Dopaminergic Modulation of Intertemporal Choice

Dissertation zur Erlangung des Doktorgrades
der Humanwissenschaftlichen Fakultät der Universität zu Köln
nach der Promotionsordnung vom 18.12.2018
vorgelegt von:

Ben Jonathan Wagner
aus
Freiburg im Breisgau

09/2021

Diese Dissertation wurde von der Humanwissenschaftlichen
Fakultät der Universität zu Köln im Januar 2022 angenommen.

Content

Abbrevations 5

Preface 6

Introduction 7

Continuous Dimensions of Mental Health 10

Theoretical Background 13

Cortico-Striatal-Thalamo-Cortical Loops 13

Cortex to Striatum 14

Striatum to Cortex 15

Subprocesses Involved in Intertemporal Choice 16

Valuation 16

Self-Control 17

Future Imagination 18

Dopamine 21

Physiology and Anatomy 21

Reward Prediction Error 23

Incentive Salience 25

Theories Synthesizing Dopamine´s Role in Learning and Motivation 27

Study Specific Summaries 31

Methods 35

Computational Modelling of Intertemporal Choice 35

Hierarchical Bayesian Parameter Estimation 40

Studies 42

Dopaminergic Modulation of Human Intertemporal Choice: A Diffusion Model Analysis using the D2-Receptor Antagonist Haloperidol (2020). The Journal of Neuroscience. 42

Gambling environment exposure increases temporal discounting but improves model-based control in regular slot-machine gamblers (2021). Preprint BioRiv 55

Temporal discounting in adolescents and adults with Tourette syndrome (2021). PLOS One 111

Chronic deep brain stimulation of the human nucleus accumbens region disrupts the stability of inter-temporal preferences (2020). Preprint BioRiv 128

General Discussion 148

Effects of the D_2-Receptor Antagonist Haloperidol in Healthy Participants 148

Contextual Effects of Gambling Environments 149

Intertemporal Choice in Patients with Tourette Syndrome 152

Effects of Chronic Deep Brain Stimulation 153

Computational Cognitive Modelling 154

Limitations 156

Outlook 159
Conclusion 161
Publication bibliography 162
Contribution Statement 187

Abbrevations

ACC anterior cingulate cortex

ALIC anterior limb of the internal capsule

BG basal ganglia

CSTC cortico-striatal-thalamo-cortical

DA dopamine

DAergic dopaminergic

dACC dorsal anterior cingulate cortex

DAT dopamine active transporter

DDM drift diffusion model

DS dorsal striatum

fMRI functional magnetic resonance imaging

HDI highest density interval

L-DOPA levo-3,4-dihydroxyphenylalanine

lPFC lateral prefrontal cortex

NAcc nucleus accumbens

MB model-based

MF model-free

MTL medial temporal lobe

OFC orbitofrontal cortex

OCD obsessive compulsive disorder

PFC prefrontal cortex

SN substantia nigra

RL reinforcement learning

RT reaction time

RTs reaction times

TS Tourette syndrome

vmPFC ventromedial prefrontal cortex

VS ventral striatum

VTA ventral tegmental area

Preface

We constantly make decisions and these often involve trade-offs between direct and future outcomes. Often these trade-offs are minor or irrelevant, but in other cases they have major impacts. Even small decisions can have long lasting consequences when a specific decision pattern (e.g. never taking into account the future) persists and negative effects accumulate over time. In consequence, those decision patterns can become bad habits and contribute to maladaptive behavior and harmful consequences in the long-run. An example for such a trade-off is the decision between c(going out with friends, playing a video game, listening to music) or writing a thesis. Enjoying social interactions, games and music immediately pays off, while writing a thesis on the trade-off between smaller-sooner and larger-but-later rewards (probably) pays off in the future. How should one decide? Daily life requires these frequent trade-offs in various forms and different processes like valuation of decision options, prospection into the future, self-control and context orchestrate their outcome.

This dissertation tries to contribute to a better understanding of so-called intertemporal choices. Thereby it will especially focus on the role of dopamine in modulating decisions with time trade-offs in various populations, from healthy controls to humans with gambling problems and participants with neurological and psychiatric disorders. In view of the above and below I would like to thank my supervisor Prof. Dr. Jan Peters for the tremendous support and for everything I was allowed to learn during these times. Thank you for the opportunity to eventually become a scientist. I would like to thank my grandmother Renate Duderstadt for her kind support and always believing in me, even when no one else did. Finally, I would like to thank my girlfriend Nadja Hödl, for all her patience, humor and love.

Introduction

In cognitive science, decisions between mutually exclusive outcomes that unfold at temporally divergent points in time are known under the concepts of intertemporal choice, temporal or delay discounting (used interchangeably here). When faced with such intertemporal choices in the context of rewards humans typically devalue or discount rewards as a function of time to delivery. Thus, humans do often prefer smaller-sooner (SS) over larger-but-later (LL) rewards rendering intertemporal choice as one measure for choice impulsivity (Ainslie 1975; Mazur and Coe 1987; Peters and Büchel 2011; Lempert et al. 2019). Overall, the discounting of rewards has been broadly studied across economics, psychology and cognitive neuroscience for quite some time (Samuelson 1937; Loewenstein and Elster 1992; Grüne-Yanoff 2015; Peters and Büchel 2011). Scientific studies assessing this construct often rely on questionnaires such as the Kirby Monetary Choice Questionnaire (Kirby and Maraković 1996) or on computerized tasks (so-called intertemporal choice or delay discounting tasks), where participants have to make exclusive choices between rewarding options. These rewards can be primary reinforcers such as food or secondary reinforces like money and usually vary with respect to magnitude (e.g. 5€, 10€ or 50€) and delay (e.g. days, weeks or months) to receipt (Lempert et al. 2019).

The quantification of intertemporal choice, nowadays, is typically assessed via a mathematical model. Estimated model parameter then describe the degree of future devaluation/discounting, i.e. how short-sighted (impulsive; SS preference; steep discounting) or future oriented (self-controlled; LL preference; shallow discounting) the individual is in contrast to others in the sample (Lempert et al. 2019). Historically, stories of human myopia date back over 1000 years. A well-known example is part of the story of Odysseus. Odysseus had to be tied to a mast by his comrades while listening to the song of the sirens. All his sailors had to stuff beeswax in their ears to resist the sirens' appealing but deadly song. Metaphorically, this tale can be interpreted as a story of impulsivity and self-control in the face of desirable incentives.

Around 900 years later the first recorded scientific studies on intertemporal choice were dominated by economists. During the end of the 19th and beginning of the 20th century economists like Jevons and Böhm-Bawerk concluded that “a future feeling is always more important than a present one” or that “short-sighted choice is the consequence of an inability to imagine the future in detail” (Jevons 2013; Böhm-Bawerk 1891; Loewenstein and Elster 1992). Interestingly, nowadays it is a given fact that future imagination modulates delay discounting (Peters and Büchel 2010b; Gershman and Bhui 2020; Rösch et al. 2021). During these times

the most famous contribution likely stems from Samuelson´s (1937) model of discounted utility, which was widely adopted across multiple research disciplines (Samuelson 1937; Loewenstein and Elster 1992). One of the first important psychological experiments on intertemporal choice is nowadays known under the Stanford marshmallow experiment (Mischel and Ebbesen 1970) and a few years later, the finding of preference reversals in pigeons (Ainslie 1974) questioned the dominating view proposed by economists. Preference reversals describe the observation that immediate (SS) options are sometimes preferred as the time to reward delivery decreases. Note, that these inconsistencies are incompatible the economists view of "homo economicus", describing humans as rational deciders and are further inconsistent with fundamental assumptions proposed by the model of discounted utility, i.e. exponential discounting (see methods). In consequence, Georg Anslie in 1975 proposed that discounting of future rewards follows a hyperbolic function (Ainslie 1975). This idea was further elaborated by Mazur and Coe in 1987 and is still the most common model for describing the devaluation of reward over time (Mazur and Coe 1987). Thus, while it was long known that most humans do discount future rewards, the first experimental finding of a relationship between steep discounting and mental illnesses date back to 1968 (Shybut 1968). Since then, countless studies assessed links between discounting and mental health conditions.

During the 1990s research on the neurotransmitter dopamine (DA) expanded from its known role in movement and movement disorders (Coyle and Snyder 1969) to a role in predicting reward (Schultz et al. 1993) and causing motivation (Robinson and Berridge 1993). Since then the catecholamine DA is known to play a fundamental role reward-related decision-making and is likewise implicated in nearly all psychiatric diseases (Beaulieu and Gainetdinov 2011). Interestingly, both steep discounting (impulsive choice; extreme SS preference) and DA neurotransmission are associated with a range of potentially maladaptive behaviors ranging from substance use disorders (Taber et al. 2012; Bickel et al. 2014; Bickel et al. 2019; Rodriguez-Moreno et al. 2021), attention-deficit hyperactivity disorder (Shiels et al. 2009; Jackson and MacKillop 2016), obesity (Volkow et al. 2011; Volkow and Baler 2015; Amlung et al. 2016) and behavioral addictions, such as gambling disorder (Dixon et al. 2003; Potenza 2013; Wiehler and Peters 2015; Potenza 2018). Further, steep discounting and dopaminergic disturbances are observed in major depression, schizophrenia, borderline personality disorder and binge eating disorder (Friedel 2004; Tye et al. 2013; Maia and Frank 2017; Bickel et al. 2019; Amlung et al. 2019). On the other side, shallow discounting (extreme LL preference) is associated with increased suicidality, anorexia nervosa and in some cases with obsessive compulsive personality disorder (Amlung et al. 2019; Lempert et al. 2019). Interestingly, in

recent years many aspects of human cognitive functioning have proven to lie on larger and continuous dimensions (Casey et al. 2014). Given these associations (see above) and establishing view of continuous dimensions of behavior it was suggested that intertemporal choice/delay discounting also lies on a continuous dimension and likely constitutes a trans-diagnostic trait (for an illustration see Figure 1) (Lempert et al. 2019).

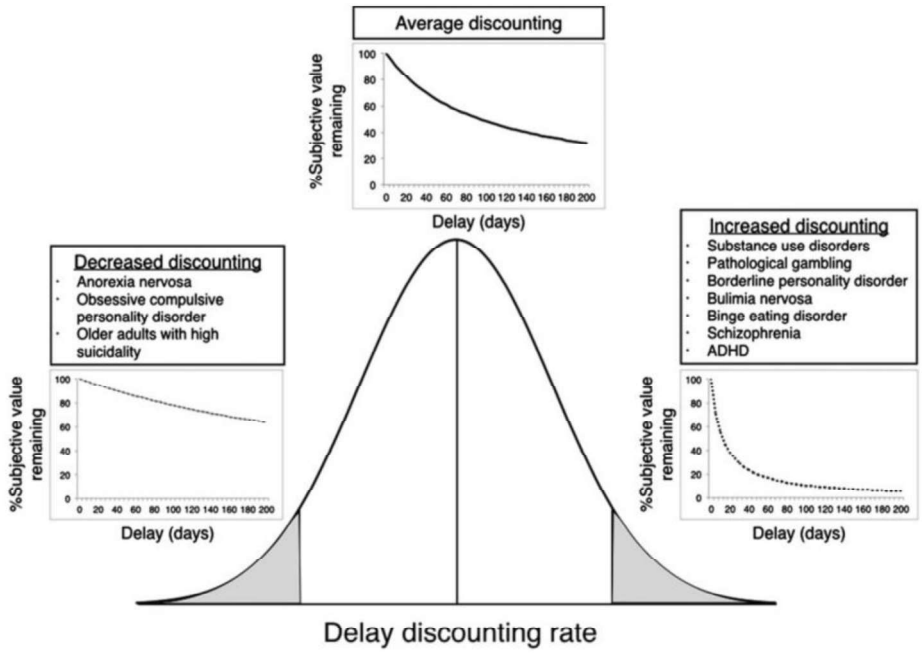


Figure 1. Adapted from Lempert et al. (2019). Illustration of delay discounting as a transdiagnostic process on a continuum.

Continuous Dimensions of Mental Health

In 2009 the National Institute of Menthal Health (NIMH) of the United States launched the Research Domain Criteria (RDoC). The mission of this ongoing initiative is to implement a new framework for investigating human functioning with the central goal of understanding mental disorders (Insel et al. 2010). The RDoC homepage describes the framework as a research strategy implemented in a dynamic matrix of elements spanning six domains of human functioning (for an overview see Figure 2 below). These consist of valence systems, motor-sensory- and regulatory circuits and social processing, each under constant influences of environmental and neurodevelopmental context. Importantly, each of these domains contains specific constructs designed to combine information, if applicable from genes over neural circuits to behavior (Kozak and Cuthbert 2016). The overall promise of this systematic approach is to foster our understanding of the continuous facets of mental health.

Intertemporal choice or the discounting of delayed rewards falls into the domain of "positive valence systems" under the construct of "reward valuation" (Lempert et al. 2019; U.S. Department of Health and Human Services, National Institutes of 2016). This construct is of high significance since reward, valuation of reward and motivation to work for reward all play crucial roles in human everyday life and mental health (Costello 1972; Alloy et al. 2016; Dutcher and Creswell 2018; Lempert and Phelps 2016; Lempert et al. 2019). In detail the construct aims to capture internal or external processes that influence the subjective valuation of reward (i.e. probability, delay or valence) and its subprocesses that might go awry in disease.

Delay discounting is believed to live up to the promise of the RDoC framework because research results suggest that it spans multiple levels of human functioning (it formally spans multiple matrix elements of the RDoC framework) (Lempert et al. 2019). First, findings suggest that delay discounting captures a trait-like variable. Delay discounting is relatively stable over time (Kirby 2009; Peters and Büchel 2011) and thus provides a reliable marker of choice preference. Further, studies linking behavior to genetics proposed that intertemporal preferences are partly inherited (Anokhin et al. 2011; Anokhin et al. 2015). Moreover, social processing, i.e. the context or social evaluation, plays an important role. For example, delay discounting changes as a function of community or social trust (Michaelson et al. 2013; Jachimowicz et al. 2017), parenting practices (Schneider et al. 2014), poverty (Lawrance 1991) and income in general (Green et al. 1996; Hampton et al. 2018). Specific contexts (Dixon et al. 2006), option framing (Peters and Büchel 2010b) and pharmacological manipulations (Weber et al. 2016; Pine et al. 2010) have shown to modulate choice preferences. In consequence, controlled experiments can assess long-term trait-like and short-term state-like effects in detail

(Peters and Büchel 2011). Manipulations of DA neurotransmission can inform the interplay of modulating neurotransmission of a specific neurotransmitter and behavioral choice (Pine et al. 2010; Weber et al. 2016; Foerde et al. 2016). Studying contextual manipulation of different populations, i.e. in specific cue-reactivity designs can provide deeper insights of decision-making processes in addiction. For example, gambling related cues (Miedl et al. 2014) or real-life environmental gambling contexts (Dixon et al. 2006) have shown to increase delay discounting and alter neural reward representations in pathological gamblers (Miedl et al. 2012). In contrast the framing of LL reward options via individualized future related cues has proven to decrease discounting in controls (Peters and Büchel 2010b). Intra-individual differences in discounting have successfully been linked to neural processes like value representation (Peters and Büchel 2010a), functional or structural connectivity (van den Bos et al. 2014) and specific valuation processes in adolescence (Huang et al. 2017).

In terms of neurodevelopment it is known that delay discounting decreases from adolescence to adulthood (Ripke et al. 2012). Likewise the prefrontal cortex (PFC) is undergoing crucial changes during maturation in adolescence (Caballero et al. 2016). These changes are associated with the development of cognitive control (Crone and Steinbeis 2017) and decreases in impulsive choice (Steinbeis et al. 2016). Delay discounting has repeatedly been linked to real-life behavior. A study found that choice preferences in to childhood to some degree are predictive of adolescent competence (Mischel et al. 1988) and academic success (Kirby et al. 2005). Delay discounting in adults predicted creditworthiness (Meier and Sprenger 2012) and retirement savings (Hershfield et al. 2011). The steepness of delay discounting further, was associated with smoking (Peters et al. 2011), alcohol consumption (Rossow 2008), risky sexual-practices (Chesson et al. 2006) and impulsivity in general (Reimers et al. 2009). It has further been associated with risk for relapse (Sheffer et al. 2014) and has shown to hold predictive value for treatment outcome (Stanger et al. 2012; Athamneh et al. 2017) and was suggested a potential target for therapeutic intervention (Odum 2011).

Importantly, delay discounting is easy to quantify and allows for relatively fast, intuitive and standardized protocols (Lempert et al. 2019). Parameters can easily be compared within or between individuals and are, due to their intuitive nature, relatively robust against biases (Grimm 2010) and other problems with questionnaire assessment (Kaplan and Saccuzzo 2018; Alwin 2006). Overall, intertemporal choice has high potential for linking behavioral choice to neural circuits, developmental processes, context and might hold predictive value for treatment outcome and therefore suits the systematic approach suggested by the RDoC framework (see Figure 2 for an overview).

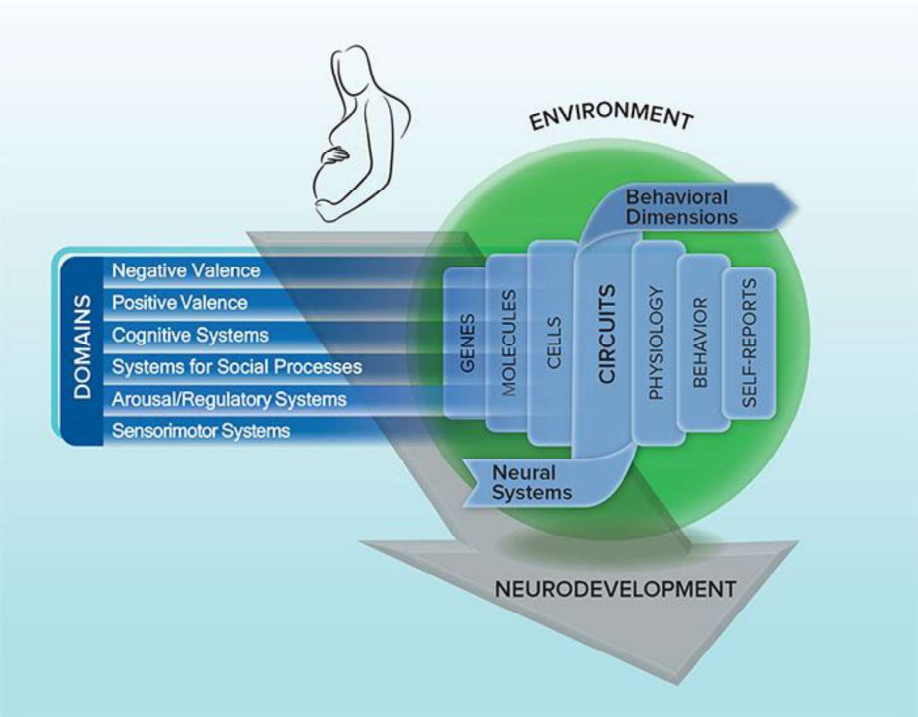


Figure 2. Illustration of the RDoC framework. Domains are listed in rows, elements in columns. Adapted from the RDoC Homepage (U.S. Department of Health and Human Services, National Institutes of 2016).

Theoretical Background

The Studies in this dissertation asses intertemporal choice in different populations ranging from healthy individuals to problem gamblers and patients with neurological disorders. The grand topic moreover focusses on the role of the neurotransmitter DA. In the following section I will briefly introduce the anatomy, neural systems and neuropharmacology that is believed to orchestrate reward-related decision-making. I will first illustrate concepts of the cortico-striatal-thalamo-cortical (CSTC) loops, the basics in anatomy and highlight the role of important basal ganglia (BG) pathways.

Further I will use this framework and integrate known concepts of neural processes involved in intertemporal choice. The first overview in parts will follow findings of one state-of-the-art publication on CSTC anatomy (Haber and Knutson 2010) and well-studied subprocesses in intertemporal choice as primarily summarized by Peters and Büchel (Peters and Büchel 2011). I will then continue to introduce the basic anatomy and physiology of DA. DA has an important role in modulating activity in CSTC loops (Satoh et al. 2003; Saddoris et al. 2015; Rogers 2011; Westbrook et al. 2021). Therefore, the last part of this chapter focusses on DAs qualitative role as a neurotransmitter in modulating circuits within the BG.

Cortico-Striatal-Thalamo-Cortical Loops

The CSTC loops are neural circuits projecting from the cortex to the BG, the thalamus and back to cortical areas. Said differently, projections originating in cortical cells, associated with action plans and goal representations, form a loop connecting these regions to the BG, the dopaminergic midbrain (both associated with reward and motivational processing) and then project back to regions in spatial proximity to where they originated (Haber and Knutson 2010). Until the late 1970s, these CSTC loops were primarily associated with their role in motor and sensory functions, most likely due to their known involvement in movement disorders such as Parkinson's disease (Coyle and Snyder 1969; Haber and Knutson 2010). The general idea was that movement plans originating in cortex are modulated along the way, resulting for example in the automatic execution of motor plans or the learning and refinement of motor skills. Following the work on the ventral striatum (VS), this idea was revised with the concept of an additional limbic loop associated with emotional and motivational processes (Heimer et al. 1982; Haber and Knutson 2010). Further findings led to the view of at least three domain specific and segregated circuits. This point of view suggested multiple in parallel working CSTC loops each corresponding to either motor, associative or limbic cortical functions (Alexander et al. 1990; Parent and Hazrati 1995; Haber and Knutson 2010). However,

nowadays the concept of segregated functional CSTC-loops has been revised and replaced with the view of parallel pathways that interact and integrate information at multiple nuclei along the loop (Draganski et al. 2008; Haber and Knutson 2010). This revised view also lines up with the theoretical perspective that adaptive behavior requires the integration of multiple sources of information, e.g. sensory information, associative learning and the retrieval of information from memory plus motivational components (Haber and Knutson 2010).

