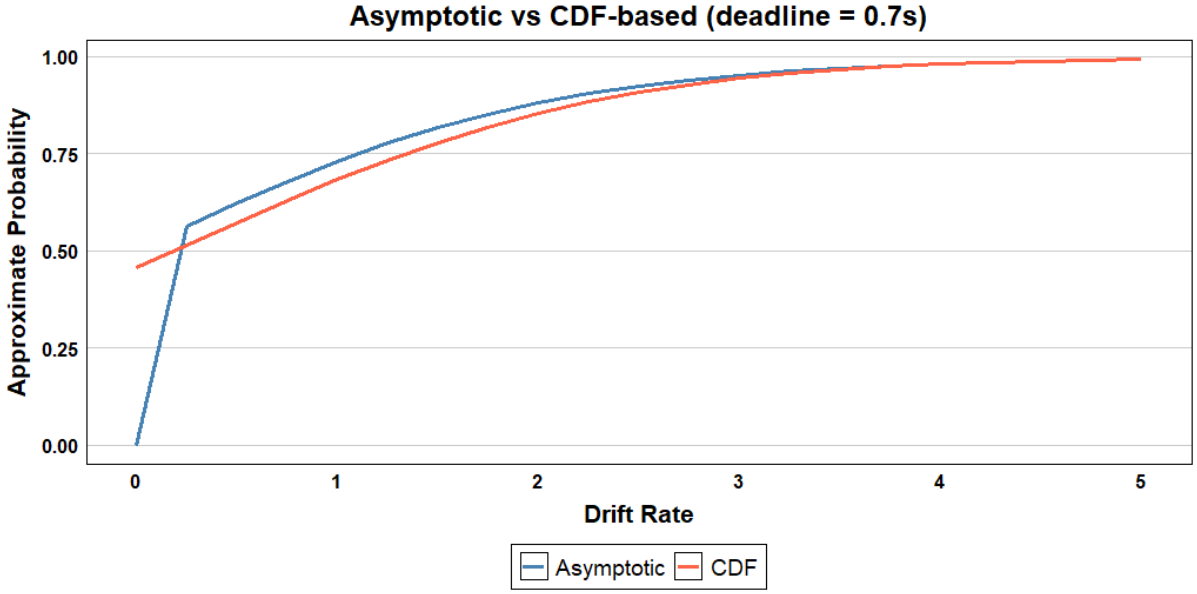
Delayed Bandit. *Estimation Error*, $b_{switch*arm}$ = .208, SE = .027, $t(3774)$ = 7.64, p < .001, 95% CI [.155, .262], η^2_p = .02. b_{switch} = .056, SE = .019, $t(3774)$ = 2.93, p = .003, 95% CI [.019, .094], η^2_p = .003, b_{arm} = .337, SE = .107, $t(3774)$ = 3.164, p = .002, 95% CI [.128, .546], η^2_p = .002, *RT*, b_{switch} = -.012, SE = .004, $t(3774)$ = -3.063, p = .002, 95% CI [-.019, -.004], η^2_p = .002.

BIS-11 Factors and Lack of Engagement Analysis. First, one might assume that higher intercept (baseline stay rate) and lower slope (sensitivity to unrewarded outcomes) would materialize as better delay estimation per our previous analysis. To test this validation, we took the random intercept and slope from each model to predict delay estimation. Overall, most effects where similar so reported values may have small variation. All random intercepts were significant, p's < .005, indicating better delay estimation per higher baseline stay rate. Additionally, some random slopes were either marginal or significant, p's = .05~, illustrating that as unrewarded trials are accrued estimation becomes more difficult. Second, one might assume a problematic cluster of individuals which do not switch in the block at all. This cluster could either be indicative of lack of engagement or an over reliance on the heuristic strategy to reduce the learning problem. To control for this, we calculated each participant's proportion of no-switches across their 9 blocks and performed bivariate spearman correlations on each self-report factor. If a correlation was observed, then the factor's degree of influence on stay rate may be an artifact of lack of engagement or over reliance on the heuristic strategy. While none of the BIS-11 factors were significant: nonplanning, r = .075, p = .28, motor, r = .031, p = .65, attentional, r = .06, p = .41, there was significance with OCDS: obsessive, r = .17, p = .01, compulsive, r = .15, p = .03, along with significance with BSSS: total, r = .15, p = .03. This does provide some

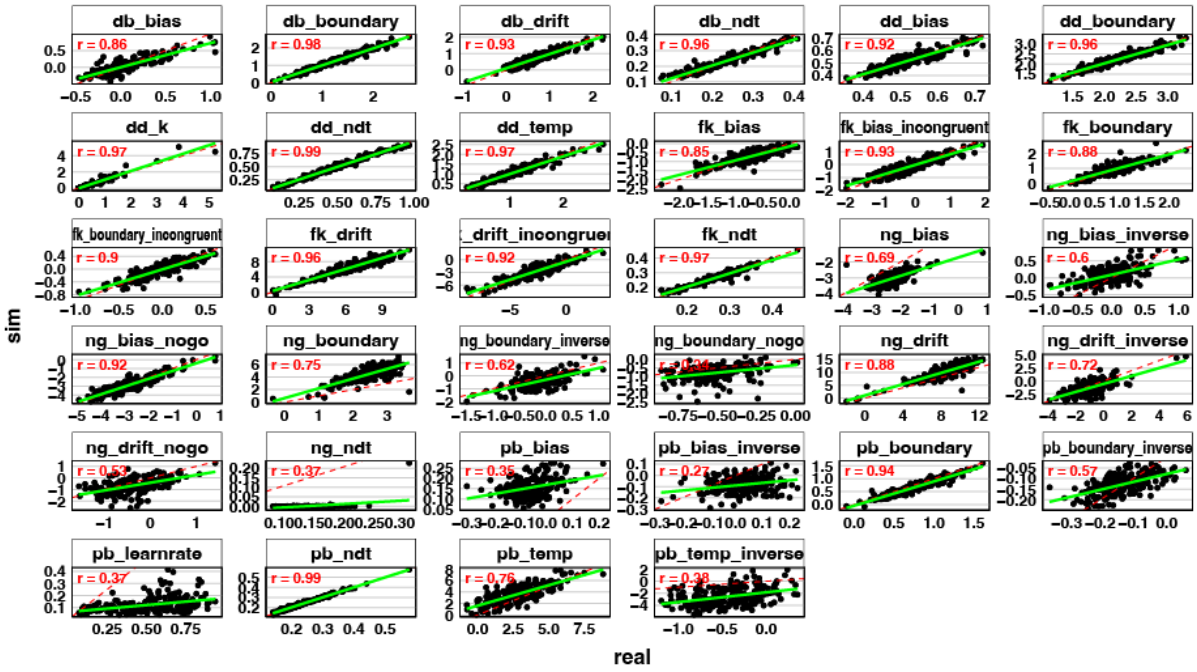
clarity that BIS-11 interaction effects seem to not be driven by lack of engagement or overreliance on heuristic strategy.