Cortex to Striatum

Cortical inputs to the striatum are organized in a topographical manner where the VS, including the nucleus accumbens (NAcc), receive inputs from the ventro-medial-prefrontal cortex (vmPFC) and broader orbitofrontal-cortex (OFC) areas, both associated with reward valuation (Peters and Büchel 2010a). Importantly, the most medial part of the VS, the NAcc, also receives most inputs from the medial part of the PFC, the vmPFC (see Figure 3 below). In contrast the dorsolateral striatum (head of the caudate nucleus) receives more inputs from sensorimotor areas which also originate more caudally in cortex. However, none of these topographic patterns are exclusive and cortical projections often interface between different projection areas and therefore provide evidence of functional integration (Haber and Knutson 2010). Still, this distinction between movement and reward associated projections, namely the topographic organization is also evident in striatal afferents that originate in the dopaminergic (DAergic) midbrain. DA neurons from the ventral tegmental area (VTA; mesolimbic pathway) project most exclusively to the VS especially the NAcc.

These anatomical projections are in line with activations commonly found in human imaging studies on reward valuation and subjective value (Peters and Büchel 2010a; Haber and Knutson 2010). Substantia nigra (SN) neurons project more laterally to the dorsal striatum (DS) (the so-called nigrostriatal pathway). The lateral prefrontal cortex (lPFC) also projects most densely to dorsal parts of the striatum. The dorsal anterior cingulate cortex (dACC), associated with conflict monitoring (Braem et al. 2017; Ebitz et al. 2020), projects mostly to areas directly ventral to those of lPFC. The dACC is further deeply integrated within other prefrontal areas, i.e. the OFC and especially the vmPFC. Anatomically, hippocampal and amygdaloid fibers project most densely to the VS and especially the NAcc region making the VS the main entry point where motivational and emotional information converge and enter the BG (Haber and Knutson 2010). The VS is also the region where fibers from vmPFC overlap with those from amygdala and hippocampus making it one spot of informational integration (Haber and Knutson

2010). In intertemporal decision-making findings support the importance of information integration in these regions, for example in terms of future imagination (see below).

Striatum to Cortex

The striatum is the main entry point to the BG and receives major inputs from cerebral cortex and the DAergic midbrain. The most important and major striatal outputs, both from the ventral and dorsal part, project through the remaining BG nuclei, that is via the globus pallidus internal and external segments to the thalamus before projecting back to cortex (Alexander et al. 1990).

In detail, concepts dissociate two different classes of pathways the so-called direct pathway and the so-called indirect pathway (Frank and O'Reilly 2006), both of which reach out to the thalamus via different routes. The direct pathway projects from striatal areas to the internal segment of the globus pallidus (GPi). Activity in this pathway results in inhibition of the GPi. While normally, the GPi inhibits the thalamus, inhibition of the GPi releases the thalamus from inhibition. This does not activate the thalamus directly but enables other direct excitatory projections to excite the thalamus and therefore provides a gating function (Frank and O'Reilly 2006).

In contrast, the indirect pathway first projects to the external segment of the globus pallidus (GPe) which regularly tonically inhibits the GPi via the subthalamic nucleus. Activity in the indirect pathway inhibits the GPe and therefore takes away the inhibition of the GPi, which then results in overall increased inhibition of the thalamus (Parent and Hazrati 1995; Frank 2005; Frank and O'Reilly 2006). Several models propose that activity in both pathways compete, and in consequence modulate action plans originating in cortical areas (Frank and O'Reilly 2006; Parent and Hazrati 1995; Collins and Frank 2014). Importantly, DA is deeply involved in further modulating these already modulatory CSTC loops via its action on both direct and indirect BG pathways (see section on DA below) (Frank and O'Reilly 2006; Frank 2005; Collins and Frank 2014; Westbrook et al. 2021; Westbrook et al. 2020).

Subprocesses Involved in Intertemporal Choice

Intertemporal decisions are the product of at least three dissociable but integrated subprocesses (Peters and Büchel 2011). In what follows I will integrate these processes within the CSTC-loop framework introduced above.

Valuation

Valuation is subjective and it is thus plausible that this subjective value is represented in neural activity. Research in decision neuroscience has identified strong evidence for value-correlates in the OFC, especially the ventromedial part, the vmPFC (Peters and Büchel 2010a; Bartra et al. 2013; Seaman et al. 2018). The vmPFC is a highly integrated region with strong connections to other prefrontal areas such as the lPFC and other reward associated regions like the VS and DAergic midbrain, both of which play important roles in valuation processes. Together these regions make up a so-called valuation network (FitzGerald et al. 2009; Peters and Büchel 2010a; Bartra et al. 2013). Evidence for a unique domain spanning account of this concept comes from empirical studies ranging from various choice paradigms and stimulus types. For example, Peters and Büchel (2009) showed that subjective value of delayed monetary rewards (estimated via hyperbolic discounting) correlated with activity in the VS and some parts of OFC (Peters and Büchel 2009). Hare and colleagues and others identified a vmPFC subjective value signal in a food-choice task (Hare et al. 2009; Tabibnia et al. 2011; Harris et al. 2013). Hariri and colleagues further reported an association of VS activity and choice preferences in intertemporal choice (Hariri et al. 2006).

Importantly, in the past there was an ongoing debate whether potentially dissociable striatal and prefrontal value signals contribute to choice during intertemporal decision-making. In this dual-systems view, a VS/vmPFC value signal corresponds to the concrete immediate outcomes and a prefrontal valuation signal corresponds to the value of more abstract LL rewards . Choice is then modelled as a function of a specific weighting/ competition between both valuation systems (McClure et al. 2004). However, new findings suggest one unique valuation systems for all types of rewards (Kable and Glimcher 2007) where the lPFC exerts top-down control in support of self-controlled choices (Hare et al. 2009; Figner et al. 2010; Peters and D'Esposito 2016). This view is also supported by studies providing evidence for the integration of costs and benefits of decision value within one vmPFC value signal (Basten et al. 2010).
To sum up, human imaging studies and primate anatomy converge on evidence for prefrontal and striatal recruitment for various types of rewards. Here primary and secondary rewards generally recruit OFC, especially the vmPFC the VS and the DAergic midbrain (see Figure 3)

(Peters and Büchel 2009; Haber and Knutson 2010). Empirical evidence emerged in studies on food-choice (Hare et al. 2009; Harris et al. 2013), risky decision-making (Peters and Büchel 2009), intertemporal choice (Peters and Büchel 2009; Hare et al. 2014) and others (Clithero and Rangel 2014; Peters and Büchel 2010a).

Self-Control

A central distinction between humans and other living organisms is the high degree of selectively in choosing to which sensory information humans react and what they ignore. Navigating complex environments, filtering information and resisting distraction is an essential ability to achieve short- and long-term goals (Miller 2000). In cognitive neuroscience research on these mechanisms is often summarized under the term of cognitive control. Cognitive control is anatomically attributed to the PFC. The PFC is characterized by widespread connections through which it can exert a modulating influence on a wide range of cognitive processes (see Miller 2000 for an overview).

With respect to intertemporal choice, cognitive control is especially useful when long-term goals require to resist immediate gratification. The concept of cognitive control in such a situation can thus be viewed as one type of self-control (Harris et al. 2013). Evidence for an involvement of self-control in intertemporal choice stems from several studies. For example, artificially impairing dorsolateral-prefrontal cortex (dlPFC) results in less far sighted choice/steeper discounting (Figner et al. 2010). Functional connectivity between regions associated with valuation and PFC during the choice period is predictive of successful and unsuccessful dieters (Hare et al. 2009; Harris et al. 2013). Hare and colleagues (2014) provided further evidence that increased lPFC activity is especially evident when subjects prefer LL over immediate SS options (Hare et al. 2014). Moreover, there is evidence of a scaling effect, i.e. dlPFC activity increases even more with choice difficulty particular in self-controlled individuals (Jimura et al. 2018). Another study by van den Bos and colleagues confirmed these functional effects and extended the findings in terms of structural differences. That is, far-sighted choice was associated with higher density of structural integration of lPFC and VS (van den Bos et al. 2014). Another well-replicated effect on temporal discounting refers to the observation that the rate of temporal discounting decreases (less impulsive choice) with increasing reward magnitude (Green et al., 1997). This so-called magnitude-effect also depends on lPFC processing (Ballard et al., 2017). Further, non-substance use addictions like pathological gambling (characterized by increased discounting) are associated with decreased PFC response during decision-making (Tanabe et al. 2007). Likewise, evidence for a pivotal

role of the PFC stems from the developmental trajectory of self-control. It has been proposed that the developmental trajectory of PFC can be linked to vulnerability for addiction in adolescents (Volkow and Boyle 2018) and the developmental trajectory of cognitive control and intertemporal choice (Water et al. 2014; Achterberg et al. 2016).

Taken together, self-control is an important modulator of intertemporal choice. Studies suggest that self-control resides mostly within regions in the PFC, likely in the dorsolateral part. The dlPFC can modulate vmPFC and subcortical value representations to foster long-term achievements and resist immediate incentives. Changes in structural and functional connectivity within these circuits have been linked to self-controlled choice in both adolescents (van den Bos et al. 2015) and adults (van den Bos et al. 2014) and the developmental trajectory from adolescence to adulthood [(van den Bos et al. 2015; Anandakumar et al. 2018); see Figure 3]

Future Imagination

The human ability to foresee and simulate the future may reach beyond those of any other (known) species. While retrospection refers to memorize and re-experience the past, prospection refers to our ability to imagine or pre-experience what might be tomorrow (Gilbert and Wilson 2007). Research in memory and future imagination revealed striking similarities and suggested a common network underlying both (Schacter et al. 2012). For example patients with specific hippocampal- or general medial temporal lobe (MTL) damages show significant impairments in both memory performance and future imagination (Hassabis et al. 2007; Race et al. 2011). And indeed one aspect of hippocampus function is the evaluation of possible decision outcomes (Lebreton et al. 2013).

Likewise intertemporal decisions often require imagination of future outcomes. In rodents, damage to hippocampal areas are associated with increased discounting (Mariano et al. 2009). In humans, research has found that (positive) future imagination can reduce delay discounting in both adults and adolescence, i.e. trials framed with personalized episodic cues were associated with more farsighted choice (Peters and Büchel 2010b; Bromberg et al. 2017; Rösch et al. 2021). One possible mechanism may be that imagining an event makes it seem more certain to happen (Bulley et al. 2016) or that imagining a positive (reward associated) event might induce some kind of positive anticipatory affect. The combination of positive affect and imagery might result in a less abstract representation and thus attenuated discounting (Benoit et al. 2018). Peters and Büchel (2011b) showed that the extend of episodic imagery predicted the reduction strength in delay discounting. In detail, the episodic future thinking

(EFT) condition differed from the control condition in activations in the right amygdala, anterior cingulate cortex (ACC) and dlPFC. Further, this was accompanied by an enhanced ACC-hippocampal coupling likely modulating subjective value via an upregulation of neural value signals (Peters and Büchel 2010b). Interestingly, a number of studies have associated the amygdala with the certainty and timing of reward delivery (Bermudez et al. 2012; van Holstein et al. 2020) and amygdala to striatal structural and functional connectivity is predictive of differences in intertemporal choice (van den Bos et al. 2014).

To sum up, MTL regions (see Figure 3) associated with both memory and future imagination, namely the hippocampus (memory, future imagination) and the amygdala (reward proximity and probability) are both associated with the effects of EFT on intertemporal choice (Peters and Büchel 2010b). Further, studies propose that these regions are integrated within a neural circuit responsible for integrating spontaneous future imagination into valuation processes (Peters and Büchel 2010b, 2011; Rösch et al. 2021; Masuda et al. 2020).

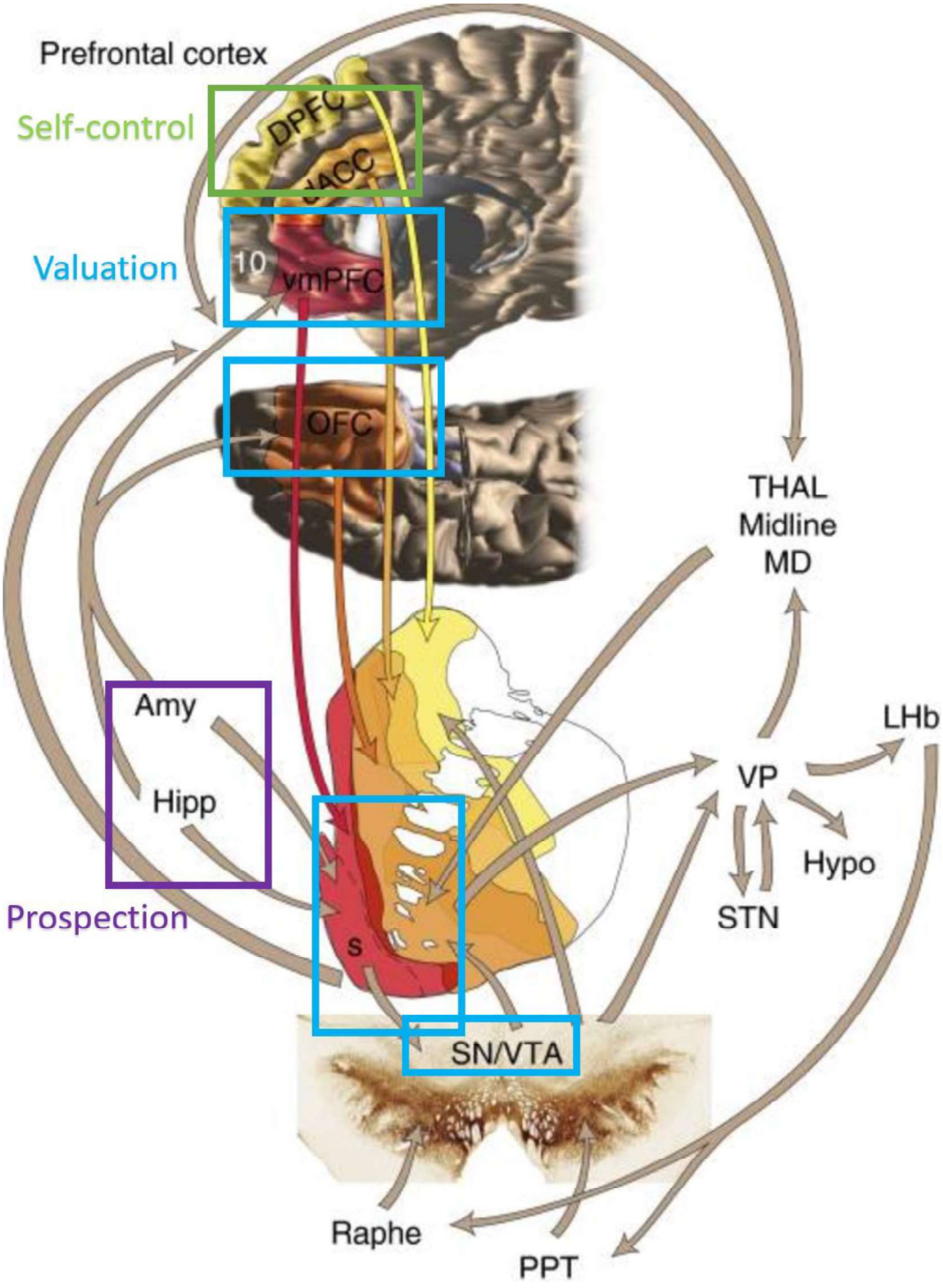


Figure 3. Schematic illustration of key structures and pathways of the human-reward circuit; Adapted from Haber and Knutson (2010) and modified. Regions associated with neural processing of intertemporal choice: valuation processes (blue), prospection (purple) and self-control (green). Regions associated with valuation: vmPFC = ventral medial prefrontal cortex; s = shell of the NAcc/ventral striatum; SN/VTA = substantia nigra/ventral tegmental area located in the midbrain; OFC = orbital frontal cortex. Medial temporal lobe regions associated with future imagination: Amy = amygdala; Hipp = hippocampus. Regions associated with self-control: DPFC = dorsal prefrontal cortex; dACC = dorsal anterior cingulate cortex. Other regions within the CSTC loops: Raphe = Raphe nucleus; PPT = pedunculopontine nucleus; hypo = hypothalamus; OFC = orbital frontal cortex; STN = subthalamic nucleus; Thal = Thalamus; VP = ventral pallidum; LHb = lateral habenula.

Dopamine

DA plays a fundamental role in modulating the CSTC loops and is implicated in mediating multiple functions like reward learning (Schultz 2015), reward valuation (Schultz et al. 2017), motivation (Wise 2004; Berridge 2012; Westbrook et al. 2020) and cognitive control (Cools 2008). The next pages briefly summarize the pharmacology of DA and its receptors and then continue on theories of qualitative function of DA neurotransmission.

Physiology and Anatomy

DA (4-[2-Aminoethyl]benzene-1,2-diol) is a monoamine compound of the family of catecholamines. It consists of a benzene ring with two hydroxyl side groups and a side-chain amine. DA is directly synthesized in the cytosol of DAergic neurons and packed into vesicles located within the presynaptic terminals of those (Beaulieu and Gainetdinov 2011; Meiser et al. 2013). Biochemically, DA it is the product of an enzymatic reaction, derived from the amino acid tyrosine and part of a bigger reaction cycle that also metabolizes epinephrine (adrenaline) and nor-epinephrine (nor-adrenaline) (Meiser et al. 2013). In the human brain around 80 % of DA neurons reside within the midbrain. Two regions, specifically the substantia nigra (SN) and the ventral tegmental area (VTA) contain the highest number of DA neuron cell bodies (Beaulieu et al. 2015).

There are three major DAergic pathways (given a classical distinction of DA pathways), each originating from the neighboring regions of the SN and VTA. Axons of the nigrostriatal pathway mostly project from cells located in the SN to the caudate nucleus (see CSTC loops above). This pathway is mostly associated with the control of voluntary movement and action selection and further plays a crucial role in the development of Parkinson's disease (Bernheimer et al. 1973; Haber and Knutson 2010). The so-called mesolimbic pathway projects from the VTA to the VS and most densely to the NAcc area. Further, projections of this pathway reach out to MTL regions such as the amygdala and hippocampus. This pathway is associated with processing of reward, reward magnitude (NAcc), probability (amygdala) and reward associated cues and memory (hippocampus) (Haber and Knutson 2010; Beaulieu and Gainetdinov 2011). The last major DA pathway projects from the VTA to cortical structures. This so-called mesocortical pathway is associated with executive functions like working memory, attention and cognitive control in general (Cools 2008; Ayano 2016). At first, all DA pathways travel through the median fore brain bundle and internal capsule before extensively branching out (~500,000 synapses per neuron) to their corresponding nuclei (Doucet et al. 1986; Björklund and Dunnett 2007).

In striatal target regions, DA neurons most exclusively synapse with medium spiny neurons (MSNs) of the direct and indirect BG pathways. DA can interact with pre- and postsynaptic DA receptors located on both DA and GABAergic MSN neurons, respectively. Here D_1 (D_1 and D_5) and D_2 (D_2-D_4) like DA receptors represent two main categories (Ayano 2016; Beaulieu et al. 2015). D_1-like receptors are bound to $G\alpha_{s/olf}$ proteins which stimulate signaling via signal-cascade that increases cyclic-adenosin-monophosphat (cAMP) levels. D_2-like receptors play an inhibiting role via decreasing cAMP levels and therefore decrease the probability of action potentials in the preceding neuron (Ford 2014; Beaulieu et al. 2015; Ayano 2016). Thus, when DA binds to D_1-like receptors (located on MSNs), the probability that these neurons depolarize rises, whereas when DA binds to D_2 - like receptors these neurons hyperpolarize. Besides these direct effects on polarization DA can initiate signal cascades that modulate further processes including long term potentiation (LTP) and long term depression (LTD) (Ayano 2016; Schultz 2016). The density/ratio of DA D_1- and D_2-receptors differs between structurally divergent BG pathways. GABAergic neurons projecting directly to the internal segment of the globus pallidus (direct pathway) have a higher ratio of D_1-receptors. MSNs of the pathway projecting indirectly (via GPE and the subthalamic nucleus; indirect pathway) to the GPi show a higher ratio of D_2 receptors (Aizman et al. 2000; Ayano 2016).

The literature differentiates between two distinct modes of DA neuron activity: Spike-dependent DA release via phasic bursts following high frequency cell firing (e.g. prediction error) and low frequency tonic DA release (Robinson et al. 2004; Schultz 2007; Ford 2014). One suggested mechanism that induces tonic DA release are low frequency single action potentials from projections from the PFC or other regions (e.g. amygdala, hippocampus, midbrain) that synapse with DAergic neurons in striatum (Nieoullon et al. 1978; Grace and Bunney 1984; Zhang et al. 2009). Following exocytosis and signaling, DA can be degraded by various mechanisms in the synaptic cleft. The DA active transporter (DAT) uses a symport transport mechanism to pump DA back out of the synaptic cleft into the presynaptic terminal, that is the cytosol of DAergic neurons. DAT is an integral membrane protein (Giros and Caron 1993) and the primary mechanism (Ciliax et al. 1999) through which DA is cleared from the synapse. Cleared DA is then repacked into synaptic storage or degraded via specific enzymes. Enzymatic degradation of DA is conducted mostly inside neuronal- or neighboring glial-cells by Monoamine-Oxidase or Catechol-O-Methyltransferase, both of which catalyze slightly different reactions that degrade DA to its various metabolites, e.g. homovanillicacid (Meiser et al. 2013).

One further important site of DA action are D_2-autoreceptors located in the presynaptic membrane directly on DA neurons. D_2-autoreceptores are G-protein coupled inhibitory receptors ($G_{i/o}$) that modulate the synthesis, release, cell firing and reuptake of DA via an inhibitory feedback mechanism (Ford 2014; Ayano 2016). D_2-autoreceptors are expressed in high density in striatal regions and to a lower degree in the hippocampus. D_2-autoreceptors in these regions are believed to regulate locomotor- and reward related behavior (Ford 2014). For example, animals lacking D_2-autoreceptors have shown hyperactive behavior and hypersensitivity to cocaine (Bello et al. 2011). Also, altered autoreceptor function correlates with changes in impulsivity (Buckholtz et al. 2010). D_2-antagonists like haloperidol are associated with differential effects as a function of dosage. For example, acute administration of low haloperidol concentrations are believed to primarily interact with D_2-autoreceptors (Frank and O'Reilly 2006). The resulting antagonistic inhibition of these autoreceptors in consequence can inhibit the DA feedback loop and therefore enhance tonic (Zhang et al. 2009) and/or phasic DA neurotransmission (Benoit-Marand et al. 2011; Frank and O'Reilly 2006) predominantly in regions with the highest D_2 density, e.g. the striatum and hippocampus (Ford 2014). Note, that all of these mechanisms mentioned above result in a complex interplay of DA release, pre- and postsynaptic DA receptors and reuptake that could be individually affected in neuropsychiatric disease and in consequence to medication.

Reward Prediction Error

One prominent theory suggests that DA neurons encode expectancy violations, that is they react in response to unexpected reward or when expected reward is omitted (Schultz 2013, 2015, 2016; Kobayashi and Schultz 2014). In the 1990s Schultz and colleagues (1993) studied firing patterns of DA neurons in rhesus monkeys. In detail, they recorded DA neuron activity in the context of salient stimuli presentation while these monkeys were learning a behavioral task. While, Schultz et al. knew a priori that the performance of such cognitive tasks was impaired when DA neurons were lesioned (Schultz and Romo 1990), they were able to show that around 75% of these neurons responded with phasic short latency firing to unpredicted liquid reward but responded with a reduction in firing rate when a reward was already predicted (Schultz et al. 1993). Shortly thereafter, this finding turned out as an important scientific achievement and marked the beginning of DAs role as an error signal. Further discoveries confirmed that the neuronal responsiveness of midbrain DA neurons is indeed determined by the particular context and timing of reward delivery, namely by its predictability (Schultz and Romo 1990; Schultz et al. 1993; Lak et al. 2016; Mirenowicz and Schultz 1994). These findings

also lined up with assumptions of learning theory suggesting that the rate of learning is proportionally related to predictability of reward (Sutton 1988; Sutton and Barto 1990, 1998). Computationally, these observations were compatible with simple temporal-difference reinforcement learning (RL) algorithms (Sutton 1988). These algorithms track an estimate of the average reward rate of a specific state or action. Said differently, they encode a summary value that each action or state-transition has produced in the past. In consequence, values are learned on the fly and new knowledge is integrated at each moment the agent (human or animal) revisits a specific state or action, but only when the obtained reward differs from what was predicted (Bayer and Glimcher 2005; Sutton and Barto 1998; Dayan and Niv 2008).

These perspectives (RPE and RL) were finally united in a 1996 collaboration of Montague, Dayan and Schultz (1996). Imaging studies later confirmed the neural correlates of prediction error in temporal difference learning models (Knutson and Cooper 2005). Further research expanded on this work to account for higher order decision problems. So called model-based (MB) RL tackles situations in which an agent has to infer the structure of the environment, evaluate and update the transition probabilities leading to specific states and rewards. In line with work on simple temporal-difference learning (also known as model-free [MF] RL) DA was found to report discrepancies between assumed and observed state-transitions in higher order MB RL (Gläscher et al. 2010). Over the following years the reward prediction error (RPE) hypothesis further evolved. For example, the signal is nowadays believed to consist of distinct components. A first sensory component is believed to represent physical salience (sensory impact), novelty or surprise, while shortly thereafter the signal rapidly transitions into detection and valuation of reward (Schultz 2016). Thus, the first component of the signal is relatively unrelated to reward and might explain why DA bursts have been associated with aversive stimuli (Joshua et al. 2008). Electrophysiological and voltammetry studies confirm this two-component prediction error response (Schultz 2016).

Interestingly, a salient context associated with reward beforehand can modulate both response components (Kobayashi and Schultz 2014; Kobayashi and Hsu 2019). A recent publication (Dabney et al. 2020) provided first evidence for distributional coding of DA neurons, that is individual neurons might encode different expectations resulting in a complete probability distribution. This is especially exciting in light of theories proposing that the brain computationally mirrors Bayesian principles and approximates probability densities of outcomes (Friston 2010, 2012). However, future research needs to clarify the mechanisms behind such theories.

Incentive Salience

While first formulated in 1993, incentive salience provides a framework for the involvement of DA in addiction and formulates evidence for motivational effects of DA neurotransmission (Robinson and Berridge 1993). While referring to the emergence of their theory Berridge and Robinson (2016) explain that the (an)hedonia hypothesis (Wise 1980), that assigned DA the role of a "pleasure neurotransmitter", has provided inconsistent evidence. In fact, studies provided evidence that DA is not necessary to produce liking reactions (Berridge et al. 1989; Berridge 2000; Berridge and Robinson 1998). The theory of incentive saliences highlights these differences between liking (pleasure of reward consumption) and wanting (craving for reward). For example, while pharmacological manipulation or impairments of DA activity changed wanting reactions, liking was relatively unaffected in animals (Berridge and Robinson 1998; Berridge et al. 1989; Cannon and Palmiter 2003) and human subjects (Sienkiewicz-Jarosz et al. 2005; Sienkiewicz-Jarosz et al. 2013; Leyton et al. 2005). For example, DA is not necessary for mice to preferably consume sweet solutions or to find sweet tastes rewarding (Cannon and Palmiter 2003). It was thus suggested, that liking per se is rather modulated by specific hedonic hotspots belonging to the opioid system than by DA (Peciña and Berridge 2005; Peciña et al. 2006). Secondly, Robinson and colleagues (2005) showed that DA depleted mice were still able to learn about rewards (Robinson et al. 2005). Another study confirmed these findings (Cagniard et al. 2006). Thus, DA is not always necessary to learn about reward, contradicting one important assumption of the RPE hypothesis. These findings are compatible with other research highlighting the motivational aspects of DA, i.e. its role in working for reward and willingness to expend cognitive resources (Cagniard et al. 2006; Day et al. 2010; Nunes et al. 2010; Westbrook and Braver 2016; Berke 2018).

Berridge and Robinson (1998, 2016) further argue that reward associated DA neurons assign "magnet-like" properties to specific internal representations or external stimuli. For example, specific cues or contexts that predict subjective reward. These representations are then rendered as appetitive and in consequence preferably induce approach-like behavior (Berridge and Robinson 1998; Berridge 2016). Incentive sensitization, a theoretical derivative of incentive salience, further elaborates on the exact role of DA in addiction and refines these ideas (Robinson and Berridge 2001). In detail, incentive sensitization posits that DAs systems are rendered hypersensitive (in response to specific cues), due to a history of excessive DA release in the context of drug-abuse or behavioral addiction. Since DA is supposed to induce wanting (see above), this hypersensitive DA system induces pathological craving for drugs of abuse, even though liking may be absent or reduced. Further, this should be especially evident in

contexts that are associated with addiction-related cues. Said differently, DA transfers wanting to associations, contexts and cues that predict the desired state (Berridge and Robinson 2016). Theoretically, this sensitization might be the result of a first excessive mesencephalic DA release as a function of pharmacology, i.e. DAT blockade (cocaine or amphetamine) or extensive prediction errors due to behavior alone (e.g. a first gambling win). However, the exact mechanisms are puzzling and for example studies on DA in gambling disorder showed inconsistent results. While some studies suggest that gamblers and controls show no overall difference in DA release while gambling (Joutsa et al. 2012; van Holst et al. 2018) others propose differences in striatal synthesis capacity (van Holst et al. 2018). Still, a bias in response to gambling cues (Oberg et al. 2011) and blunted responses to other incentives (Sescousse et al. 2013) are compatible with DA assigning "magnet-like" properties to addiction-specific cues.

Interestingly, there is evidence that incentive sensitization is associated with functional and structural changes within DAergic systems (Steketee and Kalivas 2011). Finally and contradicting with the suggested distinction between wanting (DAergic system) and liking (opiod system), in some recent studies DA sometimes correlated with subjective pleasantness (Kühn and Gallinat 2012), or enhancement of DA led to elevated liking (Ferreri et al. 2019). In consequence, the exact interplay of DA, the opioid system and reward liking might be more complex than previously thought and future studies are necessary to resolve these issues.

Theories Synthesizing Dopamine's Role in Learning and Motivation

Over the years considerable progress shaped our understanding of DAs effects on both learning (Schultz 2013, 2015, 2016) and motivation (Smith et al. 2011; Berridge 2012). However, both theories have also been prominently contrasted (Berridge 2012; Colombo 2014). Nevertheless, recent attempts aimed to synthesize both of these perspectives into more unifying accounts of DAs qualitative role in human functioning (Collins and Frank 2014; FitzGerald et al. 2015; Westbrook and Braver 2016; Berke 2018). Likewise theories on regional differences emerged, e.g. it was suggested that while RPE-scaled midbrain DA bursts signal learning, separate striatal DA fluctuations signal motivation and arise independently from midbrain DA cell bursts (Mohebi et al. 2019).

One promising model to combine both perspectives was suggested by Collins and Frank in 2014 (Collins and Frank 2014). Their approach builds on earlier work, namely the actor-critic model as illustrated in Sutton and Barto (1998). In this RL model, the critic and actor learn the state- and action-values of specific states and actions via temporal difference mechanism, respectively (Dayan and Niv 2008; Montague et al. 1996). While this model provided a compelling explanation of some findings, i.e. learning via RPE, others like asymmetries in context dependent choice or different weighting of learned values, are not well explained. For example, DA signals can ramp up in the presence of incentives which can boost task performance independently of learned action values and directly influence choice (Wassum et al. 2011; Smith et al. 2011; Berridge 2012). Thus, a simple actor-critic architecture can explain DAs effects on learning specific values, but is unable to account for DAs effects during choice (incentive salience /motivational properties).

The modified opponent actor and critic model (OPAL) can do both, while at the same time better approximating BG anatomy and function. In detail, Collins and Frank (2014) implemented positive (GO signal) and negative (NO-GO signal) action values (also known as Q-values; Q+ and Q-) and further added separate weighting parameters (β GO and β NO-GO) that simulate DA activity during choice. Note, these Q-values (Q+ and Q-) are updated (learned) via temporal-difference mechanism (Sutton and Barto 1998) and the weighting parameters β are just weights on these learned Q-values simulating DAs effects on both values, separately. Thus, DA during choice increases the β for Q+ (GO-singal) and decreases the β for Q- (NO-GO signal). In analogy to BG anatomy OPAL assumes these different action values do mirror the direct (GO) and indirect (NO-GO) pathway gain and further allows for context dependent weighting of these "learned" gains. Said differently, each pathway gain, modelled via separate action values, is learned. Mechanisms that shape learning are phasic DA bursts (positive

prediction errors) or dips below the baseline rate of DA neuron activity (negative prediction errors), both after choice. This translates to BG anatomy in the following way: Stimulating D_1-receptors increases the signal to noise ratio and amplifies activity (GO signal) in the direct pathway (increased action value for GO; see Figure 1B). A lack of DA at D_2-receptors after choice (DA dip) in the indirect pathway results in increased inhibition via the indirect pathway (increased action values for NO-GO; Figure 1C) (Collins and Frank 2014). Motivation is then reflected in the context dependent weighting of these already learned direct (GO) and indirect (NO-GO) action-values. For example, when a phasic burst or dip of DA neuron activity occurs during choice it immediately effects the direct pathway (GO; D_1-receptor mediated) MSNs or indirect pathway (NO-GO; D_2 mediated) MSNs (Shen et al. 2008; Collins and Frank 2014; Frank 2005) and is modelled via an effect on direct (GO) β weights or indirect (NO-GO) β weights, respectively. This allows the model to account for context effects, i.e. incentives or other cue effects without new learning due to a different environmental context or otherwise motivating stimuli (see Figure 1D).

Simulations using OPAL showed that it accounts more flexibly for data than simpler reinforcement learning approaches (Collins and Frank 2014) and experimental findings are compatible with some of its main assumptions. For example, pharmacological enhancement of indirect pathway activity results in avoidance learning (Nunes et al. 2010) confirming the suggested role of this BG pathway. Moreover, specific D_1 or D_2 MSN optogenetic stimulation proved to enhance or diminish action values (Tai et al. 2012) and likewise confirm the roles of the direct and indirect pathways in approach and avoidance learning (Hikida et al. 2010; Kravitz et al. 2012). Enhancing DA function restores cognitive motivation in Parkinson´s disease (McGuigan et al. 2019). Maia and Frank (2017) further show that, under specific assumptions of DAergic disturbances, OPAL architecture can account for positive and negative symptoms in schizophrenia. The authors argue that spontaneous task/stimulus-irrelevant DA transients might induce learning of irrelevant stimulus-response associations which in consequence can be exaggerated by excessive DA during choice. Further, decreased stimulus relevant DA transients might contribute to a lack of motivation and so-called negative symptoms [for further details see: (Maia and Frank 2017)].

These ideas are also compatible with research on the role of striatal DA in decisions that require cognitive effort. While cognitive control is generally associated with the maintenance of information in working memory and prefrontal control processes (Cools 2008), the willingness to control or invest in effort for control processes fundamentally requires motivation (Haber and Knutson 2010). Moreover fast striatal gating processes control which information

represented in cortex is updated or maintained (van Schouwenburg et al. 2010). Thus, Westbrook and colleagues argue that striatal DA and its known role in reward and motivation are crucial for understanding the willingness to exert cognitive effort (Westbrook and Braver 2016; Westbrook et al. 2020; Westbrook et al. 2021). They tackle this idea with a task that requires to allocate effort and find that participants with low (dorsal) striatal DA synthesis capacity were less likely to invest in cognitive effort. However, methylphenidate or a selective D_2 antagonist (presynaptic mechanism) restored motivation especially in participants with low baseline synthesis capacity (Westbrook et al. 2020). Using the Drift Diffusion Model (DDM), Westbrook and colleagues demonstrate that DA shifts attention (early in the decision process) to benefits (reward information) away from costs (effort information). In other words, increased baseline synthesis capacity or methylphenidate/sulpiride increase the weight of benefits on evidence accumulation (drift-rate) and therefore increase the willingness to spend cognitive resources. Interestingly, these findings are in line with an motivational effect of DA during choice (see Figure 1D above) and with the perspective that the direct (GO) and indirect (NO-GO) pathway qualitatively signal benefits and costs of actions (Westbrook et al. 2021).

Therefore the effect of DA on cognitive effort is compatible with a dual-pathway model where DA emphasizes processing in the direct D_1 pathway and suppresses processing in the indirect D_2 pathway (Frank and O'Reilly 2006; Collins and Frank 2014; Maia and Frank 2017). Note, that other theories are also compatible with these findings, i.e. a general effect of DA on the willingness to spend energy resources (Berke 2018) or the idea that DAs basically modulates reward sensitivity (FitzGerald et al. 2015), however, these are less specific in terms of the exact mechanisms.

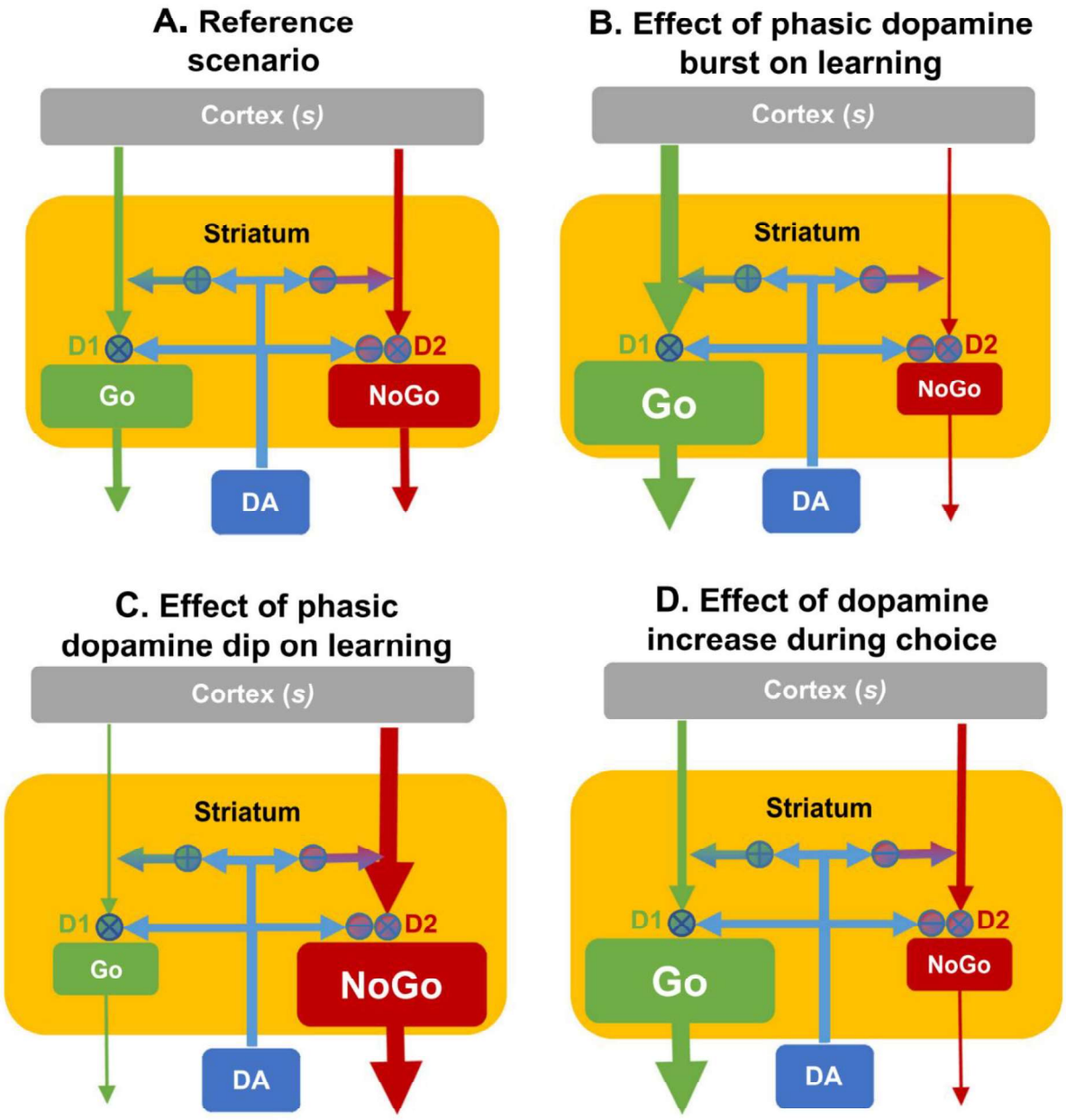


Figure 4. The effect of dopamine (DA) on learning and motivation. Adapted from Maia and Frank (2017). A, reference scenario (without prior learning). B, learning scenario where unpredicted reward resulted in a phasic response of DA neurons (positive prediction error). C, learning scenario where predicted but omitted reward resulted in a dip of DA neuron activity (negative prediction error). D, scenario of increased DA during decision-making/choice reflecting motivational effects of DA biasing approach (benefits) vs. avoidance (costs).

Study Specific Summaries

The overall goal of the studies presented in this dissertation is to improve our understanding of DA-associated changes in intertemporal preferences. Understanding these dopamine-mediated relationships is essential to our understanding of the continuous dimensions of human functioning and the promise of the RDoC framework (U.S. Department of Health and Human Services, National Institutes of 2016). *Study 1* in this dissertation investigates DAergic modulation of intertemporal choice in healthy adult participants using the DA D_2-receptor antagonist haloperidol and state-of-the-art computational approaches to further decompose the decision-process (see methods section below). *Study 2* takes behavioral testing beyond the lab into real-life environments and assesses the effects of addiction related environments on intertemporal preferences and MB RL in regular slot machine gamblers. In *Study 3* we examine whether patients with Tourette Syndrome (TS) show aberrations in intertemporal choice. This is of particular interest because TS is associated with reward sensitivity and disturbances in DA neurotransmission. In *Study 4* we investigate short- and long-term stability of intertemporal preferences as a function of acute and chronic deep brain stimulation (DBS) in a cohort of OCD-patients. OCD is associated with pathological activity in CSTC loops and the NAcc is a central hub of the reward circuit. In what follows I will further introduce specific background information on each work.

Study 1: "Dopaminergic Modulation of Human Intertemporal Choice: A Diffusion Model Analysis Using the D2-Receptor Antagonist Haloperidol"

DAergic modulation of frontostriatal-circuits (CSTC loops) is associated with reward learning, motivation and is implicated in cost-benefit decision-making (see above). Reduced discounting has especially been observed under moderate doses of amphetamine (Wit 2002) and D_2-receptor antagonists (Weber et al. 2016; Arrondo et al. 2015). Note, D_2-receptor antagonists in small doses are likely to increase DA neurotransmission via a presynaptic mechanism. In contrast Pine and colleagues (2010) did not find an effect of the D_2-receptor antagonist haloperidol but observed increased discounting following administration of L-DOPA (Note, this study consisted of a small sample size [n = 13]). Using a greater sample than (Pine et al. 2010) and state-of-the-art computational methods, *Study 1* of this dissertation tries to resolve and clarify the effects of haloperidol. We therefore examined the effect of a single (2 mg) dose of the D_2-receptor antagonist to healthy human participants. Further, our approach extends previous studies via applying state-of-the-art temporal discounting drift-diffusion models (see methods below) to further evaluate the effects of haloperidol on individual decision components. Moreover our approach, using two different magnitude conditions, allows us to

examine the effects of DA on modulating the magnitude effect. The magnitude effect describes the phenomenon that larger rewards are discounted less than smaller rewards (Ballard et al. 2017). Results are discussed with respect to the effects of haloperidol on decision components and the literature on DA and intertemporal choice in general.