Appendix B. Supplemental Plots

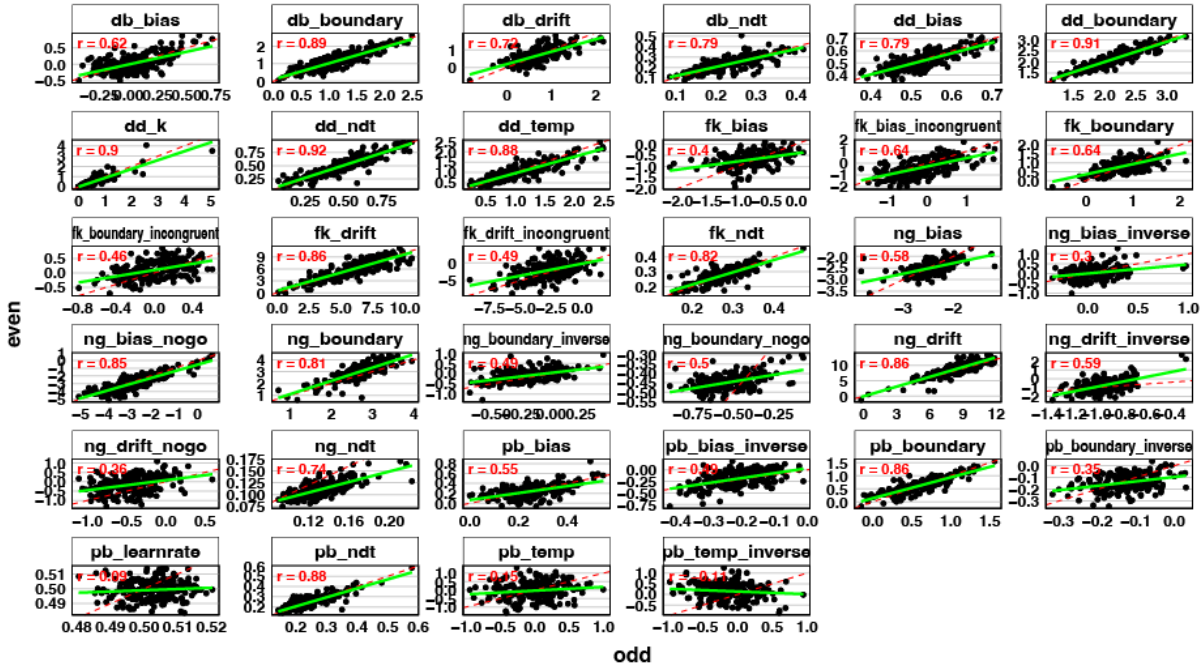
S1. Comparison between censored probability approximations. Two approximations of integrating the area under the curve using the asymptotic first-passage (Wiecki et al., 2013) and cumulative density function (Henrich & Klauer, 2026) approximations.



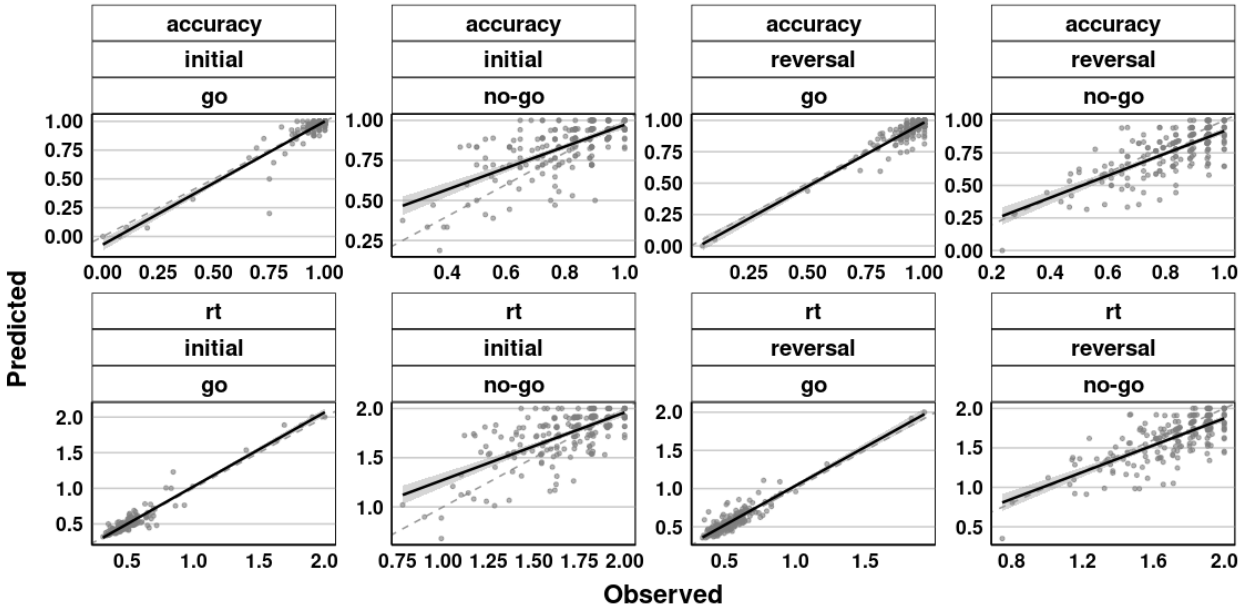
S2. Parameter recovery for all parameters. Reference line (red-dotted) and linear fit to real and simulated data (green-solid).



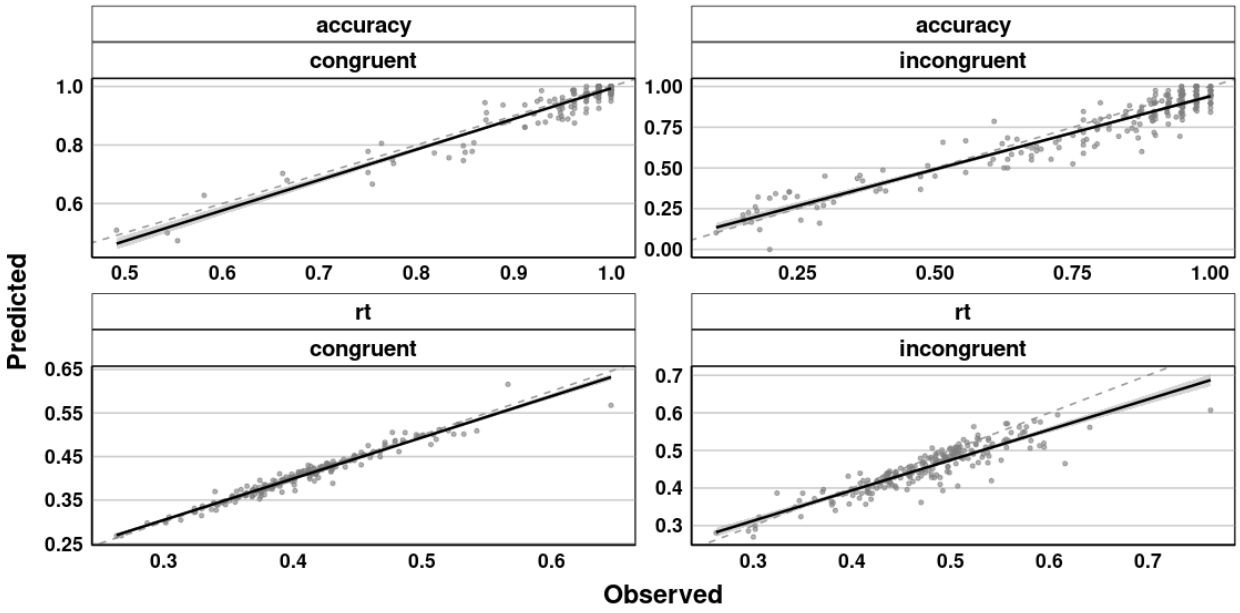
S3. Parameter split half reliability for all parameters. Reference line (red-dotted) and linear fit to even and odd trial-data (green-solid). Caution as some parameters do not make sense in split half estimation (learning rate and inverse temperature).