Study 2 "Gambling Environment Exposure Increases Temporal Discounting but Improves Model-Based Control in Regular Slot-Machine Gamblers"

Substance use disorders and behavioral addictions are both associated with increased discounting (Lempert et al. 2019; Lempert and Phelps 2016; Bickel et al. 2019). Incentive sensitization theory (Robinson and Berridge 1993; Robinson and Berridge 2008) provides a theoretical framework that links contextual effects, i.e. addiction-related cues to a highly sensitized DA system (see *Incentive salience* above). It thus provides a theoretical framework that renders addiction related environments as powerful triggers of craving and their role in relapse. One example for a highly context dependent addiction with huge societal impact is gambling disorder. Continuous gambling often negatively impacts personal finances, work, relationships and mental health (Blaszczynski and Nower 2002; Muggleton et al. 2021). Interestingly, steep discounting has been consistently observed in substance use disorders and gambling disorder (Reynolds 2006; Bickel et al. 2012; MacKillop et al. 2011; Bickel et al. 2019). One further candidate that might contribute to a range of psychiatric conditions are impairments in MB control during RL (Daw 2011). While MF control operates on stimulus-response associations (Balleine and O'Doherty 2010; Doll et al. 2012; Daw 2011; Voon et al. 2017), MB control refers to computationally more expensive goal-directed strategies that utilize models of the environment. Such reinforcement learning strategies are typically assessed via the so-called Two-Step task (Daw 2011). Contextual modulation of intertemporal choice has been consistently observed in laboratory tasks that include gambling-related cues (Miedl et al. 2014; Genauck et al. 2020; Dale et al. 2019). Further, one study confirms these findings in real-life gambling environments (Dixon et al. 2006). While there is evidence that MB reinforcement learning might be reduced in gambling disorder (Wyckmans et al. 2019) it is unclear if this process is under contextual control. In *Study 2* of this dissertation we therefore use the rare opportunity to investigate the impact of real-life gambling environments on both of these computational markers of addiction. We extend previous findings of contextual modulation of intertemporal choice via an assessment and analysis of important predictors of gambling behavior, i.e. gambling related cognitive distortions and a gambling severity compound score. Both constructs are analyzed using state-of-the-art hybrid temporal discounting and RL DDMs and results are discusses with respect to neural models of addiction.

Study 3: “Temporal Discounting in Adolescents and Adults with Tourette Syndrome”

TS is a neuropsychiatric disorder with a typical onset during childhood. TS patients are characterized by so called tics, i.e. involuntary eye blinking or muscle concentrations, phonic repetitive sounds and others (Leckman 2002; M. M. Robertson 2012). TS symptom severity generally improves at the end of adolescence in around 80% of patients (Coffey et al. 2004; Bloch and Leckman 2009). However the rate of comorbidities like impulsive control disorders, attention-deficit/hyperactive disorder (ADHD) or obsessive compulsive disorder (OCD), are relatively high and only absent in around 25% of participants (Robertson 2012; Hirschtritt et al. 2015; Groth et al. 2017). Clinical and neuroscientific studies both highlight neurological aberrations within the CSTC-loops (Baldermann et al. 2016; Dwyer 2018) especially with respect to DA that strongly modulates these circuits (Frank and O'Reilly 2006; Denys et al. 2013). Theories on the developmental underpinnings of TS range from DA receptor super-sensitivity (Singer 2013) over an imbalance in tonic-phasic DA function to a presynaptic DA dysfunction (Buse et al. 2013; Singer et al. 2002) and DA hyperinnervation (Buse et al. 2013; Maia and Conceição 2018). DA in fronto-striatal circuits plays a key role in both motor control (Smith et al. 2018; Canário et al. 2019) and choice impulsivity (Pine et al. 2010; Weber et al. 2016; Freund et al. 2019). TS has already been associated with aberrations in reward sensitivity in RL tasks (Palminteri and Pessiglione 2013; Kéri et al. 2002). However, whether choice impulsivity is impaired in TS remains ambiguous. To date, only one recent study reports slightly increased choice impulsivity in adolescent TS patients (Vicario et al. 2020). *Study 3* in this dissertation therefore compares intertemporal choice in two cohorts of TS patients (adolescents and adults) with matched healthy controls. Said differently, we examine whether TS pathophysiology is associated with changes in temporal discounting. Our results are likewise discussed with respect to neural models of temporal discounting, DAergic alterations in Tourette syndrome and the developmental trajectory of cognitive control.

Study 4 “Chronic Deep Brain Stimulation of the Human NAcc Region Disrupts the Stability of Intertemporal Preferences”

Obsessive-compulsive disorder (OCD) is a chronic disorder marked by either compulsions, obsessions or both and has a life time prevalence around 2-3 % (Björgvinsson *et al.*, 2007; Ruscio *et al.*, 2010). Patients with OCD suffer from intrusive aversive thoughts (obsessions) and ritualistic behaviors (compulsions). Around 10% of patients with OCD do not respond to conventional therapy (cognitive-behavioral therapy, serotonin reuptake inhibitors) and are considered therapy refractory (Milad and Rauch, 2012). The underlying neurobiology of OCD is not fully understood. Converging evidence suggests a CSTC dysfunction and

especially aberrations in orbitofrontal and VS activity (Modell et al. 1989; Whiteside et al. 2004; Chamberlain et al. 2008; Milad and Rauch 2012; Robbins et al. 2019). While traditional research viewed impulsivity and compulsivity as opposing concepts, new perspectives challenge this view (e.g. Robbins *et al.*, 2012). In patients with OCD, poor treatment outcome has been associated with heightened self-reported impulsivity and a likewise high prevalence of impulsive disorders (Fontenelle et al. 2005). Those findings lead to the assumption that both compulsivity and impulsivity contribute to OCD (Robbins et al. 2012; Kashyap et al. 2012). However, studies on choice impulsivity in patients with OCD have shown inconsistent results. For instance, a study by Sohn and colleagues (2014) patients with OCD discounted significantly more than controls (Sohn et al. 2014), whereas other studies found no difference in discounting behavior between patients with OCD and controls (Pinto et al. 2014; Steinglass et al. 2017; Carlisi et al. 2017). Thus it is not clear if OCD is generally characterized by abnormalities in delay discounting. When conventional therapy is ineffective, deep brain stimulation (DBS) can be considered as a treatment option for patients with OCD. DBS is a neuromodulation procedure that relies on the implantation of electrodes in subcortical structures (Cleary et al. 2015). DBS is widely used with high success rates in areas of movement disorders (Larson 2014). For patients with OCD, DBS targeted for the anterior limb of the internal capsule and the NAcc (ALIC/NAcc) area leads to symptom reduction in about 40 % of therapy refractory patients (Denys et al. 2010; Kohl et al. 2014; Alonso et al. 2015). Thus, hypotheses that justify the application of DBS for psychiatric patients (including OCD) are based on the assumption that DBS modulates CSTC function (Figee et al. 2013; Wu et al. 2020), and in our case directly in regions at the heart of the reward circuit (i.e. the NAcc).

While some studies showed no effect of acute subthalamic nucleus (Seinstra et al. 2016; Aiello et al. 2019) and NAcc (Peisker et al. 2018) stimulation protocols it is especially unclear if chronic long-term DBS can modulate intertemporal choice. In the context of a DBS treatment-efficacy study (Huys et al. 2019) we examined whether phasic or long-term effects of chronic DBS to the ALIC/ NAcc area are associated with changes in intertemporal preferences in OCD patients.

Methods

Computational Modelling of Intertemporal Choice

Over the years scientists developed multiple methods to quantify intertemporal preferences. Choice proportions, for instance, are a simple ratio of how often specific options, i.e. the SS or LL rewards are chosen. So-called points of indifference capture the specific value were preferences do switch from one choice option (i.e. SS) to the other (i.e. LL). One would therefore fit a sigmoid function into delay-specific choices between rewards of varying magnitude. The inflection point of this sigmoid function is qualitatively equivalent to the point of indifference, the amount where the decision-maker becomes indifferent between the SS and LL reward (FitzGerald et al. 2009).

Connecting multiple indifference points in a graph results in a continuous function. Integrating the area under this empirical discounting curve (AUC) provides a model-free estimate of intertemporal choice. One would therefore normalize the subjective values (the points of indifference) and delays (proportions of the maximum delay) and compute the integral (in a step-wise numerical approximation) over the normalized y- and x-axis data points. The resulting AUC would then vary within the boundaries of 0 to 1, whereas a lower value indicates steeper discounting and values next to 1 indicate shallow to no discounting (Myerson et al. 2001). All of these so called “model-free” methods (mentioned above) have in common that they are easy to compute and provide a central tendency without further theoretical assumptions. This rather agnostic approach has the advantage to provide a neutral measurement without a theoretical framework, but also lacks further explanation and does not yield specific testable predictions.

Computational models are often build on the promise for further explanatory power. That is, they try to formalize testable assumptions (theoretical frameworks), ideally match underlying processes (so-called process models) and manage to predict data not yet observed (Farrell and Lewandowsky 2018). Early computational approaches to delay discounting modelled the relationship of delay and reinforcer using an exponential function within the framework of discounted utility (Samuelson 1937). In its rewritten form this function can be expressed as follows $SV = A * \exp(-kD)$. The subjective value (SV) of an amount (A) is devalued via an exponential function weighting delay (D) by an individual discounting parameter (k). A lower value of the discount-rate indicates shallower discounting, whereas a higher value of the discount-rate reflects steeper discounting. However in practice, experimental validation and comparison of computational models (McKerchar et al. 2009; Ainslie 1974) led to the notion, that the exponential model often fails to accurately describe

empirical data. In general people tend to discount more (are increasingly impatient) when choosing between options in the near future, but show less discounting when options are further in the future. For example people might prefer 25 € in three months over 20 € in two months, but 20 € today over 25 € in a month. This phenomenon is referred to as preference reversals and date back to observations in pigeons around 50 years ago (Ainslie 1974; Ainslie 1975). Building on these findings and further validation of experimental data the hyperbolic model SV = A / (1 + kD) proved superior to fit delay discounting data (Mazur and Coe 1987). The main mathematical difference is that the exponential model assumes a constant rate of devaluation, i.e. a neutral time preference, and the hyperbolic model assumes steeper discounting of the near future. Nowadays parameter estimates of the discount-rate parameter derived from the hyperbolic model are determined the standard procedure to quantify tendencies in intertemporal choice/delay discounting (Mazur and Coe 1987; Green and Myerson 2004; Peters and Büchel 2011). Formally a simple hyperbolic function describes how values change as a function of delay (Eq. 1).

$$SV(LL_t) = \frac{A_t}{1 + \mathrm{k} * D_t} \quad (Eq. 1)$$

Here, A_t is the numerical reward amount of the LL option on trial *t*. A trial is simply one choice between a SS and LL option. D_t is the corresponding delay in days on that trial. The model has only one free parameters, the hyperbolic discounting rate *k*. A standard softmax action-selection rule (Sutton and Barto 1998) can then be used to model choice probabilities (in this case the probability for choosing the LL reward) as a sigmoid function of value differences, that is the difference between the SS and discounted LL in a trial-wise fashion.

$$P(LL)_t = \frac{\exp\left(\beta * SV(LL_t)\right)}{\exp\left(\beta * SV(SS_t)\right) + \exp\left(\beta * SV(LL_t)\right)} \quad (Eq. 2)$$

The softmax model thus accounts for two things. First, it maps the discounted subjective LL and SS values onto the observed decisions. Secondly, the β parameter models choice stochasticity, that is as β increases choices become more dependent on the option values. As β approaches zero, choices become purely random (Sutton and Barto 1998). Thus, the softmax choice-rule or its β parameter can inform overall tendencies in terms of choice noisiness but do not allow for further insights into the decision process itself.

One recent advancement in modelling value-based decisions is the implementation of sequential sampling models (Pedersen et al. 2017; Fontanesi et al. 2019; Peters and D'Esposito 2020). A key feature of these models, here with focus on the drift diffusion model (DDM) is that they explain both choices and reaction time (RT) distributions (Ratcliff and McKoon 2008). According to Forstmann et al. (2016) sequential sampling models date back to the early days of probability theory where much effort was attributed to problems devoted to gambling. For example, early random walk models described the process when gamblers with different starting capitals and win probabilities played against each other multiple times. The resulting random walk only depended on the starting capital and win probabilities for each player and is theoretically equal to an accumulation of evidence over time (Forstmann et al. 2016; Feller 2003). The evidence accumulation in a DDM itself, however, historically stems from modelling Brownian motion, that is the random fluctuations of physical particles given a specific temperature. The first observation of this phenomenon date back to 1827 where botanist Robert Brown described the random like behavior of pollen in water (Brown 1828). Later in 1905 Einstein referred to Brown in his work on the molecular-kinetic theory of heat where he statistically modelled Brownian motion for the first time (Einstein 1905). This work was later described as one of his most important contribution to modern science indirectly proving the existence of atoms (Mazo 2009). Another important step to modern sequential sampling models was the work of Norbert Wiener in the 1920s. Wiener worked on the mathematical properties of noisy stochastic processes and his contributions were later fundamental to prove the existence of the "Wiener Process" using methods from probability theory (Nelson 1967).

In 1978 Roger Ratcliff built on these findings to propose the first mathematical diffusion model in psychology capable of predicting RT distributions in the framework of memory retrieval (Ratcliff 1978). Since then the DDM has been successfully implemented in the broad field of psychology and cognitive neuroscience (Forstmann et al. 2016). In detail, the DDM models participant´s choices as two distinct response boundaries, each reflecting the whole RT distributions corresponding to each choice-option (see Figure 5). A choice is initiated if the cumulated rate of evidence crosses one of two response-boundaries. In its simple form the standard DDM has four parameters. First, the drift-rate (average speed of evidence accumulation) parameter (v) models the rate at which evidence for one of both response boundaries is accumulated (see Figure 5). Second, the starting point of this process (z). This so-called bias can account for a bias to either response boundary. Third, the distance between both response boundaries is captured by boundary separation parameter (a). This parameter adjusts the speed-accuracy trade-off, i.e. the amount of evidence that is needed until one of the two

alternative boundaries is reached. Fourth, the non-decision time (t0) captures the time for stimulus encoding (e.g. sensory processing of option space) and motor preparation until the evidence accumulation process starts. While hyperbolic discounting in combination with softmax-action selection maps the different values (SS and subjective LL) onto participants choices, the DDM in its standard form maps RTs onto choice boundaries, but ignores further value information. In other words, standard DDM thus allows for the studying of individual RT components of this decision process but ignores trial-wise value differences.

In *Study 1* and *Study 2* within this dissertation we therefore aimed to combine the strength of the DDM´s ability in capturing individual components of the decision process and trial-wise option valuation of hyperbolic discounting. To do this we build on recent successful implementations of the DDM in RL (Pedersen et al. 2017; Fontanesi et al. 2019; Shahar et al. 2019; Miletić et al. 2020). In this combined model, the two boundaries correspond to the RT distributions of the choices (stimulus coding) of the SS and subjective LL (via hyperbolic discounting) rewards, whereas the drift-rate is modelled as a linear (Pedersen et al. 2017) or sigmoid (Fontanesi et al. 2019) trial-wise function of value differences (see Eq. 3 and Eq. 4 for the linear version, and the methods sections on *Study 1* and *Study 2* for other versions). Using a DDM as choice-rule has some major advantages over the softmax choice-rule introduced above. Instead of simply mapping value-differences (differences of SS and subjective LL values) onto binary choices it allows for further decomposition with the advantage of studying each component individually while taking option valuation into account.

$$v_t = v_{coeff} * \left(SV(LL_t) - SV(SS_t)\right) \quad (Eq. 3)$$

$$RT_t \sim wfpt(\alpha\,, \tau\,, z\,, v_t) \quad (Eq. 4)$$

Value-based diffusion models are a promising candidate for a further approximation to a real process model in value based choice, i.e. a model that not only manages to explain patterns in data but also captures underlying processes (Farrell and Lewandowsky 2018). This point of view is also supported by neurobiological evidence. For example, neural correlates of gradual evidence accumulation in monkeys (Selen et al. 2012; Bichot et al. 2001) and humans (Selen et al. 2012; Shadlen and Kiani 2013) support some of the DDM´s assumptions like for example, evidence accumulation under time-pressure (Shadlen and Kiani 2013).

In *Study 2* we additionally model behavior in a reinforcement learning (RL) task. This so-called Two Step Task (Daw 2011) dissociates model-based (MB) from model-free (MF) RL strategies. MB control refers to computationally more expensive goal-directed strategies that rely on a probabilistic map of the environment, contrasting with MF control that operates on simpler stimulus-response associations (Balleine and O'Doherty 2010; Daw 2011; Voon et al. 2017). Behavior in this task is modelled using a Q-learning algorithm that estimates both MF and MB Q-values (state-action values). The standard-approach to analyze this task is the hybrid RL model introduced by Daw (2011). The model updates MF state-action values through prediction errors. MB state-action values are then computed from the transition and reward estimates using the Bellman Equation. This approach then models the strength of MF and MB RL strategies during task performance via a softmax function. Like in our analysis of temporal discounting we additionally replaced the softmax function with a DDM choice rule (Shahar et al. 2019). Further in deepth explanations on task and computational methods are available in the methods section of *Study 2*.

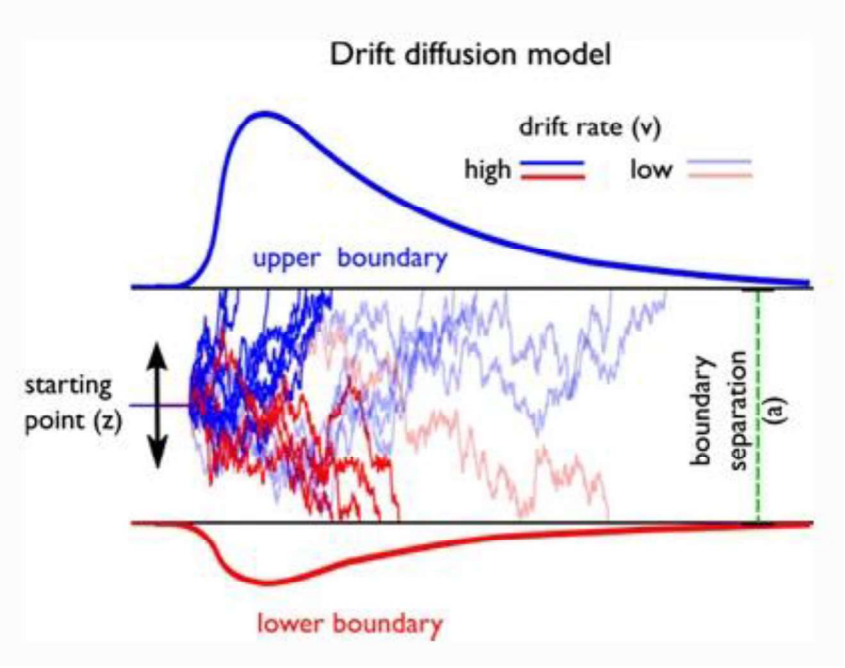


Figure 5 Illustration of the drift diffusion model. Adapted from (Pedersen et al. 2017).

Hierarchical Bayesian Parameter Estimation

This chapter gives a brief overview of cognitive modelling implemented in a hierarchical Bayesian framework. Cognitive modelling in general provides a mathematical description that preferably describes an underlying process (Farrell and Lewandowsky 2018). A Bayesian estimation scheme implies that the parameters describing this process are updated (given the data) in the sense of probability distributions. Bayesian frameworks have the advantage of estimating the entire posterior distribution of parameters. This allows to quantify the degree of belief in these values via a probability estimate, i.e. the probability density of a continuous parameter (Wagenmakers et al. 2018). In the special case of hierarchical models the hierarchy can involve multiple levels to reflect some kind of hierarchical dependencies within the data, that could be a different experimental manipulation or a specific environmental context. In consequence subject level parameters that describe individual behavior, i.e. the discounting of rewards, are drawn from separate hyperparameter distributions depending on the specific context or group. The subject- and group level parameters then form a joint parameter space, are estimated simultaneously and mutually inform each other. In consequence group comparisons can be intuitively implemented via a comparison of overlapping probability density estimates of group- or condition-specific hyperparameter distributions and their corresponding highest density invervalls (Kruschke 2011; Baldwin and Fellingham 2013; Farrell and Lewandowsky 2018; Wagenmakers et al. 2018).

In a first step one needs to define a model that maps the observed data (experiment or simulation) to outcomes via a likelihood function in the setting of probability distributions. Such a model for example consists of a hyperbolic function like Eq.1 in combination with a softmax choice-rule Eq. 2 (Sutton and Barto 1998) . In terms of intertemporal choice one main parameter of interest is the individual discount-rate *k*. However, we cannot directly observe this parameter. Instead our model defines the relationship of a suggested underlying process (hyperbolic devaluation) and a function, i.e. softmax to map the estimated values onto choice probabilities and in consequence to binary decision-outcomes. Before one can begin to numerically solve this one needs to define a plausible range of prior values (defined as probability distributions) that constrain our parameters of interest (discount-rate and softmax-temperature; see above) in a plausible way. If there is no prior knowledge a uniform prior distribution, that assigns equal probability to each possible parameter in a plausible range is appropriate (e.g. that let the discount-rate account for both the possibility of very steep discounting [only SS choices] or no discounting [only LL choices]). Bayes Rule (Eq. 5) for conditional probabilities formalizes the relationship of our priors $p(\theta)$, the likelihood of the data

given the parameters p(D|θ) and the posterior that quantifies the probability of a parameter given that we have seen the data p($\theta|D$). In other words, the observed data is used to update our prior information, resulting in an updated posterior belief. The term $p(D)$, the so called evidence/general probability of the data or the marginal likelihood, serves as a normalizing constant such that the posterior distribution is a proper probability density function where the integral sums to 1 (Wagenmakers et al. 2018). As this is a scaling parameter a rule of thumb states that the posterior is equivalent to the prior times the likelihood (Eq. 6).