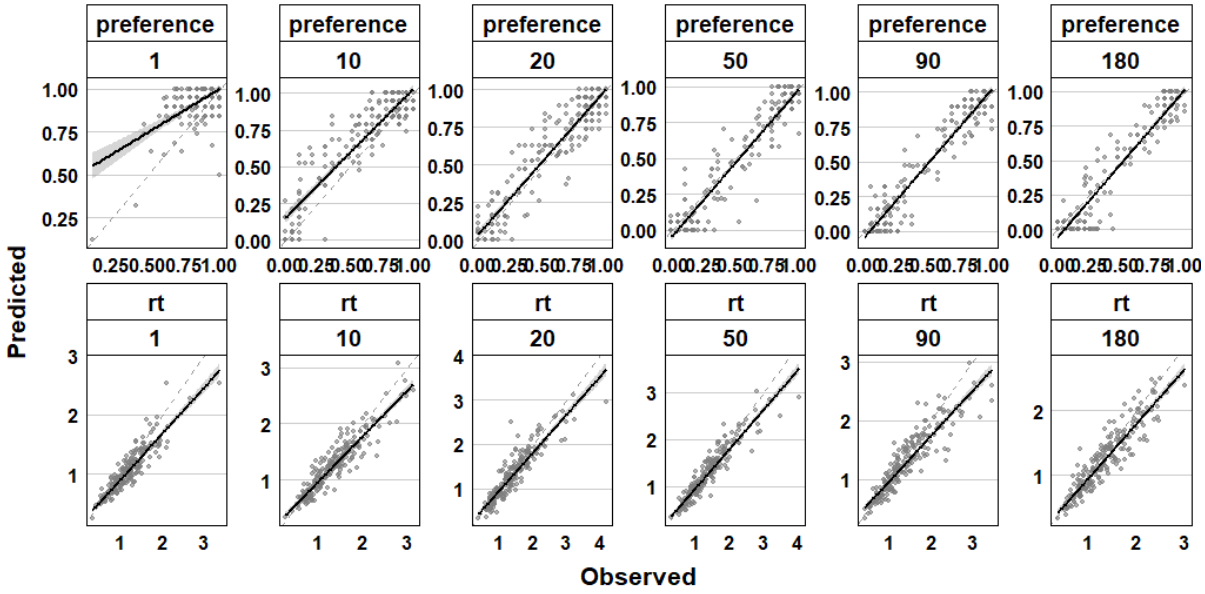
S4. Go/No-Go observed and predicted correlation of behavioral metrics. Facet names refer to conditions of being in a Go or No-Go trial and initial before and after target letter reversal.



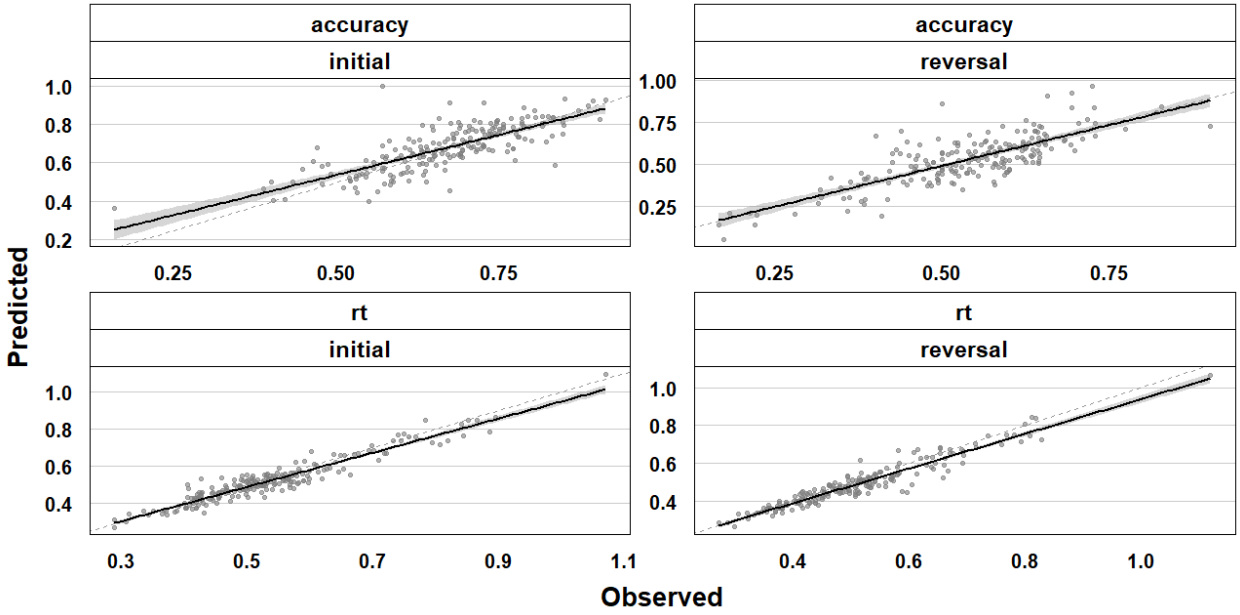
S5. Flanker observed and predicted correlation of behavioral metrics. Facet terms refer to condition of congruent and neutral (collapsed into congruent) and incongruent.



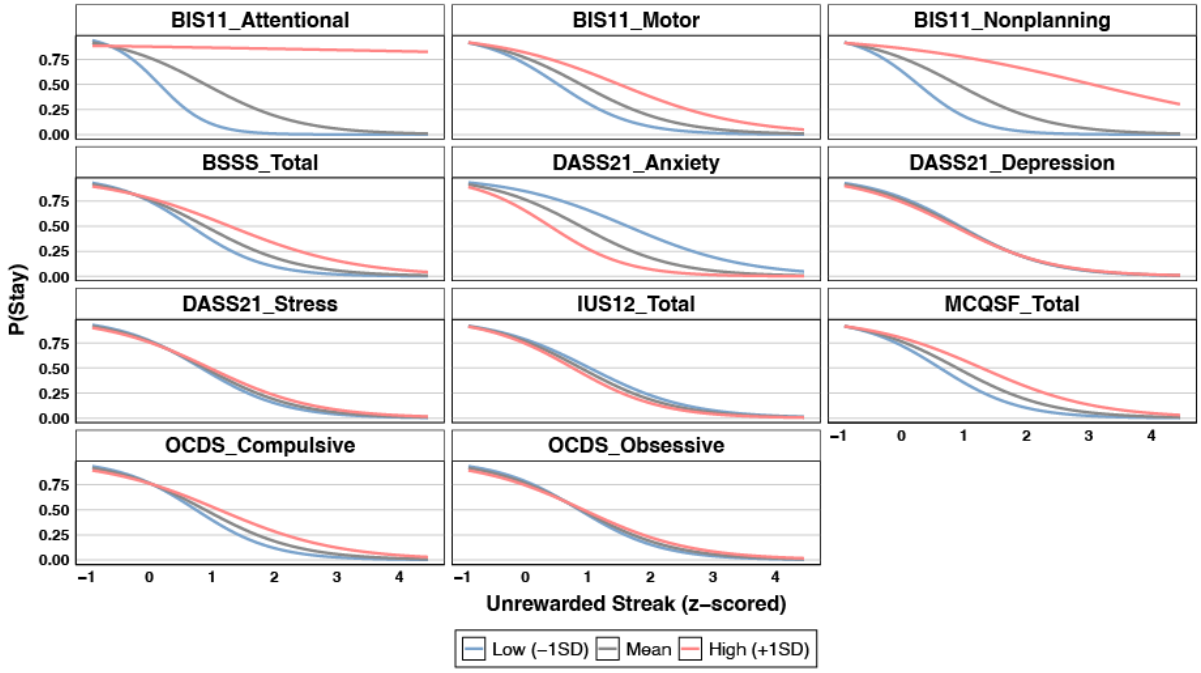
S6. Delay discounting observed and predicted correlation of behavioral metrics. Facet numbers (1, 10, 20, 50, 90, 180) refer to delay in number of days before receiving delayed reward.



S7. Probabilistic bandit observed and predicted correlation of behavioral metrics. Facet terms refer to initial block before and after reversal.



S8. Interaction effects (factor * unrewarded streak) from individually fit logistic regressions predicting stay behavior on the delayed bandit task.



References

Ahn, W.-Y., Gu, H., Shen, Y., Haines, N., Hahn, H. A., Teater, J. E., Myung, J. I., & Pitt, M. A. (2020). Rapid, precise, and reliable measurement of delay discounting using a Bayesian learning algorithm. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-68587-x

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 95–110. https://doi.org/10.1037/a0020684

Akaishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue Environment. *Journal of Neuroscience*, *36*(4), 1096–1112. https://doi.org/10.1523/JNEUROSCI.3159-15.2016

American Psychiatric Association, A. (2022). *Diagnostic and statistical manual of mental disorders.* American Psychiatric Association Publishing.

Ames, S. L., Wong, S. W., Bechara, A., Cappelli, C., Dust, M., Grenard, J. L., & Stacy, A. W. (2014). Neural correlates of a Go/NoGo task with alcohol stimuli in light and heavy young drinkers. *Behavioural Brain Research*, *274*, 382–389. https://doi.org/10.1016/j.bbr.2014.08.039

Amlung, M., Marsden, E., Holshausen, K., Morris, V., Patel, H., Vedelago, L., Naish, K. R., Reed, D. D., & McCabe, R. E. (2019). Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. *JAMA Psychiatry*, *76*(11), 1176. https://doi.org/10.1001/jamapsychiatry.2019.2102

Anton, R. F., Moak, D. H., & Latham, P. (1995). The Obsessive Compulsive Drinking Scale: A self-rated instrument for the quantification of thoughts about alcohol and drinking behavior. *Alcoholism: Clinical and Experimental Research*, *19*(1), 92–99.

Asaad, W. F., Lauro, P. M., Perge, J. A., & Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment Problem. *The Journal of Neuroscience*, *37*(29), 6995–7007. https://doi.org/10.1523/JNEUROSCI.3311-16.2017

Bailey, A. J., Romeu, R. J., & Finn, P. R. (2021). The problems with delay discounting: A critical review of current practices and clinical applications. *Psychological Medicine*, *51*(11), 1799–1806. https://doi.org/10.1017/S0033291721002282

Banca, P., Vestergaard, M. D., Rankov, V., Baek, K., Mitchell, S., Lapa, T., Castelo-Branco, M., & Voon, V. (2015). Evidence Accumulation in Obsessive-Compulsive Disorder: The Role of Uncertainty and Monetary Reward on Perceptual Decision-Making Thresholds. *Neuropsychopharmacology*, *40*(5), 1192–1202. https://doi.org/10.1038/npp.2014.303

Bargh, J. A., & Chartrand, T. L. (1999). The Unbearable Automaticity of Being. *American Psychologist*, 18.

Barratt, E. S. (1959). Anxiety and impulsiveness related to psychomotor efficiency. *Perceptual and Motor Skills*, *9*(3), 191–198.

Barratt, E. S. (1983). Impulsivity: Cognitive, behavioral, and psychophysiological correlates. *Biological Bases of Sensation Seeking, Impulsivity, and Anxiety*.

Baumann, A. A., & Odum, A. L. (2012). Impulsivity, risk taking, and timing. *Behavioural Processes*, *90*(3), 408–414. https://doi.org/10.1016/j.beproc.2012.04.005

Berg, J. M., Latzman, R. D., Bliwise, N. G., & Lilienfeld, S. O. (2015). Parsing the heterogeneity of impulsivity: A meta-analytic review of the behavioral implications of the UPPS for

psychopathology. *Psychological Assessment*, *27*(4), 1129–1146. https://doi.org/10.1037/pas0000111

Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Koffarnus, M. N., & Gatchalian, K. M. (2012). Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: Emerging evidence. *Pharmacology & Therapeutics*, *134*(3), 287–297. https://doi.org/10.1016/j.pharmthera.2012.02.004

Bickel, W. K., Koffarnus, M. N., Moody, L., & Wilson, A. G. (2014). The behavioral- and neuro-economic process of temporal discounting: A candidate behavioral marker of addiction. *Neuropharmacology*, *76*, 518–527. https://doi.org/10.1016/j.neuropharm.2013.06.013

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of Abbreviated Nine-Item Forms of the Raven's Standard Progressive Matrices Test. *Assessment*, *19*(3), 354–369. https://doi.org/10.1177/1073191112446655

Bitsakou, P., Psychogiou, L., Thompson, M., & Sonuga-Barke, E. J. S. (2009). Delay Aversion in Attention Deficit/Hyperactivity Disorder: An empirical investigation of the broader phenotype. *Neuropsychologia*, *47*(2), 446–456. https://doi.org/10.1016/j.neuropsychologia.2008.09.015

Boehm, U., Annis, J., Frank, M. J., Hawkins, G. E., Heathcote, A., Kellen, D., Krypotos, A.-M., Lerche, V., Logan, G. D., Palmeri, T. J., Van Ravenzwaaij, D., Servant, M., Singmann, H., Starns, J. J., Voss, A., Wiecki, T. V., Matzke, D., & Wagenmakers, E.-J. (2018). Estimating across-trial variability parameters of the Diffusion Decision Model: Expert

advice and recommendations. *Journal of Mathematical Psychology*, *87*, 46–75. https://doi.org/10.1016/j.jmp.2018.09.004

Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624.

Botvinick, M., Weinstein, A., Solway, A., & Barto, A. (2015). Reinforcement learning, efficient coding, and the statistics of natural tasks. *Current Opinion in Behavioral Sciences*, *5*, 71–77. https://doi.org/10.1016/j.cobeha.2015.08.009

Broos, N., Schmaal, L., Wiskerke, J., Kostelijk, L., Lam, T., Stoop, N., Weierink, L., Ham, J., De Geus, E. J. C., Schoffelmeer, A. N. M., Van Den Brink, W., Veltman, D. J., De Vries, T. J., Pattij, T., & Goudriaan, A. E. (2012). The Relationship between Impulsive Choice and Impulsive Action: A Cross-Species Translational Study. *PLoS ONE*, *7*(5), e36781. https://doi.org/10.1371/journal.pone.0036781

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

Bruinsma, S., Petzschner, F., & Nassar, M. (2024). Perseverative Behavioral Sequences Aid Long-Term Credit Assignment. *Conference on Cognitive Computational Neuroscience*.

Burghoorn, F., Scheres, A., Monterosso, J., Guo, M., Luo, S., Roelofs, K., & Figner, B. (2025). Pavlovian impatience: The anticipation of immediate rewards increases approach behaviour. *Cognitive, Affective, & Behavioral Neuroscience*, *25*(2), 358–376. https://doi.org/10.3758/s13415-024-01236-2

Carragher, N., Krueger, R. F., Eaton, N. R., & Slade, T. (2015). Disorders without borders: Current and future directions in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology*, *50*(3), 339–350. https://doi.org/10.1007/s00127-014-1004-z

Casey, B. J., Somerville, L. H., Gotlib, I. H., Ayduk, O., Franklin, N. T., Askren, M. K., Jonides, J., Berman, M. G., Wilson, N. L., Teslovich, T., Glover, G., Zayas, V., Mischel, W., & Shoda, Y. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences*, *108*(36), 14998–15003. https://doi.org/10.1073/pnas.1108561108

Chen, S., Yao, N., Qian, M., & Lin, M. (2016). Attentional biases in high social anxiety using a flanker task. *Journal of Behavior Therapy and Experimental Psychiatry*, *51*, 27–34. https://doi.org/10.1016/j.jbtep.2015.12.002

Cockburn, J., Ding, W., Oh, S., Charpentier, C. J., Feusner, J. D., Tadayonnejad, R., & O'Doherty, J. (2024). *Characterizing heterogeneity in human reinforcement learning and the arbitration of behavioral control*. PsyArXiv. https://doi.org/10.31234/osf.io/egq3x

Cockburn, J., Man, V., Cunningham, W. A., & O'Doherty, J. P. (2022). Novelty and uncertainty regulate the balance between exploration and exploitation through distinct mechanisms in the human brain. *Neuron*, *110*(16), 2691-2702.e8. https://doi.org/10.1016/j.neuron.2022.05.025

Collins, A. G. E., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *The Journal of Neuroscience*, *34*(41), 13747–13756. https://doi.org/10.1523/JNEUROSCI.0989-14.2014

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis: Working memory in reinforcement learning. *European Journal of Neuroscience*, *35*(7), 1024–1035. https://doi.org/10.1111/j.1460-9568.2011.07980.x

Cosenza, M., & Nigro, G. (2015). Wagering the future: Cognitive distortions, impulsivity, delay discounting, and time perspective in adolescent gambling. *Journal of Adolescence*, *45*(1), 56–66. https://doi.org/10.1016/j.adolescence.2015.08.015

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*(1), 1–8.