There are multiple methods available for solving or fitting these Bayesian or hierarchical Bayesian models. So-called "Monte-Carlo Markov Chain" (MCMC) algorithms implemented in programs like JAGS (Plummer 2003) or STAN (Stan Development Team 2020) generate a conditioned walk through parameter space. Even though the details of both algorithms (Metropolis Hastings [Stan] or Gibbs Sampling [JAGS]) differ, both result in a chained walk that visits those values with higher probability more often, than those with lower or no probability yielding a converging representative sample of the true posterior given the model and data (Kruschke 2015). One major advantage of models implemented in these programming languages is that functions are easily replace- and extendable allowing for fast modification and model comparison. For example, one can easily replace softmax action-selection with a DDM to further account for response time distributions. Furthermore, parameters in hierarchical models mutually inform and constrain each other (partial pooling), such that meaningful estimates can be derived even with limited data per subject. Moreover, the decision rule can accept the null value and provides an effect size estimate (Kruschke 2011, 2013, 2015).

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \quad \text{(Eq. 5)}$$

$$posterior \sim prior * likelihood \quad \text{(Eq. 6)}$$

Behavioral/Cognitive

Dopaminergic Modulation of Human Intertemporal Choice: A Diffusion Model Analysis Using the D2-Receptor Antagonist Haloperidol

Ben Wagner,[1] Mareike Clos,[2] Tobias Sommer,[2] and Jan Peters[1,2]
[1]Department of Psychology, Biological Psychology, University of Cologne, Cologne 50969, Germany, and [2]Department of Systems Neuroscience, University-Medical-Center Hamburg-Eppendorf, Hamburg 20246, Germany

The neurotransmitter dopamine is implicated in diverse functions, including reward processing, reinforcement learning, and cognitive control. The tendency to discount future rewards over time has long been discussed in the context of potential dopaminergic modulation. Here we examined the effect of a single dose of the D2 receptor antagonist haloperidol (2 mg) on temporal discounting in healthy female and male human participants. Our approach extends previous pharmacological studies in two ways. First, we applied combined temporal discounting drift diffusion models to examine choice dynamics. Second, we examined dopaminergic modulation of reward magnitude effects on temporal discounting. Hierarchical Bayesian parameter estimation revealed that the data were best accounted for by a temporal discounting drift diffusion model with nonlinear trialwise drift rate scaling. This model showed good parameter recovery, and posterior predictive checks revealed that it accurately reproduced the relationship between decision conflict and response times in individual participants. We observed reduced temporal discounting and substantially faster nondecision times under haloperidol compared with placebo. Discounting was steeper for low versus high reward magnitudes, but this effect was largely unaffected by haloperidol. Results were corroborated by model-free analyses and modeling via more standard approaches. We previously reported elevated caudate activation under haloperidol in this sample of participants, supporting the idea that haloperidol elevated dopamine neurotransmission (e.g., by blocking inhibitory feedback via presynaptic D2 auto-receptors). The present results reveal that this is associated with an augmentation of both lower-level (nondecision time) and higher-level (temporal discounting) components of the decision process.

Key words: **computational modeling; decision making; dopamine; haloperidol; intertemporal choice; pharmacology**

Significance Statement

Dopamine is implicated in reward processing, reinforcement learning, and cognitive control. Here we examined the effects of a single dose of the D2 receptor antagonist haloperidol on temporal discounting and choice dynamics during the decision process. We extend previous studies by applying computational modeling using the drift diffusion model, which revealed that haloperidol reduced the nondecision time and reduced impulsive choice compared with placebo. These findings are compatible with a haloperidol-induced increase in striatal dopamine (e.g., because of a presynaptic mechanism). Our data provide novel insights into the contributions of dopamine to value-based decision-making and highlight how comprehensive model-based analyses using sequential sampling models can inform the effects of pharmacological modulation on choice processes.

Introduction

Future rewards are discounted in value (Peters and Büchel, 2011) such that humans and many animals prefer smaller-sooner (SS) rewards over larger-but-later (LL) rewards (temporal discounting). Steep discounting of reward value is associated with a range of maladaptive behaviors ranging from substance use disorders (Bickel et al., 2014), attention-deficit hyperactivity disorder (Jackson and MacKillop, 2016), and obesity (Amlung et al., 2016) to behavioral addictions, such as gambling disorder (Wiehler and Peters, 2015). Temporal discounting has thus been suggested to constitute a transdiagnostic process (Amlung et al.,

Received Mar. 11, 2020; revised Sep. 3, 2020; accepted Sep. 9, 2020.
Author contributions: B.W. and J.P. contributed unpublished reagents/analytic tools; B.W. and J.P. analyzed data; B.W. wrote the first draft of the paper; B.W. and J.P. wrote the paper; M.C., T.S., and J.P. designed research; M.C. performed research; M.C. and T.S. edited the paper.
This work was supported by Deutsche Forschungsgemeinschaft PE1627/5-1 to J.P. and SO952/3-1 to T.S.
The authors declare no competing financial interests.
Correspondence should be addressed to Ben Wagner at ben.jonathan.wagner@uni-koeln.de or Jan Peters at jan.peters@uni-koeln.de.
https://doi.org/10.1523/JNEUROSCI.0592-20.2020

Table 1. Demographic and working memory data[a]

	Placebo	Haloperidol	Group comparison
Age (yr)	24.4 ± 3.4	23.3 ± 2.5	$t_{(45,614)} = 1.40$, $p = 0.17$
Sex (M/F)	7/19	6/17	$\chi^2_{(1)} = 0.001$, $p = 1$
WM baseline (*z* score)	−0.0453 ± 0.665	0.0943 ± 0.556	$t_{(46,826)} = -0.80$, $p = 0.43$
Weight (kg) (M/F)	70.7 ± 3.39/63.5 ± 3.19	80.5 ± 2.80/62.5 ± 2.32	$t_{(36,702)} = -0.68$, $p = 0.50$

[a]Data are mean ± SD.

2019; Lempert et al., 2019) with relevance for many psychiatric conditions.

Dopamine (DA) plays a central role in addiction (Robinson and Berridge, 1993). In rodents, reductions versus moderate increases in DA transmission led to increases and decreases in discounting, whereas the corresponding human literature is small and more heterogeneous (D'Amour-Horvat and Leyton, 2014). For example, de Wit et al. (2002) found that acute administration of D-amphetamine decreased impulsivity, such that temporal discounting was reduced under D-amphetamine. However, a later study did not replicate this effect (Acheson and de Wit, 2008). Administration of the D2/D3 receptor agonist pramipexole did not affect measures of impulsivity in another study ($n = 10$) from the same group (Hamidovic et al., 2008). In contrast, Pine et al. (2010) observed increased temporal discounting following administration of the catecholamine precursor L-DOPA compared with placebo in healthy control participants ($n = 13$), while the D2-receptor antagonist haloperidol did not modulate discounting. In a recent within-subjects study using L-DOPA in a substantially larger sample ($n = 87$), there was no overall effect on temporal discounting (Petzold et al., 2019). Rather, effects depended on baseline impulsivity, which the authors interpreted in the context of the inverted-U-model of DA effects on cognitive control functions (Cools and D'Esposito, 2011). Two recent studies have reported a reduction in discounting following administration of the selective D2/D3-receptor antagonist amisulpride (Weber et al., 2016) as well as the D2 receptor antagonist metoclopramide (Arrondo et al., 2015). Although the latter is primarily used clinically for its peripheral effects, it can pass the blood-brain barrier and act centrally (Shakhatreh et al., 2019).

A similar heterogeneity is evident when considering model-based reinforcement learning (RL) (Doll et al., 2012), which in some studies (Shenhav et al., 2017), but not others (Solway et al., 2017), was associated with reduced temporal discounting. However, in contrast to temporal discounting (see above), L-DOPA instead increased reliance on model-based RL in healthy controls (Wunderlich et al., 2012) and Parkinson's disease patients (Sharp et al., 2016). Notably, this overall effect was not observed in a recent study in a substantially larger sample ($n = 65$) (Kroemer et al., 2019). Here, increased model-based RL under L-DOPA was restricted to participants with high working memory capacity.

One well-replicated behavioral effect in temporal discounting (magnitude effect) refers to the observation that the rate of temporal discounting decreases with increasing reward magnitude (Green et al., 1997). In humans, this effect depends on lateral PFC processing (Ballard et al., 2017); and in rodents, D-amphetamine effects on temporal discounting are more pronounced for large-magnitude conditions (Krebs et al., 2016). However, it is unclear whether DA impacts the magnitude effect in humans.

In the present study, we examined these processes using a between-subjects double-blind placebo-controlled pharmacological study with the D2-receptor antagonist haloperidol (2 mg). We previously reported increased dorsal striatal activation under haloperidol versus placebo in these participants (Clos et al., 2019a,b), compatible with a predominantly presynaptic effect of haloperidol that increases striatal dopaminergic signaling. Importantly, we extended previous pharmacological studies by applying a temporal discounting modeling framework based on a combination of discounting models with the drift diffusion model (DDM) (Pedersen et al., 2017; Fontanesi et al., 2019; Shahar et al., 2019; Peters and D'Esposito, 2020), allowing us to comprehensively examine drug effects on response time (RT) components related to both valuation and non–valuation-related processes.

Materials and Methods

Participants

Fifty-four healthy participants were initially enrolled in the study. Participants were screened by a physician for current diseases and current intake of prescription drugs or drugs of abuse. All participants were presently in good health and had no history of neurologic or psychiatric disorder with no current intake of prescription medication. Only healthy subjects were allowed to participate. Twenty-seven participants were randomly assigned to each group (placebo/haloperidol). Two participants from the haloperidol group did not complete the temporal discounting task. Technical problems led to working memory data loss from 4 participants (3 from the haloperidol, 1 from the placebo group), but these participants were still included in the temporal discounting data analysis.

Following filtering of RTs (see below; the fastest and slowest 2.5% of trials were excluded per participant), we examined the individual RT histograms for each subject (see Extended Data Fig. 1-1). This revealed that, even after filtering, the 3 participants with the fastest minimum RTs (2 from the haloperidol group and 1 from the placebo group) still showed implausibly fast responses on a number of trials (minimum RTs of 2, 2, and 234 ms, in Subjects 24, 25, and 41, respectively) such that the minimum RTs were substantially faster than those in the remaining participants (all min(RT) *z* scores of −2.04, −2.04, and −1.7; see Extended Data Fig. 1-2). These subjects were therefore excluded from further modeling.

We verified that there were no significant differences in demographic background in terms of age or baseline working memory capacity (Table 1). Potential side effects of the medication were monitored via multiple blood pressure and pulse measurements and evaluated via mood questionnaires. These analyses did not reveal significant group differences in terms of reported mood, side effects, or physiological parameters, as reported in our previous study (Clos et al., 2019b). Before enrollment, participants provided informed written consent, and all study procedures were approved by the local institutional review board (Hamburg Board of Physicians).

Experimental design

General procedure. The study consisted of two testing sessions performed on separate days. On the first day (T0), participants completed a background screening and a set of working memory tasks (see below). On the second day (T1), participants received either placebo or haloperidol (2 mg). In line with the pharmacokinetics of haloperidol (Franken et al., 2017), testing on T1 was performed 5 h after drug administration to ensure appropriate plasma levels of haloperidol. During the first 2.5 h, participants were under constant observation, and pulse as well as blood pressure levels were checked 30 min and 2 h after drug administration.

During the waiting period, participants filled out questionnaires on current mood and medication effects. Participants then completed a number of unrelated tasks during an fMRI scanning session (total scan time 2.5 h.). Following scanning, they first completed the temporal discounting task outlined below, followed by a set of working memory tasks (digit span forward and backward, block span forward and backward, complex working memory span) (for detailed results, see Clos et al., 2019b).

Temporal discounting task. Participants performed 210 trials of a temporal discounting task where on each trial they made a choice between an SS reward available immediately and an LL reward. SS and LL rewards were randomly displayed on the left and right sides of the screen, and participants were free to make their choice at any time. For half the trials, the SS reward consisted of 20€; and for the remaining trials, the SS reward was fixed at 100€. These trials were presented randomly intermixed. LL options were computed via all combinations of a set of LL reward amounts (constructed by multiplying the SS reward with [1.01, 1.02, 1.05, 1.10, 1.20, 1.50, 1.80, 2.50, 2, 3, 4, 5, 7, 10, 13]) and LL delays (1, 2, 3, 5, 8, 30, 60 d), yielding 105 trials in total per magnitude condition. As is typically the case for temporal discounting tasks investigating magnitude effects (Green et al., 1997), all choices were hypothetical.

Computational modeling

Temporal discounting model. We applied a simple single-parameter hyperbolic discounting model to describe how value changes as a function of delay (Mazur, 1987; Green and Myerson, 2004) as follows:

$$SV(LL_t) = \frac{A_t}{1+\exp(k + s_k * I_t) * D_t} \quad (1)$$

Here, A_t is the numerical reward amount of the LL option on trial t, D_t is the LL delay in days on trial t, and I_t is an indicator variable that takes on a value of 0 for trials from the large-magnitude condition (SS amount = 100€) data and 1 for trials from the small-magnitude condition (SS amount = 20€). The model has two free parameters: k is the hyperbolic discounting rate from the large-magnitude condition (modeled in log-space) and s_k is a weighting parameter that models the degree of change in discounting for small versus large SS rewards (i.e., higher values in s_k reflect a greater magnitude effect) (Green et al., 1997).

Softmax action selection

Softmax action selection models choice probabilities as a sigmoid function of value differences (Sutton and Barto, 1998) as follows:

$$P(LL)_t = \frac{\exp\big((\beta + s_\beta * I_t) * SV(LL_t)\big)}{\exp\big((\beta + s_\beta * I_t) * SV(SS_t)\big) + \exp\big((\beta + s_\beta * I_t) * SV(LL_t)\big)} \quad (2)$$

Here, SV is the subjective value of the risky reward according to Equation 1 and β is an inverse temperature parameter, modeling choice stochasticity (for $\beta = 0$, choices are random and as β increases, choices become more dependent on the option values). $SV(SS_t)$ was fixed at 100 for the large-magnitude condition and fixed at 20 for the small-magnitude condition. I_t is again the dummy-coded condition regressor, and s_β models the magnitude effect on β.

Temporal discounting DDMs

To more comprehensively examine dopaminergic effects on choice dynamics, we additionally replaced Softmax action selection with a series of DDM-based choice rules. In the DDM, choices arise from a noisy evidence accumulation process that terminates as soon as the accumulated evidence exceeds one of two response boundaries. In the present setting, the upper boundary was defined as selection of the LL option, whereas the lower boundary was defined as selection of the SS option.

RTs for choices of the SS option were multiplied by −1 before model fitting. We furthermore used a percentile-based cutoff, such that, for each participant, the fastest and slowest 2.5% of trials were excluded from the analysis. We then first examined a null model (DDM_0) without any value modulation. Here, the RT on each trial t is distributed according to the Wiener First Passage Time (*wfpt*) as follows:

$$RT_t \sim wfpt(\alpha + s_\alpha * I_t, \tau + s_\tau * I_t, z + s_z * I_t, v + s_v * I_t) \quad (3)$$

The parameter α models the boundary separation (i.e., the amount of evidence required before committing to a decision), τ models the nondecision time (i.e., components of the RT related to motor preparation and stimulus processing), z models the starting point of the evidence accumulation process (i.e., a bias toward one of the response boundaries, with $z > 0.5$ reflecting a bias toward the LL boundary, and $z < 0.5$ reflecting a bias toward the SS boundary), and ν models the rate of evidence accumulation. For each parameter x, we also include a parameter s_x that models the change in that parameter from the high-magnitude (SS = 100€) to the low-magnitude (SS = 20€) condition (coded via the dummy-coded condition regressor I_t).

As in previous work (Pedersen et al., 2017; Fontanesi et al., 2019; Peters and D'Esposito, 2020), we then set up temporal discounting diffusion models by making trialwise drift rates proportional to the difference in subjective values between options. First, we set up a linear modeling scheme (DDM_{lin}) (Pedersen et al., 2017) as follows:

$$v_t = \big(v_{coeff} + s_{v_{coeff}} * I_t\big) * \big(SV(LL_t) - SV(SS_t)\big) \quad (4)$$

Here, the drift rate on trial t is calculated as the scaled value difference between the LL and SS rewards. As noted above, RTs for SS options were multiplied by −1 before model estimation, such that this formulation predicts SS choices whenever SV(SS) > SV(LL) (the trialwise drift rate is negative) and predicts longest RTs for trials with the highest decision conflict (i.e., in the case of SV(SS) = SV(LL) the trialwise drift rate is zero). We next examined a DDM with nonlinear trialwise drift rate scaling (DDM_S) that has recently been reported to account for the value dependency of RTs better than the DDM_{lin} (Fontanesi et al., 2019; Peters and D'Esposito, 2020). In this model, the scaled value difference from Equation 4 is additionally passed through a sigmoid function with asymptote v_{max} as follows:

$$v_t = S\Big[\big(v_{coeff} + s_{v_{coeff}} * I_t\big) * \big(SV(LL_t) - SV(SS_t)\big)\Big] \quad (5)$$

$$S(m) = \frac{2 * (v_{max} + s_{v_{max}} * I_t)}{1 + \exp(-m)} - (v_{max} + s_{v_{max}} * I_t) \quad (6)$$

All parameters, including v_{coeff} and v_{max}, were again allowed to vary according to the reward magnitude condition, such that we included s_x parameters for each parameter x that were multiplied with the dummy-coded condition predictor I_t (see above).

Hierarchical linear regression

Here we used the median posterior log(k) parameter of each participant from the DDM_S model (see above) to compute the discounted values for all LL options. We then computed the trialwise decision conflict as the absolute difference between the subjective value of the LL reward and the corresponding smaller sooner reward. To ensure that the intercept in the regression model corresponds to the RT for the lowest decision conflict and to account for the strongly skewed distribution of value differences, we took the inverse of the absolute difference in SS and discounted LL values in each trial. To further avoid numerical instabilities when taking the inverse of absolute differences < 1 (high conflict, e.g., SV(LL) = 20.10€, SS = 20€), these value differences were capped at 1 before computing the inverse. We then ran a hierarchical linear regression model in JAGS with 1/RT (to account for the skewed RT distribution) as dependent variable and decision conflict (inverse of the absolute value difference) as a predictor.

Statistical analyses

Hierarchical Bayesian models. Models were fit to all trials from all participants using a hierarchical Bayesian modeling approach with

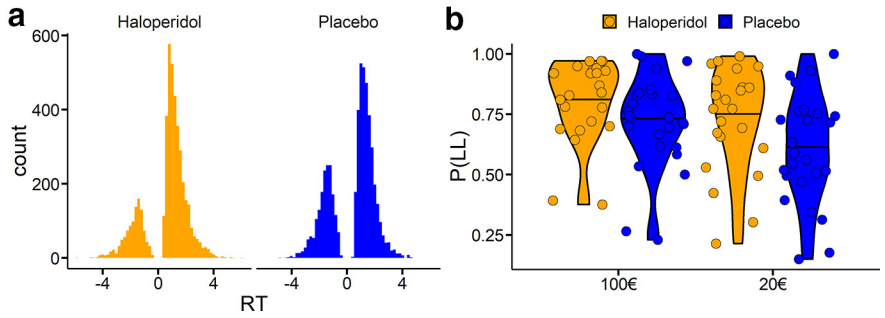


Figure 1. ***a***, Overall RT distributions for the placebo group ($n = 26$) and the haloperidol group ($n = 23$). Negative RTs reflect choices of the SS option, whereas positive RTs reflect choices of the LL option. It can be seen that participants in the placebo group made numerically more SS selections than participants in the haloperidol group. For individual subject RT distributions, see Extended Data Figure 1-1. For minimum RTs following trial filtering, see Extended Data Figure 1-2. ***b***, Proportion of LL choices per group and magnitude condition.

separate group-level distributions for all parameters for the placebo and haloperidol groups. Model fitting was performed using Markov Chain Monte Carlo as implemented in the JAGS software package (Plummer, 2003) (version 4.3) using the Wiener module for JAGS that implements the Wiener First Passage Time (Wabersich and Vandekerckhove, 2014) (see Eq. 3) in combination with R (version 3.4) and the R2Jags package. For group-level means, we used uniform priors defined over numerically plausible parameter ranges (see Code and data availability). For all s_x parameters modeling condition effects on model parameters, we used Gaussian priors with means of 0 and SDs of 2. For group-level precisions, we used γ distributed priors (0.001, 0.001). We initially ran 2 chains with a burn-in period of 900,000 samples and thinning of two. Chain convergence was then assessed via the Gelman-Rubinstein convergence diagnostic $\hat{R}$ and sampling was continued until $1 \leq \hat{R} \leq 1.1$ for all group-level and individual-subject parameters. This occurred after a maximum of 1.3 million samples. For most parameters, $1 \leq \hat{R} \leq 1.01$ (Softmax: all parameters, DDM_0: all parameters, DDM_{lin}: 5 parameters $1.01 \leq \hat{R} \leq 1.1$, DDM_S: 9 parameters $1.01 \leq \hat{R} \leq 1.1$). Relative model comparison was performed via the deviance information criterion (DIC), where lower values reflect a superior fit of the model (Spiegelhalter et al., 2002). A total of 10,000 additional samples were then retained for further analysis. We then show posterior group distributions for all parameters of interest as well as their 85% and 95% highest density intervals (HDIs). For group comparisons, we report Bayes factors (BFs) for directional effects Kass and Raftery, 1995 for the hyperparameter difference distributions of placebo-haloperidol, estimated via kernel density estimation using R (version 4.01) via RStudio (version 1.3) interface. These are computed as the ratio of the integral of the posterior difference distribution from 0 to ∞ versus the integral from 0 to $-\infty$. Using common criteria (Beard et al., 2016), we considered BFs between 1 and 3 as anecdotal evidence, BFs >3 as moderate evidence, and BFs >10 as strong evidence. BFs >30 and >100 were considered as very strong and extreme evidence, respectively, whereas the inverse of these reflect evidence in favor of the opposite hypothesis.

Parameter recovery analyses. To ensure that the parameters underlying the data-generating process could be recovered using our modeling procedures, we performed posterior predictive checks for the best-fitting model (DDM_S). During model estimation, we generated 10,000 datasets simulated from the posterior distribution of the DDM_S. Ten of these simulated datasets were randomly selected and refit with the DDM_S (see previous section) (Fontanesi et al., 2019; Peters and D'Esposito, 2020). Parameter recovery was then assessed in two ways. For group-level parameters, we examined whether the estimated 95% highest posterior density intervals contained the true generating parameters. For subject-level parameters, we examined scatter plots of generating versus estimated single-subject parameters, pooled across all 10 simulations.

Posterior predictive checks. To check whether the best-fitting model indeed captured key aspects of the data, in particular the value dependency for RTs, we performed posterior predictive checks (Peters and D'Esposito, 2020) as follows. For each individual participant, we binned trials into five bins, according to the absolute difference in LL versus SS value ("decision conflict," computed according to each participant's median posterior log(k) parameter from the DDM_S, and separately for the high- and low-magnitude conditions). For each participant and condition, we then plotted the mean observed RTs as a function of decision conflict, as well as the mean RTs across 10,000 datasets simulated from the posterior distributions of the DDM_0, DDM_{lin} and DDM_S.

Code and data availability

Model code is available on the Open Science Framework (https://osf.io/wm7ud/). Raw choice data are available from Zenodo.org (https://doi.org/10.5281/zenodo.4006531) for researchers meeting the criteria for access to confidential data.

Results

Subjective and physiological drug effects

As reported in detail in our previous papers (Clos et al., 2019a,b), there were no significant group differences with respect to reported side effects, subjective mood, heart rate, or blood pressure relative to baseline. Likewise, groups did not differ with respect to the actual and guessed drug condition (haloperidol vs placebo) (Clos et al., 2019b).

Model free analysis of temporal discounting

Figure 1*a* shows the overall RT distributions per group with choices of the LL option coded as positive RTs and choices of the SS option coded as negative RTs. As a model-free measure of temporal discounting, we examined proportions of LL choices as a function of group (placebo vs haloperidol) and condition (100€ vs 20€ reference reward). Raw proportions of LL choices are plotted in Figure 1*b*. ANOVA on arcsine-square-root transformed proportion values with the within-subject factor magnitude (high [100€] vs low [20€] SS reward) and the between-subject factor drug (placebo vs haloperidol) confirmed a significant magnitude effect ($F_{(1,47)} = 96.86$, $p < 0.001$) such that participants overall made more LL selections in the high-magnitude condition. Furthermore, effects of drug ($F_{(1,47)} = 3.47$, $p = 0.068$) and drug × magnitude ($F_{(1,47)} = 3.31$, $p = 0.075$) showed trend-level significance.

Softmax choice rule

First, we analyzed our data using a standard Softmax choice rule (Fig. 2). This analysis revealed an overall drug effect on log(k), such that discounting was substantially lower in the haloperidol group compared with the placebo group (Fig. 1*a*). Examination of BFs indicated that a decrease in log(k) in haloperidol versus placebo was ~116 times more likely than an increase (Table 2).

Model comparison

We next compared three versions of the DDM that varied in the way that they accounted for the influence of value differences on trialwise drift rates, based on the DIC (Spiegelhalter et al., 2002). In each model, we included separate group-level distributions for the two drug conditions (haloperidol vs placebo). Furthermore, for each parameter x, we included a shift parameter s_x modeling the change in parameter x from the high-magnitude condition (SS reward =100€) to the low-magnitude condition (SS

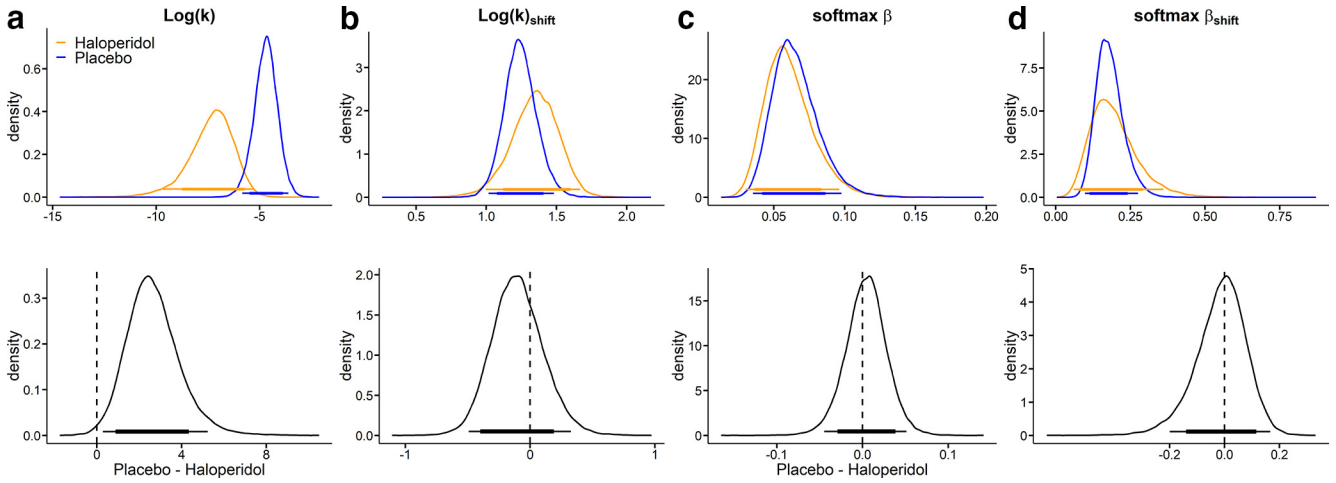


Figure 2. Modeling results (blue: placebo, orange: haloperidol) from a hierarchical Bayesian Model with softmax choice rule. ***a***, Log(k) is the log(discount rate) from the high magnitude condition (smaller-sooner reward = 100€). ***b***, $Log(k)_{shift}$ is the change in log(k) from the high magnitude condition to the low magnitude condition (smaller-sooner reward = 20€). ***c***, is the inverse temperature parameter. ***d***,$_{shift}$ the corresponding shift in inverse temperature from the high to low magnitude condition. The thin (thick) horizontal lines denote 95% (85%) highest posterior density intervals.

Table 2. Summary of group differences in model parameters for the temporal discounting Softmax model[a]

	Baseline		Magnitude effect	
Parameter	M_{diff}	dBF	M_{diff}	dBF
Log(k)	2.66	116.34	−0.10	0.42
Temp	0.03	1.36	−0.01	0.89

[a]For each parameter, we report mean posterior group differences (M_{diff}) and BFs (dBF) testing for directional effects on both the baseline parameter in the 100€ condition (left columns) and on the magnitude effect on each parameter (right columns). BFs < 0.33 reflect evidence for placebo < haloperidol, whereas BFs >3 reflect evidence for placebo > haloperidol. For details, see Statistical analyses.

Table 3. Model comparison of three variants of the DDM based on the DIC (Spiegelhalter et al., 2002) where lower values indicate a better model fit[a]

	DIC		
Model	Placebo	Haloperidol	Full model
DDM_0	11792.1	10034.5	21833.8
DDM_{lin}	10835.0	10092.1	20923.9
DDM_s	8586.5	8161.7	16771.8

[a]The data were generally better accounted for by a temporal discounting DDM with DDM_S compared with DDM_{lin} and DDM_0.

Table 4. Proportion of correctly predicted binary choices for each group and model[a]

	Placebo	Haloperidol
Softmax	0.89 (0.77–1.00)	0.90 (0.78–0.98)
DDM_0	0.73 (0.57–1.00)	0.80 (0.60–0.98)
DDM_{lin}	0.88 (0.71–0.97)	0.85 (0.62–0.98)
DDM_s	0.89 (0.81–1.00)	0.90 (0.82–0.98)

[a]Data are mean (range).

reward = 20€) (see Materials and Methods). These s_x parameters were modeled with Gaussian priors with means of zero (see Materials and Methods). DDM_0 assuming constant drift rates independent of value was also included and compared with two variants of the DDM using either linear (DDM_{lin}) (Pedersen et al., 2017) or in a nonlinear (sigmoid) drift rate scaling (Fontanesi et al., 2019; Peters and D'Esposito, 2020). In both drug conditions as well as overall (Table 3), the data were best accounted for by a DDM with nonlinear drift rate scaling (DDM_S).

We also compared the three diffusion models and the Softmax model with respect to the proportion of binary choices (LL vs SS selections) that they correctly accounted for. As can be seen from Table 4, the DDM_s performed numerically on par with the Softmax model, whereas the DDM_{lin} performed slightly worse.

Overall group differences

We next examined overall group differences in model parameters for the baseline (SS reward =100€) condition. Results are plotted in Figure 3, and BFs for all group comparisons are listed in Table 5. In both groups, there was a positive association between trialwise drift rates and value differences, as the 95% HDI for the drift rate coefficient parameter did not include 0 in either group (Fig. 3*b*). Likewise, there was a slight bias toward the SS option in both groups, as the 95% HDI for bias was <0.5 in both cases (Fig. 3*e*).

We furthermore observed substantially lower group-level discount rates log(k) in the haloperidol group compared with placebo, such that the 95% HDI of the posterior group difference in log(k) was >0 (Fig. 3*a*; Table 5). Interestingly, the nondecision time was likewise substantially lower in the haloperidol group (Fig. 3*c*; Table 5), amounting, on average, to 180 ms faster nondecision times.

Magnitude effects on model parameters

We next turned to the effects of the magnitude manipulation on diffusion model parameters, that is, the change in each parameter in the low-magnitude condition compared with the high-magnitude baseline condition. Results are plotted in Figure 4, and BFs for all directional group comparisons are listed in Table 5. There was a substantial magnitude effect on log(k), such that discounting was steeper in the low-magnitude condition (Fig. 4*a*). Interestingly, this pattern of results was not mirrored by in the magnitude effect on the starting point/bias parameter. Instead, the bias was shifted in the direction of a neutral bias (0.5) in the low-magnitude condition (Fig. 4*e*) in both groups. An additional interesting observation is that the nondecision time was increased in the low-magnitude condition by on average ~30 ms (Fig. 4*c*).

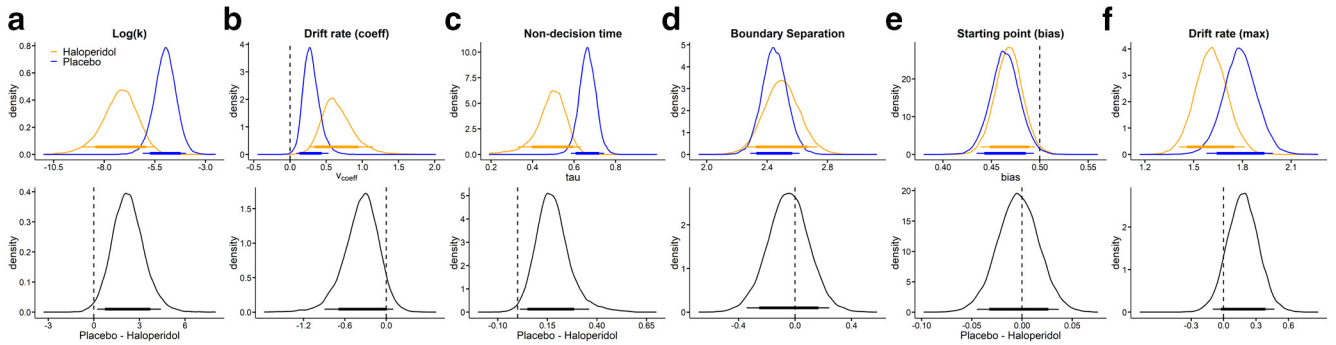


Figure 3. Posterior distributions (blue: placebo, orange: haloperidol) per parameter (top row: ***a***, Log(k); ***b***, Drift rate coefficient; ***c***, Nondecision time; ***d***, Boundary separation; ***e***, Starting point bias; ***f***, Drift rate maximum) and group differences (bottom row, placebo–haloperidol) for the baseline condition (smaller-sooner reward = 100€). Thin (thick) horizontal lines denote 95% (85%) highest posterior density intervals.

Table 5. Summary of group differences in model parameters for the temporal discounting DDM[a]

	Baseline		Magnitude effect	
Model parameter	M_{diff}	dBF	M_{diff}	dBF
Log(k)	2.26	77.9	−0.093	0.47
Drift rate coefficient	−0.365	0.061	0.020	2.73
Nondecision time	0.180	98.4	−0.0001	0.95
Boundary separation	−0.047	0.60	0.017	1.47
Starting point bias	−0.004	0.74	−0.017	0.26
Drift rate maximum	0.18	8.27	0.16	16.88

[a]For each parameter, we report mean posterior group differences (M_{diff}) and BFs (dBF) testing for directional effects on both the baseline parameter in the 100€ condition (left columns) and on the magnitude effect on each parameter (right columns). BFs < 0.33 reflect evidence for placebo < haloperidol, whereas BFs >3 reflect evidence for placebo > haloperidol. For details, see Statistical analyses.

Both drift rate components (v_{coeff} and v_{max}) were increased in the 20€ condition (Fig. 4*b*,*f*). This overall effect might in part be attributable to the fact that, in the model, these two parameters effectively scale the trialwise value differences to the appropriate scale of the DDM (Pedersen et al., 2017). Because average value differences spanned a smaller absolute range in the 20€ condition, this is compensated in the model by increasing both v_{coeff} (Fig. 4*b*) and v_{max} (Fig. 4*f*). Notably, under haloperidol, the drift rate coefficient was somewhat increased, whereas the maximum drift rate was attenuated. There might be some trade-off between the drift rate components, which could contribute to such contrasting effects, such that increases in one component can be compensated by decreases in the other. There was also some evidence for a reduced magnitude effect on the maximum drift rate (Fig. 4*f*) in the haloperidol group. This could be a reflection of the fact that the magnitude effect on LL choice proportions was numerically attenuated under haloperidol (Fig. 1*a*), leading to overall more homogeneous values in the two conditions. Difference distributions in the remaining model parameters were centered at zero, indicating no systematic group differences.

Correlation of model parameters

For descriptive purposes, we show the full correlation matrices for all single-subject median posterior parameters in Figure 5*a* for haloperidol and Figure 5*b* for placebo.

Hierarchical linear regression

We also explored whether the qualitative pattern of results could be reproduced using a hierarchical linear regression, modeling trialwise inverse RTs as a function of value differences (see Materials and Methods). Full posterior distributions of all parameters are shown in Figure 6. This analysis reproduced effects observed for the full DDM. For example, the slope was overall negative, reflecting the decrease in 1/RT for increasing conflict (Fig. 6*a*). The intercept was numerically smaller under haloperidol (dBF = 0.11; see Table 6), mirroring the drug effect on the nondecision time in the DDM_S. However, a direct comparison with DDM parameters is complicated by the fact the intercept in the regression model also captures RT components that in the DDM are reflected in the boundary separation, as well as potentially additional nonlinear aspects of the evidence accumulation process that cannot be accounted for by the slope. These effects are visualized in Figure 6*e* where we plot the 1/RT predicted by this regression model as a function of group, condition, and decision conflict. This illustrates again the slope effect in the baseline condition and the attenuated intercept under haloperidol.

Associations with working memory span

Exploratory analyses did not reveal associations between model parameters of interest (log(k), nondecision time, drift rate scaling) and working memory score (all $|r| < 0.38$).

Posterior predictive checks

We next performed extensive posterior predictive checks to ensure that the best-fitting model (DDM_S) could account for RTs of individual participants in both groups. To this end, we binned the trials of each individual participant into five bins, according to the absolute difference in LL versus SS value (computed according to each participant's median posterior log(k) parameter from the DDM_S). For each bin, participant, and condition, we then plot the mean observed RT, as well as the mean simulated RT across 10,000 datasets simulated from the posterior distributions of the DDM_0, DDM_{lin}, and DDM_S. These results are shown in Figure 7 for the placebo group and Figure 8 for the haloperidol group. As can be seen, the DDM_S provided a much better account of how RTs vary as a function of decision conflict than the DDM_{lin} in the vast majority of participants in both groups. This was mainly because the DDM_{lin} overestimated RTs with medium decision conflict and underestimated RTs in cases of very low decision conflict (Peters and D'Esposito, 2020).

Some additional nontrivial patterns in the data deserve mention. For example, while the DDM_S in most cases predicted longest RTs for choices with the highest decision conflict, this was not always the case (see, e.g., the low-magnitude condition of Participant 34 from the placebo group in Fig. 7). In this case, in the low-magnitude condition, the participant exhibited a

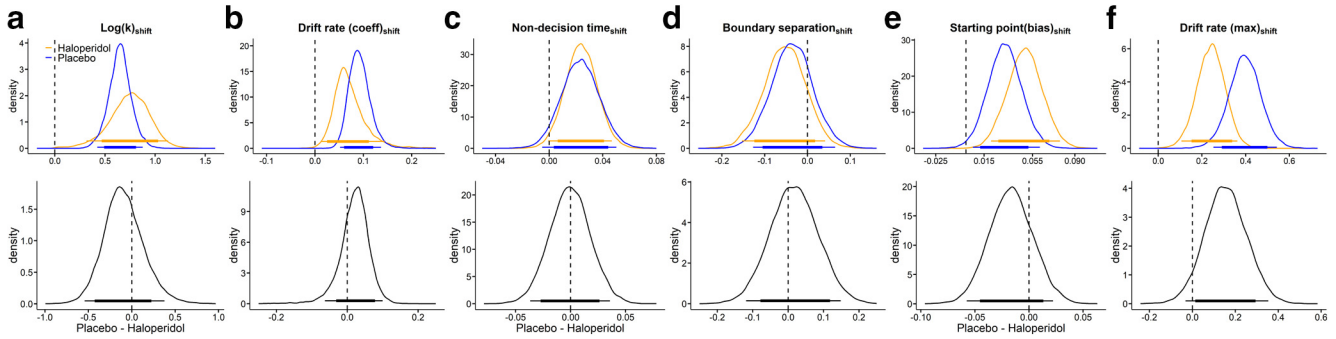


Figure 4. Posterior distributions (blue: placebo, orange: haloperidol) of the change in each parameter from the high magnitude (baseline) to the low magnitude condition (top row: ***a***, Log $(k)_{shift}$; ***b***, Drift rate coefficient$_{shift}$; ***c***, Nondecision time$_{shift}$; ***d***, Boundary separation$_{shift}$; ***e***, Starting point bias$_{shift}$; ***f***, Drift rate maximum$_{shift}$) and corresponding group differences (bottom row, placebo–haloperidol). Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals.

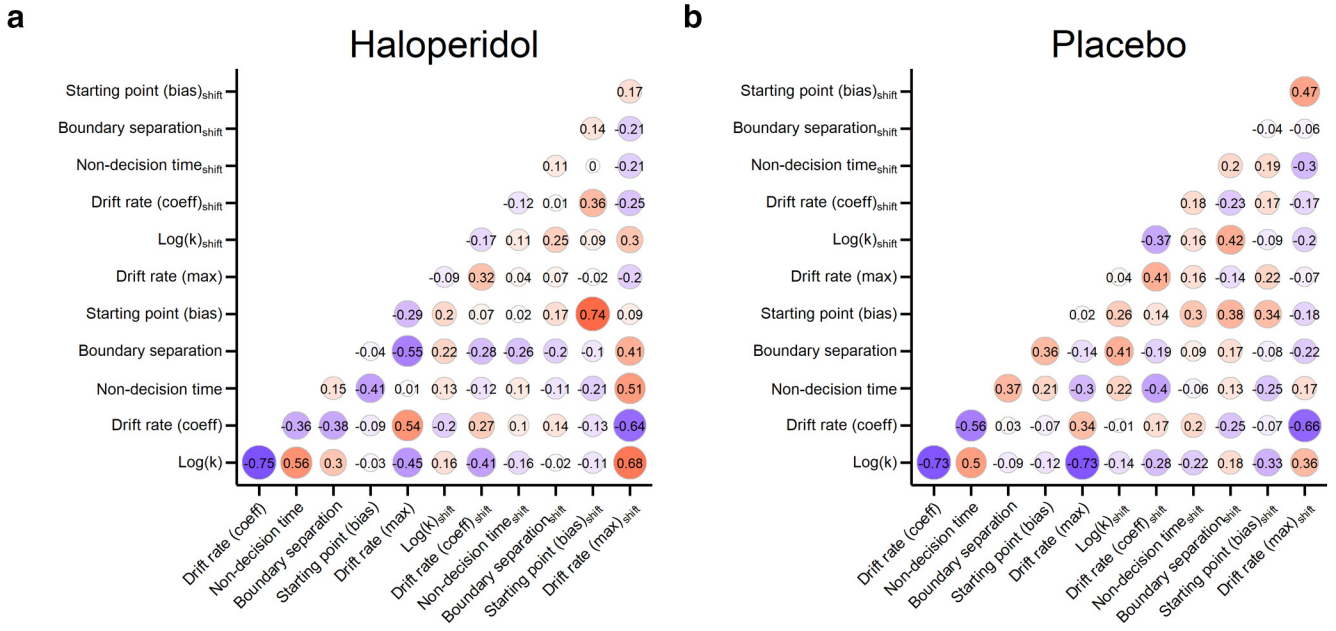


Figure 5. Correlations between all single-subject median posterior parameter estimates across participants from the haloperidol (***a***) and placebo group (***b***).