Cyders, M. A., & Coskunpinar, A. (2011). Measurement of constructs using self-report and behavioral lab tasks: Is there overlap in nomothetic span and construct representation for impulsivity? *Clinical Psychology Review*, *31*(6), 965–982. https://doi.org/10.1016/j.cpr.2011.06.001

Cyders, M. A., Smith, G. T., Spillane, N. S., Fischer, S., Annus, A. M., & Peterson, C. (2007). Integration of impulsivity and positive mood to predict risky behavior: Development and validation of a measure of positive urgency. *Psychological Assessment*, *19*(1), 107–118. https://doi.org/10.1037/1040-3590.19.1.107

Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2011). Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron*, *69*(4), 680–694. https://doi.org/10.1016/j.neuron.2011.01.020

Dalley, J. W., & Robbins, T. W. (2017). Fractionating impulsivity: Neuropsychiatric implications. *Nature Reviews Neuroscience*, *18*(3), 158–171. https://doi.org/10.1038/nrn.2017.8

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879. https://doi.org/10.1038/nature04766

Dillon, D. G., Wiecki, T., Pechtel, P., Webb, C., Goer, F., Murray, L., Trivedi, M., Fava, M., McGrath, P. J., Weissman, M., Parsey, R., Kurian, B., Adams, P., Carmody, T., Weyandt, S., Shores-Wilson, K., Toups, M., McInnis, M., Oquendo, M. A., … Pizzagalli, D. A. (2015). A computational analysis of flanker interference in depression. *Psychological Medicine*, *45*(11), 2333–2344. https://doi.org/10.1017/S0033291715000276

Dubois, M., & Hauser, T. U. (2022). Value-free random exploration is linked to impulsivity. *Nature Communications*, *13*(1), 4542. https://doi.org/10.1038/s41467-022-31918-9

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality*, *45*(3), 259–268. https://doi.org/10.1016/j.jrp.2011.02.004

Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. (2022). The interpretation of computational model parameters depends on the context. *eLife*, *11*, e75474. https://doi.org/10.7554/eLife.75474

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, *116*(12), 5472–5477. https://doi.org/10.1073/pnas.1818430116

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241.

Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology*, *146*(4), 348–361. https://doi.org/10.1007/PL00005481

Falk, A., Kosse, F., & Pinger, P. (2020). Re-revisiting the marshmallow test: A direct comparison of studies by Shoda, Mischel, and Peake (1990) and Watts, Duncan, and Quan (2018). *Psychological Science*, *31*(1), 100–104.

Feher Da Silva, C., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*, *4*(10), 1053–1066. https://doi.org/10.1038/s41562-020-0905-y

Fernández, B. P., & Bravo, A. P. (2003). Impulsivity: Historical and conceptual review. *Actas Espanolas de Psiquiatria*, *31*(4), 220–230.

Finn, P. R., Justus, A., Mazas, C., & Steinmetz, J. E. (1999). Working memory, executive processes and the effects of alcohol on Go/No-Go learning: Testing a model of behavioral regulation and impulsivity. *Psychopharmacology*, *146*(4), 465–472. https://doi.org/10.1007/PL00005492

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), 300–326. https://doi.org/10.1037/0033-295X.113.2.300

Franken, I. H. A., Van Strien, J. W., Nijs, I., & Muris, P. (2008). Impulsivity is associated with behavioral decision-making deficits. *Psychiatry Research*, *158*(2), 155–163. https://doi.org/10.1016/j.psychres.2007.06.002

Frederick, S., & Loewenstein, G. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, 96.

Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*, *204*, 104394. https://doi.org/10.1016/j.cognition.2020.104394

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). Retrospective revaluation in sequential decision making: A tale of two systems. *Journal of Experimental Psychology: General*, *143*(1), 182–194. https://doi.org/10.1037/a0030844

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, *5*, e11305. https://doi.org/10.7554/eLife.11305

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. https://doi.org/10.1016/j.neuron.2010.04.016

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389–413. https://doi.org/10.1037/0096-3445.136.3.389

Gondan, M., Blurton, S. P., & Kesselmeier, M. (2014). Even faster and even more accurate first-passage time densities and distributions for the Wiener diffusion model. *Journal of Mathematical Psychology*, *60*, 20–22. https://doi.org/10.1016/j.jmp.2014.05.002

Green, L., & Myerson, J. (2004). A Discounting Framework for Choice With Delayed and Probabilistic Rewards. *Psychological Bulletin*, *130*(5), 769–792. https://doi.org/10.1037/0033-2909.130.5.769

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, *62*(1), 154–166. https://doi.org/10.1016/j.neuroimage.2012.04.024

Gureckis, T. M., & Love, B. C. (2009). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition*, *113*(3), 293–313. https://doi.org/10.1016/j.cognition.2009.03.013

Haines, N., Beauchaine, T. P., Galdo, M., Rogers, A. H., Hahn, H., Pitt, M. A., Myung, J. I., Turner, B. M., & Ahn, W.-Y. (2020). Anxiety Modulates Preference for Immediate Rewards Among Trait-Impulsive Individuals: A Hierarchical Bayesian Analysis. *Clinical Psychological Science*, *8*(6), 1017–1036. https://doi.org/10.1177/2167702620929636

Hall, N. T., Schreiber, A. M., Allen, T. A., & Hallquist, M. N. (2021). Disentangling cognitive processes in externalizing psychopathology using drift diffusion modeling: Antagonism, but not disinhibition, is associated with poor cognitive control. *Journal of Personality*, *89*(5), 970–985. https://doi.org/10.1111/jopy.12628

Hamilton, K. R., Mitchell, M. R., Wing, V. C., Balodis, I. M., Bickel, W. K., Fillmore, M., Lane, S. D., Lejuez, C. W., Littlefield, A. K., Luijten, M., Mathias, C. W., Mitchell, S. H., Napier, T. C., Reynolds, B., Schütz, C. G., Setlow, B., Sher, K. J., Swann, A. C.,

Tedford, S. E., … Moeller, F. G. (2015). Choice impulsivity: Definitions, measurement issues, and clinical implications. *Personality Disorders: Theory, Research, and Treatment*, *6*(2), 182–198. https://doi.org/10.1037/per0000099

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But…. *Communication Methods and Measures*, *14*(1), 1–24.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Heishman, S. J., Evans, R. J., Singleton, E. G., Levin, K. H., Copersino, M. L., & Gorelick, D. A. (2009). Reliability and validity of a short form of the Marijuana Craving Questionnaire. *Drug and Alcohol Dependence*, *102*(1–3), 35–40. https://doi.org/10.1016/j.drugalcdep.2008.12.010

Henrich, F., & Klauer, K. C. (2026). Modeling truncated and censored data with the diffusion model in Stan. *Behavior Research Methods*, *58*(2), 42. https://doi.org/10.3758/s13428-025-02822-z

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, *44*(2), 227–239. https://doi.org/10.1348/014466505X29657

Hopkins, A. K., Gillan, C., Roiser, J. P., Wise, T., & Sidarus, N. (2022). *Optimising the measurement of anxious-depressive, compulsivity and intrusive thought and social withdrawal transdiagnostic symptom dimensions*. https://doi.org/10.31234/osf.io/q83sh

Horan, W. P., Johnson, M. W., & Green, M. F. (2017). Altered experiential, but not hypothetical, delay discounting in schizophrenia. *Journal of Abnormal Psychology*, *126*(3), 301–311. https://doi.org/10.1037/abn0000249

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences*, *32*(3), 401–414. https://doi.org/10.1016/S0191-8869(01)00032-0

Hoyniak, C. P., & Petersen, I. T. (2019). A meta-analytic evaluation of the N2 component as an endophenotype of response inhibition and externalizing psychopathology in childhood. *Neuroscience & Biobehavioral Reviews*, *103*, 200–215. https://doi.org/10.1016/j.neubiorev.2019.06.011

Huang, Y., Luan, S., Wu, B., Li, Y., Wu, J., Chen, W., & Hertwig, R. (2024). Impulsivity is a stable, measurable, and predictive psychological trait. *Proceedings of the National Academy of Sciences*, *121*(24), e2321758121. https://doi.org/10.1073/pnas.2321758121

Huang-Pollock, C., Ratcliff, R., McKoon, G., Shapiro, Z., Weigard, A., & Galloway-Long, H. (2017). Using the Diffusion Model to Explain Cognitive Deficits in Attention Deficit Hyperactivity Disorder. *Journal of Abnormal Child Psychology*, *45*(1), 57–68. https://doi.org/10.1007/s10802-016-0151-y

Hunter, L. E., Bornstein, A. M., & Hartley, C. A. (2018). *A common deliberative process underlies model-based planning and patient intertemporal choice* [Preprint]. Animal Behavior and Cognition. https://doi.org/10.1101/499707

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research Domain Criteria (RDoC): Toward a New Classification Framework

for Research on Mental Disorders. *American Journal of Psychiatry*, *167*(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

International Society for Research on Impulsivity. (2019). https://www.impulsivity.org/

Jackson, J. N. S., & MacKillop, J. (2016). Attention-Deficit/Hyperactivity Disorder and Monetary Delay Discounting: A Meta-Analysis of Case-Control Studies. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*(4), 316–325. https://doi.org/10.1016/j.bpsc.2016.01.007

James, W. (1890). *The principles of psychology* (Vol. 1, Issue 2). Macmillan London.

Jeffreys, H. (1998). *The theory of probability*. OuP Oxford.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625–1633. https://doi.org/10.1038/nn2007

Kable, J. W., & Glimcher, P. W. (2010). An "As Soon As Possible" Effect in Human Intertemporal Decision Making: Behavioral Evidence and Neural Mechanisms. *Journal of Neurophysiology*, *103*(5), 2513–2531. https://doi.org/10.1152/jn.00177.2009

Kahneman, D. (2011). Thinking, fast and slow. *Farrar, Straus and Giroux*.

Kanen, J. W., Ersche, K. D., Fineberg, N. A., Robbins, T. W., & Cardinal, R. N. (2019). Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: Remediating effects of dopaminergic D2/3 receptor agents. *Psychopharmacology*, *236*(8), 2337–2358. https://doi.org/10.1007/s00213-019-05325-w

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Keidel, K., Lu, X., Suzuki, S., Murawski, C., & Ettinger, U. (2024). Association of temporal discounting with transdiagnostic symptom dimensions. *Npj Mental Health Research*, *3*(1), 13. https://doi.org/10.1038/s44184-024-00060-3

Kim, M., & Seo, J. (2021). Impulsivity is related to overhasty risk learning in attention-deficit/hyperactivity disorder: A computational psychiatric approach. *Journal of Psychiatric Research*, *143*, 84–90. https://doi.org/10.1016/j.jpsychires.2021.07.044

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, *126*(1), 54.

Kobayashi, S., & Schultz, W. (2008). Influence of Reward Delays on Responses of Dopamine Neurons. *Journal of Neuroscience*, *28*(31), 7837–7846. https://doi.org/10.1523/JNEUROSCI.1600-08.2008

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., & Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454.

Krueger, R. F., Skodol, A. E., Livesley, W. J., Shrout, P. E., & Huang, Y. (2007). Synthesizing dimensional and categorical approaches to personality disorders: Refining the research agenda for DSM-V Axis II. *International Journal of Methods in Psychiatric Research*, *16*(S1), S65–S73. https://doi.org/10.1002/mpr.212

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

Kvam, P. D., Romeu, R. J., Turner, B. M., Vassileva, J., & Busemeyer, J. R. (2021). Testing the factor structure underlying behavior using joint cognitive models: Impulsivity in delay discounting and Cambridge gambling tasks. *Psychological Methods*, *26*(1), 18–37. https://doi.org/10.1037/met0000264

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478. https://doi.org/10.1162/003355397555253

Lee, J. K., Rouault, M., & Wyart, V. (2023). *Compulsivity is linked to maladaptive choice variability but unaltered reinforcement learning under uncertainty*. https://doi.org/10.1101/2023.01.05.522867

Lei, Y., & Solway, A. (2022). Conflict and competition between model-based and model-free control. *PLOS Computational Biology*, *18*(5), e1010047. https://doi.org/10.1371/journal.pcbi.1010047

Lempert, K. M., Steinglass, J. E., Pinto, A., Kable, J. W., & Simpson, H. B. (2019). Can delay discounting deliver on the promise of RDoC? *Psychological Medicine*, *49*(2), 190–199. https://doi.org/10.1017/S0033291718001770

Letkiewicz, A. M., Wakschlag, L. S., Briggs-Gowan, M. J., Cochran, A. L., Wang, L., Norton, E. S., & Shankman, S. A. (2024). Preadolescent externalizing and internalizing symptoms are differentially related to drift-diffusion model parameters and neural activation during a go/no-go task. *Journal of Psychiatric Research*, *178*, 405–413. https://doi.org/10.1016/j.jpsychires.2024.08.038

Levitt, E. E., Oshri, A., Amlung, M., Ray, L. A., Sanchez-Roige, S., Palmer, A. A., & MacKillop, J. (2023). Evaluation of delay discounting as a transdiagnostic research

domain criteria indicator in 1388 general community adults. *Psychological Medicine*, *53*(4), 1649–1657. https://doi.org/10.1017/S0033291721005110

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy*, *62*, 129–139. https://doi.org/10.1016/j.brat.2014.07.019

Lim, M. S., Jocham, G., Hunt, L. T., Behrens, T. E., & Rogers, R. D. (2015). Impulsivity and predictive control are associated with suboptimal action-selection and action-value learning in regular gamblers. *International Gambling Studies*, *15*(3), 489–505.

Lim, T. V., Cardinal, R. N., Ziauddeen, H., Regenthal, R., Sahakian, B. J., Robbins, T. W., & Ersche, K. D. (2025). Atomoxetine reduces decisional impulsivity in human cocaine addiction. *Biological Psychiatry*, *97*(6), 627–636.

Lorains, F. K., Dowling, N. A., Enticott, P. G., Bradshaw, J. L., Trueblood, J. S., & Stout, J. C. (2014). Strategic and non-strategic problem gamblers differ on decision-making under risk and ambiguity. *Addiction*, *109*(7), 1128–1137. https://doi.org/10.1111/add.12494

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*, 1494–1502.

Luria, A. R. (1973). The frontal lobes and the regulation of behavior. In *Psychophysiology of the frontal lobes* (pp. 3–26). Elsevier.

MacKillop, J., Amlung, M. T., Few, L. R., Ray, L. A., Sweet, L. H., & Munafò, M. R. (2011). Delayed reward discounting and addictive behavior: A meta-analysis. *Psychopharmacology*, *216*(3), 305–321. https://doi.org/10.1007/s00213-011-2229-0

MacKillop, J., Weafer, J., C. Gray, J., Oshri, A., Palmer, A., & De Wit, H. (2016). The latent structure of impulsivity: Impulsive choice, impulsive action, and impulsive personality

traits. *Psychopharmacology*, *233*(18), 3361–3370. https://doi.org/10.1007/s00213-016-4372-0

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology*, *11*(2), 139–145. https://doi.org/10.1037/1064-1297.11.2.139

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Marzuki, A. A., Tomić, I., Ip, S. H. Y., Gottwald, J., Kanen, J. W., Kaser, M., Sule, A., Conway-Morris, A., Sahakian, B. J., & Robbins, T. W. (2021). Association of Environmental Uncertainty With Altered Decision-making and Learning Mechanisms in Youths With Obsessive-Compulsive Disorder. *JAMA Network Open*, *4*(11), e2136195. https://doi.org/10.1001/jamanetworkopen.2021.36195

Maulhardt, S. R., Solway, A., & Charpentier, C. J. (2026). Information uncertainty influences learning strategy from sequentially delayed rewards. *PLOS Computational Biology*, *22*(2), e1013879.

Maxwell, A. L., Gardiner, E., & Loxton, N. J. (2020). Investigating the relationship between reward sensitivity, impulsivity, and food addiction: A systematic review. *European Eating Disorders Review*, *28*(4), 368–384. https://doi.org/10.1002/erv.2732

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81.

McDonald, R. P. (2013). *Test theory: A unified treatment*. psychology press.