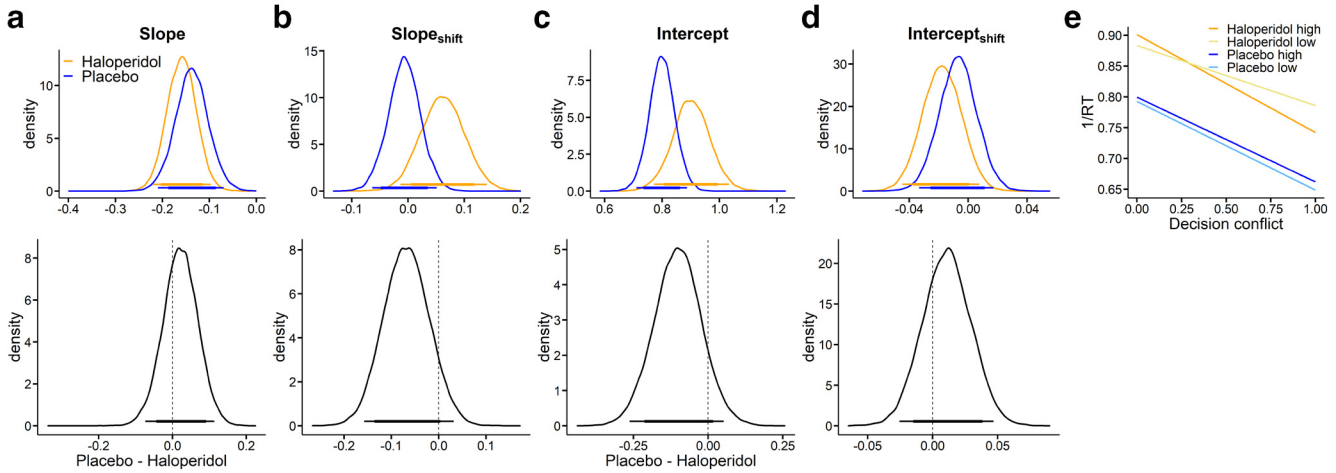


Figure 6. Modeling results (blue: placebo, orange: haloperidol) from a hierarchical linear regression with decision conflict as a predictor and 1/RT as dependent variable. Top row: The slope in ***a***, represents the influence of increasing decision conflict (decreasing value differences) on 1/RT. The intercept in ***c***, here corresponds to 1/RT for the lowest decision conflict (highest subjective value difference) from the high magnitude condition (smaller-sooner reward = 100€). Shift-parameters again reflect the change in slope and intercept (***b***, ***d***) from the high to the low magnitude condition. ***e***, Illustrates 1/RT predicted by this regression model as a function of group, condition and decision conflict. Bottom row: Corresponding group differences (placebo–haloperidol). The thin (thick) horizontal lines denote 95% (85%) highest posterior density intervals.

Table 6. Summary of group differences in model parameters for the hierarchical linear regression model[a]

	Baseline		Magnitude effect	
Model parameter	M_{diff}	dBF	M_{diff}	dBF
Slope	0.02	2.09	−0.07	0.09
Intercept	−0.10	0.11	0.01	2.59

[a]For each parameter, we report mean posterior group differences (M_{diff}) and BFs (dBF) testing for directional effects on both the baseline parameter in the 100€ condition (left columns) and on the magnitude effect on each parameter (right columns). BFs < 0.33 reflect evidence for placebo < haloperidol, whereas BFs >3 reflect evidence for placebo > haloperidol. For details, see Statistical analyses.

relatively small boundary separation (1.84) and drift rate coefficient (0.24), in combination with a bias toward the SS boundary (0.43) and a high discount rate log(k) (−0.7). In such a constellation, the bias toward the SS boundary can only be overcome when value evidence is accumulated for a relatively long time (because v_{coeff} is relatively small), giving rise to long RTs for LL choices (which in this case only occurred in the case of low decision conflict).

Parameter recovery

As a final model check, we ran a series of parameter recovery simulations. Here, we randomly selected 10 datasets simulated from the posterior distribution of the DDM_S (see Materials and Methods), and refit these synthetic data with the DDM_S. Results are shown in Figure 9 for the baseline (high magnitude 100€) parameters, and Figure 10 for the parameters modeling condition effects. As can be seen from these plots, for both baseline and condition effects, this revealed that group-level parameters (Figs. 9, 10, bottom rows) recovered well, such that the true generating parameters were generally contained in the estimated 95% HDIs.

Parameter recovery for individual-subject parameters was excellent for all baseline (100€ magnitude) parameters (Fig. 9, top row) such that the correlation between generating and estimated individual-subject parameters was >0.9 for all parameters. For the parameters modeling condition effects (magnitude effects, Fig. 10, top row), these correlations were lower for some parameters, in particular for condition effects on boundary separation and log(k). The likely reason is that the synthetic data were simulated from the actual posterior distribution, and there was overall little between-subject variance in some of these parameters in our data (see, e.g., Fig. 10*a*,*f*).

Discussion

We investigated the effects of a single dose of the D2-receptor antagonist haloperidol (2 mg) on temporal discounting in a between-subjects study in a double-blind placebo-controlled setting. A diffusion model-based analysis revealed substantially smaller log(k) parameters and a substantial reduction in nondecision times under haloperidol versus placebo.

We applied a recent class of value-based decision models based on the DDM (Pedersen et al., 2017; Fontanesi et al., 2019; Shahar et al., 2019; Peters and D'Esposito, 2020). Comprehensive RT-based analysis was not possible in previous studies because of the specifics of task timing (Pine et al., 2010) or low trial numbers (Weber et al., 2016; Petzold et al., 2019). Model comparison confirmed previous results (Fontanesi et al., 2019; Peters and D'Esposito, 2020), such that the data were better accounted for by a model assuming a nonlinear trialwise scaling of the drift rate, and this was confirmed via posterior predictive checks of single-subject data. Extensive parameter recovery analyses confirmed that group-level parameters recovered well (Fontanesi et al., 2019; Peters and D'Esposito, 2020). Recovery of individual-subject baseline parameters (100€ magnitude condition) was excellent, whereas recovery of parameters modeling condition effects was somewhat lower. This is likely because of some parameters (e.g., boundary separation shift) showing low between-subject variance. Modeling was further validated by the observation that drug effects were fully reproduced using a Softmax choice rule (Sutton and Barto, 1998) and by the finding that the magnitude effect (Green et al., 1997; Ballard et al., 2017; Mellis et al., 2017) was likewise replicated using the DDM-based approach. The qualitative pattern of RT effects was reproduced using a hierarchical linear regression model of trialwise inverse RTs as a function of decision conflict.

The human literature on DA and impulsivity is heterogeneous (D'Amour-Horvat and Leyton, 2014), and interpretation of these findings is complicated by several factors. First, effects of dopaminergic drugs might depend on baseline DA availability (Cools and D'Esposito, 2011), such that the same drug might impair or enhance performance in different participants, according to an inverted U-shaped function (or a different process-dependent function) (Floresco, 2013). Second, the action of D2-receptor antagonists is often interpreted in terms of a reduction in DA neurotransmission (Pessiglione et al., 2006; Pine et al., 2010). But such drugs might indeed enhance DA release by predominantly binding at presynaptic DA auto-receptors, at least at lower dosages (Frank and O'Reilly, 2006) as shown in animal (Pehek, 1999; Schwarz et al., 2004) and human studies (Chen et al., 2005).

Interpretation of D2-receptor antagonist effects as a presynaptically mediated elevation of DA release might reconcile a number of conflicting results. First, our finding of reduced temporal discounting under haloperidol is in line with two recent studies that reported reduced temporal discounting following administration of D2/D3-receptor antagonists (Arrondo et al., 2015; Weber et al., 2016). On the other hand, a reduction of temporal discounting following administration of haloperidol was not observed in an earlier within-subjects study in $n = 13$ participants (Pine et al., 2010) that used a slightly lower dosage of 1.5 mg (we used 2 mg). Lower dosages of D2/D3-receptor antagonists might increase (rather than decrease) DA signaling (Frank and O'Reilly, 2006), an effect mediated by inhibitory feedback through presynaptic D2 auto-receptors (Grace, 1991), which may lead to an enhancement of phasic (vs. tonic) DA signaling (Frank and O'Reilly, 2006), a point that we return to below. However, we do acknowledge that such an interpretation is not general consensus in the cognitive literature on DA drug effects (Pessiglione et al., 2006; Pine et al., 2010).

Our results advance previous findings regarding the role of D2/D3-receptor antagonists in temporal discounting in several ways. First, participants performed an unrelated memory task during fMRI directly before completing the temporal discounting task. Those data revealed an overall main effect of drug condition on trial onset-related activity in caudate nucleus (Clos et al., 2019a,b) (i.e., caudate activity was increased under haloperidol). Although this neural read-out was obtained before the discounting task, both the fMRI and temporal discounting time points were well within the time of maximum haloperidol plasma levels (Franken et al., 2017). This observation is arguably more compatible with the idea that the dosage of haloperidol applied here increased (rather than decreased) striatal DA signaling. Similar neural evidence was lacking in most previous human pharmacological studies on DA effects on discounting (de Wit et

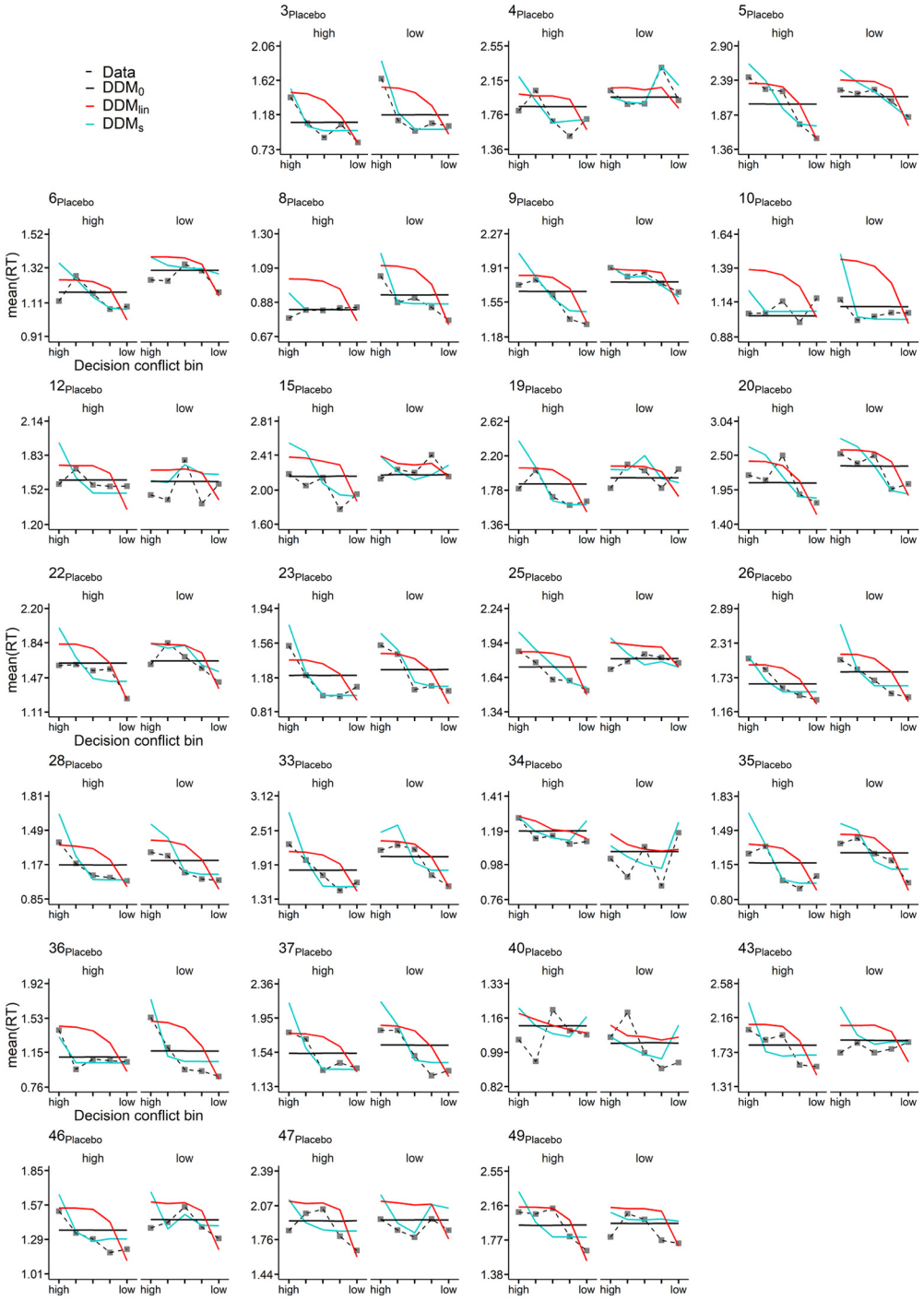


Figure 7. Placebo condition posterior predictive checks. For each participant and condition (high (left facet) represents the high magnitude condition; low (right facet) represents the low magnitude condition), trials were binned into five equal sized bins according to the absolute difference in between subjective LL and SS options (decision conflict bin). Plotted are mean observed RTs per bin (data) as well model-generated RTs (blue represents DDM_0; red represents DDM_{lin}; orange represents DDM_s) averaged >10,000 datasets simulated from the posterior distribution of each hierarchical model (blue represents DDM_0; red represents DDM_{lin}; orange represents DDM_s).

al., 2002; Hamidovic et al., 2008; Arrondo et al., 2015; Weber et al., 2016). Second, the DDM-based modeling approach adopted in the present study allowed us examine the dynamics underlying decision-making much more comprehensively than previous human pharmacological studies (de Wit et al., 2002; Hamidovic et al., 2008; Pine et al., 2010; Arrondo et al., 2015; Weber et al., 2016; Petzold et al., 2019). In addition to the drug effect on the discount rate log(k), diffusion modeling revealed substantially shorter nondecision times in the haloperidol group that amounted to ≈180 ms on average. Such a robust enhancement of lower-level motor and/or perceptual RT components is also more compatible with an increase, rather than a decrease, in DA transmission (Weed and Gold, 1998) and resonates with previous findings regarding a dopaminergic enhancement of RT-based response vigor (Guitart-Masip et al., 2011; Beierholm et al., 2013). An exploratory inspection of parameter correlations revealed that log(k) and nondecision time were positively correlated in both groups, suggesting that they might capture similar

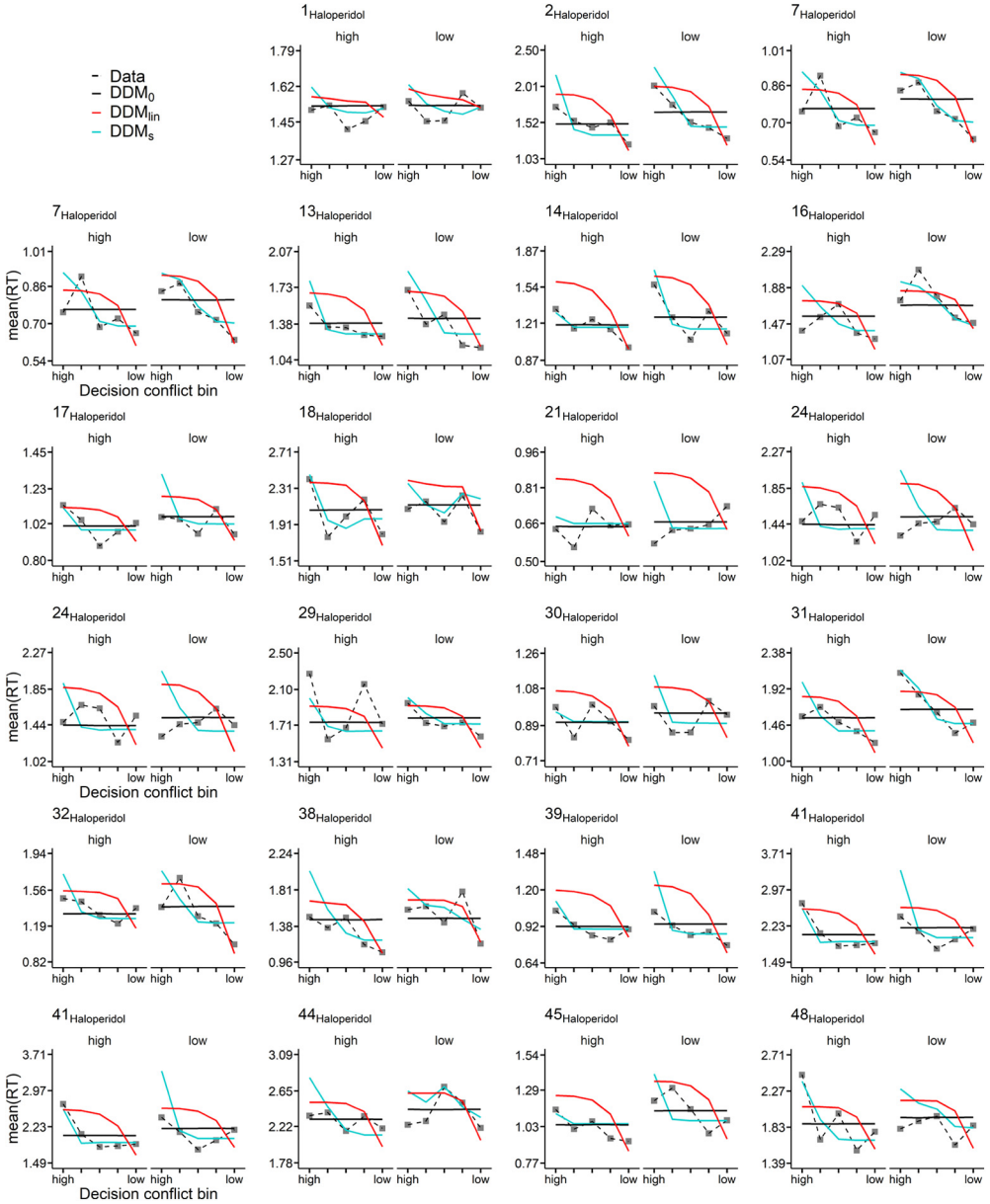


Figure 8. Haloperidol condition posterior predictive checks. For each participant and condition (high (left facet) represents the high magnitude condition; low (right facet) represents the low magnitude condition), trials were binned into five equal sized bins according to the absolute difference in between subjective LL and SS options (decision conflict bin). Plotted are mean observed RTs per bin (data) as well model-generated RTs (blue represents DDM_0; red represents DDM_{lin}; orange represents DDM_s) averaged >10,000 datasets simulated from the posterior distribution of each hierarchical model (blue represents DDM_0; red represents DDM_{lin}; orange represents DDM_s).

aspects of the data and/or might both be modulated by changes in phasic dopaminergic responses. In support of this interpretation, augmentation of DA levels in Parkinson's disease patients reduces temporal discounting (Foerde et al., 2016) and improves model-based RL (Sharp et al., 2016). Finally, this interpretation of available human D2-receptor antagonist effects would also reconcile the human and animal literature on acute dopaminergic effects on impulsivity (D'Amour-Horvat and Leyton, 2014). Together, these considerations lead us to suggest that haloperidol increased (rather than decreased) striatal DA neurotransmission, resulting in enhanced cognitive control (reduced discounting) and a substantial facilitation of motor responding (shorter nondecision times).

By what mechanism might haloperidol attenuate the impact of delay on reward valuation? According to models of basal ganglia contributions to action selection (Maia and Frank, 2011), the probability for selecting a given candidate action depends on the relative difference in activation between the direct (*go*) and the indirect (*nogo*) pathways. A similar striatal gating mechanism might underlie working memory and/or prefrontal control functions (Cools, 2011). By increasing phasic DA responses, haloperidol might increase the signal-to-noise ratio in striatal value representations, thereby increasing the likelihood that objectively smaller and/or more delayed LL rewards gain access to processing in the PFC. Naturally, other modes of action are likewise conceivable. Frontal and striatal regions are interconnected via a series of loops that follow a dorsal-to-ventral organization (Haber and Knutson, 2010), and haloperidol might impact functional interactions within these circuits (Cools, 2011), for example, related to top-down control of value representations (Hare et al., 2009,

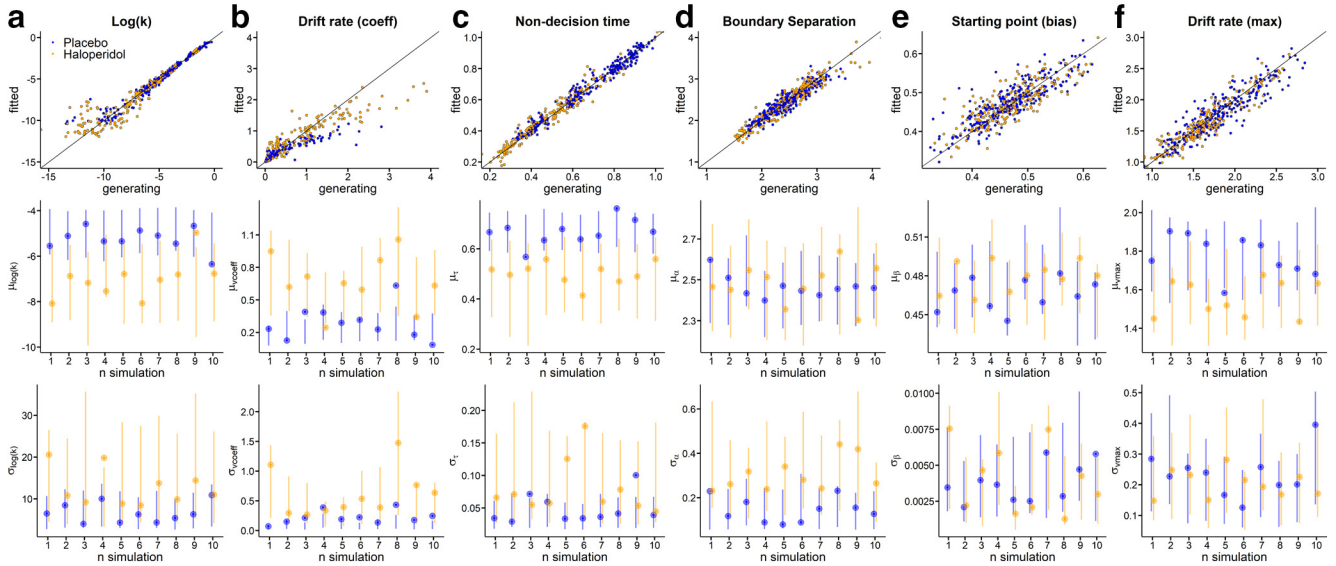


Figure 9. Parameter recovery analysis for all Baseline parameters using the DDM$_s$ (***a***, Log(k); ***b***, Drift rate coefficient; ***c***, Nondecision time; ***d***, Boundary separation; ***e***, Starting point bias; ***f***, Drift rate maximum). Top row: Generating parameters vs. fitted parameters for each subject across ten simulations for haloperidol group (yellow) and placebo group (blue). Second row: True generating group level hyperparameter means (points) and Bottom row: standard deviations (points) and estimated 95% highest density intervals (lines) per fitted simulation. For correlations between generating and estimated single-subject parameters, see Extended Data Figure 9-1.