McKerchar, T. L., Green, L., Myerson, J., Pickford, T. S., Hill, J. C., & Stout, S. C. (2009). A comparison of four models of delay discounting in humans. *Behavioural Processes*, *81*(2), 256–259. https://doi.org/10.1016/j.beproc.2008.12.017

McTeague, L. M., Goodkind, M. S., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of Psychiatric Research*, *83*, 37–46. https://doi.org/10.1016/j.jpsychires.2016.08.001

Melnikoff, D. E., & Bargh, J. A. (2018). The Mythical Number Two. *Trends in Cognitive Sciences*, *22*(4), 280–293. https://doi.org/10.1016/j.tics.2018.02.001

Metcalfe, J., & Mischel, W. (1999). A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower. *Psychological Review*, *106*(1), 3–19.

Meyers, E. A., & Koehler, D. J. (2021). Individual differences in exploring versus exploiting and links to delay discounting. *Journal of Behavioral Decision Making*, *34*(4), 515–528. https://doi.org/10.1002/bdm.2226

Miller, E., Joseph, S., & Tudway, J. (2004). Assessing the component structure of four self-report measures of impulsivity. *Personality and Individual Differences*, *37*(2), 349–358. https://doi.org/10.1016/j.paid.2003.09.008

Minsky, M. (1961). Steps toward Artificial Intelligence. *Proceedings of the IRE*, *49*(1), 8–30. https://doi.org/10.1109/JRPROC.1961.287775

Mischel, W. (1974). Processes in delay of gratification. In *Advances in experimental social psychology* (Vol. 7, pp. 249–292). Elsevier.

Mischel, W. (2013). *Personality and assessment*. Psychology Press.

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*(2), 329–337. https://doi.org/10.1037/h0029815

Mischel, W., Ebbesen, E. B., & Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *21*(2), 204.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of Gratification in Children. *Science*, *244*(4907), 933–938. https://doi.org/10.1126/science.2658056

Mobini, S., Grant, A., Kass, A. E., & Yeomans, M. R. (2007). Relationships between functional and dysfunctional impulsivity, delay discounting and cognitive distortions. *Personality and Individual Differences*, *43*(6), 1517–1528. https://doi.org/10.1016/j.paid.2007.04.009

Moeller, F. G., Barratt, E. S., Dougherty, D. M., Schmitz, J. M., & Swann, A. C. (2001). Psychiatric Aspects of Impulsivity. *American Journal of Psychiatry*, *158*(11), 1783–1793. https://doi.org/10.1176/appi.ajp.158.11.1783

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

Moran, R., Keramati, M., Dayan, P., & Dolan, R. J. (2019). Retrospective model-based inference guides model-free credit assignment. *Nature Communications*, *10*(1), 750. https://doi.org/10.1038/s41467-019-08662-8

Moreira, D., Pinto, M., Almeida, F., & Barbosa, F. (2016). Time perception deficits in impulsivity disorders: A systematic review. *Aggression and Violent Behavior*, *27*, 87–92. https://doi.org/10.1016/j.avb.2016.03.008

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area Under The Curve as a Measure of Discounting. *Journal of the Experimental Analysis of Behavior*, *76*(2), 235–243. https://doi.org/10.1901/jeab.2001.76-235

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology, Special Issue: Dynamic Decision Making*, *53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157. https://doi.org/10.1523/JNEUROSCI.2978-14.2015

Nuske, J., Nuske, L., Stevens, M. W. R., Billieux, J., Delfabbro, P. H., Hides, L., Johnson, D., & King, D. L. (2026). The association between gaming disorder and impulsivity: A systematic review and meta-analysis. *Australian & New Zealand Journal of Psychiatry*, *60*(1), 26–41. https://doi.org/10.1177/00048674251388779

O'Doherty, J. P., Lee, S. W., Tadayonnejad, R., Cockburn, J., Iigaya, K., & Charpentier, C. J. (2021). Why and how the brain weights contributions from a mixture of experts. *Neuroscience & Biobehavioral Reviews*, *123*, 14–23. https://doi.org/10.1016/j.neubiorev.2020.10.022

Okan, A., & Hallquist, M. N. (2025). Negative affect-driven impulsivity as hierarchical model-based overgeneralization. *Trends in Cognitive Sciences*, S1364661325000026. https://doi.org/10.1016/j.tics.2025.01.002

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

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive Control Predicts Use of Model-based Reinforcement Learning. *Journal of Cognitive Neuroscience*, *27*(2), 319–333. https://doi.org/10.1162/jocn_a_00709

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, *24*(4), 1234–1251. https://doi.org/10.3758/s13423-016-1199-y

Peters, J., & D'Esposito, M. (2020). The drift diffusion model as the choice rule in inter-temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and controls. *PLOS Computational Biology*, *16*(4), e1007615. https://doi.org/10.1371/journal.pcbi.1007615

Pierce, J., Hirst, R., & MacAskill, Michael. (2022). *Building Experiments in PsychoPy* (Version 3) [Computer software].

Pike-Burke, C., Agrawal, S., Szepesvari, C., & Grünewälder, S. (2017). Bandits with delayed anonymous feedback. *Stat*, *1050*, 20.

Pike-Burke, C., Agrawal, S., Szepesvari, C., & Grunewalder, S. (2018). Bandits with delayed, aggregated anonymous feedback. *International Conference on Machine Learning*, 4105–4113.

Poulton, A., & Hester, R. (2020). Transition to substance use disorders: Impulsivity for reward and learning from reward. *Social Cognitive and Affective Neuroscience*, *15*(10), 1182–1191. https://doi.org/10.1093/scan/nsz077

Qualtrics. (2005). *Qualtrics* (Version 2024) [Computer software]. https://www.qualtrics.com

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2018). Modeling individual differences in the go/no-go task with a diffusion model. *Decision*, *5*(1), 42–62. https://doi.org/10.1037/dec0000065

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychological Review*, *120*(3), 697–719. https://doi.org/10.1037/a0033152

Redish, A. D., Jensen, S., & Johnson, A. (2008). A unified framework for addiction: Vulnerabilities in the decision process. *Behavioral and Brain Sciences*, *31*(4), 415–437. https://doi.org/10.1017/S0140525X0800472X

Reiter, A. M. F., Deserno, L., Wilbertz, T., Heinze, H.-J., & Schlagenhauf, F. (2016). Risk Factors for Addiction and Their Association with Model-Based Behavioral Control. *Frontiers in Behavioral Neuroscience*, *10*. https://doi.org/10.3389/fnbeh.2016.00026

Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II* (pp. 64–99). Appleton-Century-Crofts, New York.

Reynolds, B., Ortengren, A., Richards, J. B., & de Wit, H. (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences*, *40*(2), 305–315. https://doi.org/10.1016/j.paid.2005.03.024

Reynolds, B., & Schiffbauer, R. (2004). Measuring state changes in human delay discounting: An experiential discounting task. *Behavioural Processes*, *67*(3), 343–356. https://doi.org/10.1016/j.beproc.2004.06.003

Robbins, T. W., Gillan, C. M., Smith, D. G., De Wit, S., & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: Towards dimensional psychiatry. *Trends in Cognitive Sciences*, *16*(1), 81–91. https://doi.org/10.1016/j.tics.2011.11.009

Rodriguez, C. A., Turner, B. M., Van Zandt, T., & McClure, S. M. (2015). The neural basis of value accumulation in intertemporal choice. *European Journal of Neuroscience*, *42*(5), 2179–2189. https://doi.org/10.1111/ejn.12997

Romeu, R. J., Haines, N., Ahn, W.-Y., Busemeyer, J. R., & Vassileva, J. (2020). A computational model of the Cambridge gambling task with applications to substance use disorders. *Drug and Alcohol Dependence*, *206*, 107711. https://doi.org/10.1016/j.drugalcdep.2019.107711

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Sadeghiyeh, H., Wang, S., Alberhasky, M. R., Kyllo, H. M., Shenhav, A., & Wilson, R. C. (2020). Temporal discounting correlates with directed exploration but not with random exploration. *Scientific Reports*, *10*(1), 4020. https://doi.org/10.1038/s41598-020-60576-4

Šašinka, Č., Lacko, D., Čeněk, J., Popelka, S., Ugwitz, P., Řádová, H., Fabianová, M., Šašinková, A., Brančík, J., & Jankovská, M. (2023). ImGo: A Novel Tool for Behavioral Impulsivity Assessment Based on Go/NoGo Tasks. *Psychological Reports*, *126*(1), 434–476. https://doi.org/10.1177/00332941211040431

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599.

Sharma, L., Markon, K. E., & Clark, L. A. (2014). Toward a theory of distinct types of "impulsive" behaviors: A meta-analysis of self-report and behavioral measures. *Psychological Bulletin*, *140*(2), 374–408. https://doi.org/10.1037/a0034418

Shenhav, A., Rand, D. G., & Greene, J. D. (2017). The relationship between intertemporal choice and following the path of least resistance across choices, preferences, and beliefs. *Judgment and Decision Making*, *12*(1), 1–18.

Singmann, H., Brown, S., Gretton, M., Heathcote, A., Voss, A., Voss, J., & Terry, A. (2018). *Package 'rtdists.'*

Skippen, P., Matzke, D., Heathcote, A., Fulham, W. R., Michie, P., & Karayanidis, F. (2019). Reliability of triggering inhibitory process is a better predictor of impulsivity than SSRT. *Acta Psychologica*, *192*, 104–117. https://doi.org/10.1016/j.actpsy.2018.10.016

Skvortsova, V., & Hauser, T. (2022). *Distinct computational mechanisms underlying cognitive flexibility deficits in impulsivity and compulsivity* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-1280535/v1

Smith, C. L., & Hantula, D. A. (2008). Methodological considerations in the study of delay discounting in intertemporal choice: A comparison of tasks and modes. *Behavior Research Methods*, *40*(4), 940–953. https://doi.org/10.3758/BRM.40.4.940

Smith, R., Taylor, S., Wilson, R. C., Chuning, A. E., Persich, M. R., Wang, S., & Killgore, W. D. S. (2022). Lower Levels of Directed Exploration and Reflective Thinking Are Associated With Greater Anxiety and Depression. *Frontiers in Psychiatry*, *12*, 782136. https://doi.org/10.3389/fpsyt.2021.782136

Solway, A., Lin, Z., & Kaplan, C. M. (2021). Revisiting verbal recognition memory in obsessive-compulsive disorder: A computational approach. *Journal of Psychiatric Research*, *138*, 428–435. https://doi.org/10.1016/j.jpsychires.2021.04.015

Solway, A., Lin, Z., & Vinaik, E. (2021). Transfer of information across repeated decisions in general and in obsessive–compulsive disorder. *Proceedings of the National Academy of Sciences*, *118*(1), e2014271118. https://doi.org/10.1073/pnas.2014271117

Solway, A., Lohrenz, T., & Montague, P. R. (2017). Simulating future value in intertemporal choice. *Scientific Reports*, *7*(1), 43119. https://doi.org/10.1038/srep43119

Sripada, C., & Weigard, A. (2021). Impaired Evidence Accumulation as a Transdiagnostic Vulnerability Factor in Psychopathology. *Frontiers in Psychiatry*, *12*, 627179. https://doi.org/10.3389/fpsyt.2021.627179

Stan Development Team. (2024). *Stan Reference Manual* [Computer software].

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., & Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences*, *47*(5), 385–395. https://doi.org/10.1016/j.paid.2009.04.008

Strickland, J. C., & Johnson, M. W. (2021). Rejecting impulsivity as a psychological construct: A theoretical, empirical, and sociocultural argument. *Psychological Review*, *128*(2), 336–361. https://doi.org/10.1037/rev0000263

Sutton, R. S. (1984). *Temporal credit assignment in reinforcement learning*. University of Massachusetts Amherst.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (Second edition). The MIT Press.

Tanaka, S. C., Shishida, K., Schweighofer, N., Okamoto, Y., Yamawaki, S., & Doya, K. (2009). Serotonin Affects Association of Aversive Outcomes to Past Actions. *Journal of Neuroscience*, *29*(50), 15669–15674. https://doi.org/10.1523/JNEUROSCI.2799-09.2009

Thorndike, E. L. (1911). *Animal intelligence; experimental studies*.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189.

Tran, N.-H., Van Maanen, L., Heathcote, A., & Matzke, D. (2021). Systematic Parameter Reviews in Cognitive Modeling: Towards a Robust and Cumulative Characterization of Psychological Processes in the Diffusion Decision Model. *Frontiers in Psychology*, *11*, 608287. https://doi.org/10.3389/fpsyg.2020.608287

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, *185*(4157), 1124–1131.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44–62. https://doi.org/10.1037/a0021765

Vasconcelos, A. G., Sergeant, J., Corrêa, H., Mattos, P., & Malloy-Diniz, L. (2014). When self-report diverges from performance: The usage of BIS-11 along with neuropsychological tests. *Psychiatry Research*, *218*(1–2), 236–243.

Vincent, B. T. (2016). Hierarchical Bayesian estimation and hypothesis testing for delay discounting tasks. *Behavior Research Methods*, *48*(4), 1608–1620. https://doi.org/10.3758/s13428-015-0672-2

Weigard, A., Huang-Pollock, C., & Brown, S. (2016). Evaluating the consequences of impaired monitoring of learned behavior in attention-deficit/hyperactivity disorder using a

Bayesian hierarchical model of choice response time. *Neuropsychology*, *30*(4), 502–515. https://doi.org/10.1037/neu0000257

Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*(4), 669–689. https://doi.org/10.1016/S0191-8869(00)00064-7

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*. https://doi.org/10.3389/fninf.2013.00014

Wilbertz, T., Deserno, L., Horstmann, A., Neumann, J., Villringer, A., Heinze, H.-J., Boehler, C. N., & Schlagenhauf, F. (2014). Response inhibition and its relation to multidimensional impulsivity. *NeuroImage*, *103*, 241–248. https://doi.org/10.1016/j.neuroimage.2014.09.021

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, *81*(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005

Witkowski, P. P., Rondot, L., Kurth-Nelson, Z., Garvert, M. M., Dolan, R. J., Behrens, T. E., & Boorman, E. D. (2025). *Neural mechanisms of credit assignment for delayed outcomes during contingent learning*. https://doi.org/10.7554/eLife.101841.2

Wittmann, M., & Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends in Cognitive Sciences*, *12*(1), 7–12. https://doi.org/10.1016/j.tics.2007.10.004

Xu, P., González-Vallejo, C., & Vincent, B. T. (2020). Waiting in intertemporal choice tasks affects discounting and subjective time perception. *Journal of Experimental Psychology: General*, *149*(12), 2289–2313. https://doi.org/10.1037/xge0000771

Yee, D. M., Leng, X., Shenhav, A., & Braver, T. S. (2022). Aversive motivation and cognitive control. *Neuroscience & Biobehavioral Reviews*, *133*, 104493. https://doi.org/10.1016/j.neubiorev.2021.12.016

Yeh, Y.-H., Myerson, J., & Green, L. (2021). Delay discounting, cognitive ability, and personality: What matters? *Psychonomic Bulletin & Review*, *28*(2), 686–694. https://doi.org/10.3758/s13423-020-01777-w

Zhong, G., Chen, T., Zhong, N., Rezapour, T., Haghparast, A., Jiang, H., Su, H., Yuan, T., Robbins, T. W., Du, J., & Zhao, M. (2025). Transdiagnostic neuromodulation of impulsivity: Current status and future trajectories. *Translational Psychiatry*, *15*(1), 209. https://doi.org/10.1038/s41398-025-03415-2

Zou, A. R., Muñoz Lopez, D. E., Johnson, S. L., & Collins, A. G. E. (2022). Impulsivity Relates to Multi-Trial Choice Strategy in Probabilistic Reversal Learning. *Frontiers in Psychiatry*, *13*, 800290. https://doi.org/10.3389/fpsyt.2022.800290