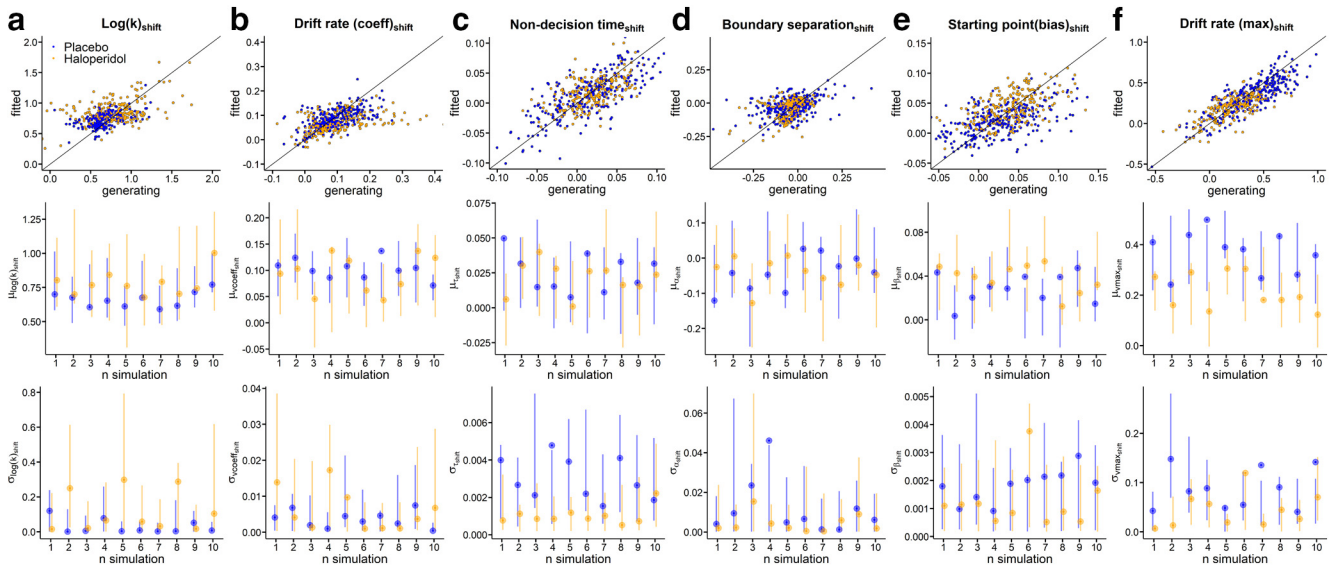


Figure 10. Parameter recovery analysis for all shift parameters using the DDM$_s$ (***a***, Log(k)$_{shift}$; ***b***, Drift rate coefficient$_{shift}$; ***c***, Nondecision time$_{shift}$; ***d***, Boundary separation$_{shift}$; ***e***, Starting point bias$_{shift}$; ***f***, Drift rate maximum$_{shift}$). Top row: Generating parameters vs. fitted parameters for each subject across ten simulations for haloperidol group (yellow) and placebo group (blue). Second row: True generating group level hyperparameter means (points) and Bottom row: standard deviations (points) and estimated 95% highest density intervals (lines) per fitted simulation.

2014; Figner et al., 2010; Peters and D'Esposito, 2016). Finally, haloperidol might have directly augmented control processes in specific PFC regions (Figner et al., 2010). However, because of the much greater expression of D2 receptors in striatum compared with PFC (Seamans and Yang, 2004), it is generally assumed that prefrontal action of D2 antagonists requires substantially higher dosages than those applied in the studies examined here (Seamans and Yang, 2004; Frank and O'Reilly, 2006).

The present study has a number of limitations that need to be acknowledged. First, we did not run a within-subjects design, which would have allowed us to account for individual-participant baseline parameters in the analysis of the drug effects. Second, this also precluded us from comprehensively analyzing potential modulatory influences of, for example, individual differences in working memory on the drug effects, which might modulate DA effects on discounting (Petzold et al., 2019) and cognitive control more generally (Cools and D'Esposito, 2011). Third, the proportion of female participants was relatively large. Given the known association of ovarian hormones with the DA system (Yoest et al., 2018), future studies would benefit from testing larger sample sizes that allow for the examination of gender effects and/or from directly controlling menstrual cycle phase. Fourth, rewards were hypothetical because of the inclusion of the high-magnitude condition. However, preferences for real and hypothetical outcomes in temporal discounting tasks show a very good correspondence (Johnson and Bickel, 2002)

and rely on similar neural circuits (Bickel et al., 2009). Also, neural haloperidol effects vary across brain regions and functions (Wächtler et al., 2020), complicating interpretation as no task-related imaging data were obtained here.

In conclusion, our data show that the D2-receptor antagonist haloperidol attenuated temporal discounting and substantially shortened nondecision times, as revealed by comprehensive computational modeling of choices and RTs using hierarchical Bayesian parameter estimation. These data are best accounted for by a model in which low dosages of haloperidol lead to an enhancement of phasic DA responses because of reduced feedback inhibition from D2 auto-receptors, leading to an augmentation of both lower-level (nondecision time) and higher-level (temporal discounting) decision components.

References

Acheson A, de Wit H (2008) Bupropion improves attention but does not affect impulsive behavior in healthy young adults. Exp Clin Psychopharmacol 16:113–123.

Amlung M, Petker T, Jackson J, Balodis I, MacKillop J (2016) Steep discounting of delayed monetary and food rewards in obesity: a meta-analysis. Psychol Med 46:2423–2434.

Amlung M, Marsden E, Holshausen K, Morris V, Patel H, Vedelago L, Naish KR, Reed DD, McCabe RE (2019) Delay discounting as a transdiagnostic process in psychiatric disorders: a meta-analysis. JAMA Psychiatry 76:1176.

Arrondo G, Aznárez-Sanado M, Fernández-Seara MA, Goñi J, Loayza FR, Salamon-Klobut E, Heukamp FH, Pastor MA (2015) Dopaminergic modulation of the trade-off between probability and time in economic decision-making. Eur Neuropsychopharmacol 25:817–827.

Ballard IC, Kim B, Liatsis A, Aydogan G, Cohen JD, McClure SM (2017) More is meaningful: the magnitude effect in intertemporal choice depends on self-control. Psychol Sci 28:1443–1454.

Beard E, Dienes Z, Muirhead C, West R (2016) Using Bayes factors for testing hypotheses about intervention effectiveness in addiction in addiction research. Addiction 111:2230–2247.

Beierholm U, Guitart-Masip M, Economides M, Chowdhury R, Duzel E, Dolan R, Dayan P (2013) Dopamine modulates reward-related vigor. Neuropsychopharmacology 38:1495–1503.

Bickel WK, Pitcock JA, Yi R, Angtuaco EJ (2009) Congruence of BOLD response across intertemporal choice conditions: fictive and real money gains and losses. J Neurosci 29:8839–8846.

Bickel WK, Koffarnus MN, Moody L, Wilson AG (2014) The behavioral- and neuro-economic process of temporal discounting: a candidate behavioral marker of addiction. Neuropharmacology 76:518–527.

Chen YC, Choi JK, Andersen SL, Rosen BR, Jenkins BG (2005) Mapping dopamine D2/D3 receptor function using pharmacological magnetic resonance imaging. Psychopharmacology (Berl) 180:705–715.

Clos M, Bunzeck N, Sommer T (2019a) Dopamine is a double-edged sword: dopaminergic modulation enhances memory retrieval performance but impairs metacognition. Neuropsychopharmacology 44:555–563.

Clos M, Bunzeck N, Sommer T (2019b) Dopamine enhances item novelty detection via hippocampal and associative recall via left lateral prefrontal cortex mechanisms. J Neurosci 39:7920–7933.

Cools R (2011) Dopaminergic control of the striatum for high-level cognition. Curr Opin Neurobiol 21:402–407.

Cools R, D'Esposito M (2011) Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol Psychiatry 69:e113-e125.

D'Amour-Horvat V, Leyton M (2014) Impulsive actions and choices in laboratory animals and humans: effects of high vs. low dopamine states produced by systemic treatments given to neurologically intact subjects. Front Behav Neurosci 8:432.

de Wit H, Enggasser JL, Richards JB (2002) Acute administration of D-amphetamine decreases impulsivity in healthy volunteers. Neuropsychopharmacology 27:813–825.

Doll BB, Simon DA, Daw ND (2012) The ubiquity of model-based reinforcement learning. Curr Opin Neurobiol 22:1075–1081.

Figner B, Knoch D, Johnson EJ, Krosch AR, Lisanby SH, Fehr E, Weber EU (2010) Lateral prefrontal cortex and self-control in intertemporal choice. Nat Neurosci 13:538–539.

Floresco SB (2013) Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. Front Neurosci 7:62.

Foerde K, Figner B, Doll BB, Woyke IC, Braun EK, Weber EU, Shohamy D (2016) Dopamine modulation of intertemporal decision-making: evidence from Parkinson disease. J Cogn Neurosci 28:657–667.

Fontanesi L, Gluth S, Spektor MS, Rieskamp J (2019) A reinforcement learning diffusion decision model for value-based decisions. Psychon Bull Rev 26:1099–1121.

Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behav Neurosci 120:497–517.

Franken LG, Mathot RA, Masman AD, Baar FP, Tibboel D, van Gelder T, Koch BC, de Winter BC (2017) Population pharmacokinetics of haloperidol in terminally ill adult patients. Eur J Clin Pharmacol 73:1271–1277.

Grace AA (1991) Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41:1–24.

Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic rewards. Psychol Bull 130:769–792.

Green L, Myerson J, McFadden E (1997) Rate of temporal discounting decreases with amount of reward. Mem Cognit 25:715–723.

Guitart-Masip M, Beierholm UR, Dolan R, Duzel E, Dayan P (2011) Vigor in the face of fluctuating rates of reward: an experimental examination. J Cogn Neurosci 23:3933–3938.

Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology 35:4–26.

Hamidovic A, Kang UJ, de Wit H (2008) Effects of low to moderate acute doses of pramipexole on impulsivity and cognition in healthy volunteers. J Clin Psychopharmacol 28:45–51.

Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. Science 324:646–648.

Hare TA, Hakimi S, Rangel A (2014) Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. Front Neurosci 8:50.

Jackson JN, MacKillop J (2016) Attention-deficit/hyperactivity disorder and monetary delay discounting: a meta-analysis of case-control studies. Biol Psychiatry Cogn Neurosci Neuroimaging 1:316–325.

Johnson MW, Bickel WK (2002) Within-subject comparison of real and hypothetical money rewards in delay discounting. J Exp Anal Behav 77:129–146.

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795.

Krebs CA, Reilly WJ, Anderson KG (2016) Reinforcer magnitude affects delay discounting and influences effects of D-amphetamine in rats. Behav Processes 130:39–45.

Kroemer NB, Lee Y, Pooseh S, Eppinger B, Goschke T, Smolka MN (2019) L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action. Neuroimage 186:113–125.

Lempert KM, Steinglass JE, Pinto A, Kable JW, Simpson HB (2019) Can delay discounting deliver on the promise of RDoC? Psychol Med 49:190–199.

Maia TV, Frank MJ (2011) From reinforcement learning models to psychiatric and neurological disorders. Nat Neurosci 14:154–162.

Mazur JE (1987) An adjusting procedure for studying delayed reinforcement. In: Quantitative analyses of behavior (Commons ML, Mazur JE, Nevin JA, Rachlin H, eds), pp 555–573. Hillsdale, NJ: Erlbaum.

Mellis AM, Woodford AE, Stein JS, Bickel WK (2017) A second type of magnitude effect: reinforcer magnitude differentiates delay discounting between substance users and controls. J Exp Anal Behav 107:151–160.

Pedersen ML, Frank MJ, Biele G (2017) The drift diffusion model as the choice rule in reinforcement learning. Psychon Bull Rev 24:1234–1251.

Pehek EA (1999) Comparison of effects of haloperidol administration on amphetamine-stimulated dopamine release in the rat medial prefrontal cortex and dorsal striatum. J Pharmacol Exp Ther 289:14–23.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–1045.

Peters J, Büchel C (2011) The neural mechanisms of inter-temporal decision-making: understanding variability. Trends Cogn Sci 15:227–239.

Peters J, D'Esposito M (2016) Effects of medial orbitofrontal cortex lesions on self-control in intertemporal choice. Curr Biol 26:2625–2628.

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 Comput Biol 16:e1007615.

Petzold J, Kienast A, Lee Y, Pooseh S, London ED, Goschke T, Smolka MN (2019) Baseline impulsivity may moderate L-DOPA effects on value-based decision-making. Sci Rep 9:5652.

Pine A, Shiner T, Seymour B, Dolan RJ (2010) Dopamine, time, and impulsivity in humans. J Neurosci 30:8888–8896.

Plummer M (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing, p 125. Technische Universit at Wien.

Robinson TE, Berridge KC (1993) The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Res Brain Res Rev 18:247–291.

Schwarz A, Gozzi A, Reese T, Bertani S, Crestan V, Hagan J, Heidbreder C, Bifone A (2004) Selective dopamine D(3) receptor antagonist SB-277011-A potentiates phMRI response to acute amphetamine challenge in the rat brain. Synapse 54:1–10.

Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog Neurobiol 74:1–58.

Shahar N, Hauser TU, Moutoussis M, Moran R, Keramati M, Dolan RJ, NSPN Consortium (2019) Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. PLoS Comput Biol 15:e1006803.

Shakhatreh M, Jehangir A, Malik Z, Parkman HP (2019) Metoclopramide for the treatment of diabetic gastroparesis. Expert Rev Gastroenterol Hepatol 13:711–721.

Sharp ME, Foerde K, Daw ND, Shohamy D (2016) Dopamine selectively remediates "model-based" reward learning: a computational approach. Brain 139:355–364.

Shenhav A, Rand DG, Greene JD (2017) The relationship between intertemporal choice and following the path of least resistance across choices, preferences, and beliefs. Judge Decis Making 12:1–18.

Solway A, Lohrenz T, Montague PR (2017) Simulating future value in intertemporal choice. Sci Rep 7:43119.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc B 64:583–639.

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: Massachusetts Institute of Technology.

Wabersich D, Vandekerckhove J (2014) Extending JAGS: a tutorial on adding custom distributions to JAGS (with a diffusion model example). Behav Res Methods 46:15–28.

Wächtler CO, Chakroun K, Clos M, Bayer J, Hennies N, Beaulieu JM, Sommer T (2020) Region-specific effects of acute haloperidol in the human midbrain, striatum and cortex. Eur Neuropsychopharmacol 35:126–135.

Weber SC, Beck-Schimmer B, Kajdi ME, Müller D, Tobler PN, Quednow BB (2016) Dopamine D2/3- and μ-opioid receptor antagonists reduce cue-induced responding and reward impulsivity in humans. Transl Psychiatry 6:e850.

Weed MR, Gold LH (1998) The effects of dopaminergic agents on reaction time in rhesus monkeys. Psychopharmacology (Berl) 137:33–42.

Wiehler A, Peters J (2015) Reward-based decision making in pathological gambling: the roles of risk and delay. Neurosci Res 90:3–14.

Wunderlich K, Smittenaar P, Dolan RJ (2012) Dopamine enhances model-based over model-free choice behavior. Neuron 75:418–424.

Yoest KE, Quigley JA, Becker JB (2018) Rapid effects of ovarian hormones in dorsal striatum and nucleus accumbens. Horm Behav 104:119–129.

Gambling environment exposure increases temporal discounting but improves model-based control in regular slot-machine gamblers.

Ben Wagner, David Mathar & Jan Peters

Department of Psychology, Biological Psychology, University of Cologne, Cologne, Germany

Correspondence

Ben Wagner (ben.jonathan.wagner@uni-koeln.de), Jan Peters (jan.peters@uni-koeln.de)

Funding

This work was funded by Deutsche Forschungsgemeinschaft (grant PE1627/5-1 to J.P.).

Abstract

Gambling disorder is a behavioral addiction that negatively impacts personal finances, work, relationships and mental health. In this pre-registered study (https://osf.io/5ptz9/) we investigated the impact of real-life gambling environments on two computational markers of addiction, temporal discounting and model-based reinforcement learning. Regular gamblers (n = 30, DSM-5 score range 3-9) performed both tasks in a neutral (café) and a gambling-related environment (slot-machine venue) in counterbalanced order. Data were modeled using drift diffusion models for temporal discounting and reinforcement learning via hierarchical Bayesian estimation. Replicating previous findings, gamblers discounted rewards more steeply in the gambling-related context. This effect was positively correlated with gambling related cognitive distortions (pre-registered analysis). In contrast to our pre-registered hypothesis, model-based reinforcement learning was improved in the gambling context. Gambling disorder is associated with increased temporal discounting and reduced model-based learning. Here we show that these effects are modulated in opposite ways by real-life gambling cue exposure. Results challenge aspects of habit theories of addiction, and reveal that laboratory-based computational markers of psychopathology are under substantial contextual control.

Introduction

Gambling disorder is a behavioral addiction that can have detrimental effects on quality of life including personal finances, work, relationships and overall mental health (Blaszczynski & Nower, 2002; Muggleton et al., 2021). Despite these negative consequences, many gamblers are motivated to continue to play, and praise the temporary excitement and pleasure (Fauth-Bühler et al., 2017). Accumulating evidence suggests similarities of gambling disorder and substance-use-disorders both on behavioral, cognitive and neural levels (Balodis & Potenza, 2020; Leeman & Potenza, 2012; Lobo et al., 2015; N. M. Petry, 2010; Singer et al., 2020). In light of these similarities, the fifth edition of the "Diagnostic and Statistical Manual of Mental Disorders" categorizes gambling disorder in the category of "Substance-related and Addictive Disorders" (Association, 2013). In contrast to substance-use-disorders, differences in behavioral and/or neural effects between gamblers and controls are unlikely to be confounded by chronic or acute drug effects (Clark et al., 2019; Peters & Büchel, 2011; Wiehler & Peters, 2015) . Gambling disorder has thus been termed a "pure addiction" (Mark Dixon, Ghezzi, et al., 2006).

Recently, categorical definitions of mental illness have increasingly been called into question. The National Institute for Mental Health of the United States proposed the Research Domain Criteria (RDoC) to foster characterization of the dimensions underlying psychiatric disorders. According to this approach, research in cognitive science should focus on the identification of continuous neuro-cognitive dimensions that might go awry in disease, i.e. trans-diagnostic markers (Nelson et al., 2016). Here we focus on two promising candidates for such trans-diagnostic processes that are affected across a range of psychiatric conditions, including gambling disorder: temporal discounting, i.e. the devaluation of delayed rewards (Bickel et al., 2019; Lempert et al., 2019; Peters & Büchel, 2011), and model-based (MB) control during reinforcement learning (Daw et al., 2011). MB control refers to computationally more expensive goal-directed strategies that utilize models of the environment, contrasting with model-free (MF) control that operates on stimulus-response associations (Balleine & O'Doherty, 2010; Daw et al., 2011; Doll et al., 2012; Valerie Voon et al., 2017).

Steep discounting has been consistently observed in substance use disorders and gambling disorder (Bickel et al., 2012; Bickel et al., 2019; MacKillop et al., 2011; Reynolds, 2006). Moreover, alterations in temporal discounting occur in a range of other disorders, including depression, bipolar disorder, schizophrenia and borderline personality disorder (Amlung et al., 2019), underlining the trans-diagnostic nature of this process. Changes in the contributions of MF and MB control have likewise been reported across disorders, including

gambling disorder (Wyckmans et al., 2019), schizophrenia (Culbreth et al., 2016), obsessive compulsive disorder (Gillan et al., 2020) and substance use disorders (Sebold et al., 2014). Reduced MB control is also reflected in sub-clinical psychiatric symptom severity (Gillan et al., 2016).

Addiction is known to be under substantial contextual control. Addiction-related cues and environments are powerful triggers of subjective craving, drug use and relapse. Incentive sensitization theory (T. Robinson & Berridge, 1993; Terry E. Robinson & Berridge, 2008) provides a theoretical framework that links such effects to a highly sensitized dopamine system that responds to drugs and addiction-related cues. Increased responses of the dopamine system to addiction-related cues ("cue-reactivity") has been consistently observed in neuroimaging studies of human addicts (Courtney et al., 2016; Moeller & Paulus, 2018), and there is evidence that trans-diagnostic behavioral traits are likewise under contextual control. For example, regular gamblers discount delayed rewards substantially more steeply when tested in a gambling-related environment as compared to a neutral environment (Mark. Dixon, Jacobs, & Sanders, 2006). Similar effects have been observed in laboratory tasks that include gambling-related cues (Dale et al., 2019; Genauck et al., 2020; Miedl et al., 2014) but whether other putative trans-diagnostic traits such as MB control are under similar contextual control is unclear. Beyond, it is unclear whether gambling severity or maladaptive control beliefs (Raylu & Oei, 2004) modulate such effects.

Though rarely examined in naturalistic settings, contextual effects on trans-diagnostic dimensions of decision-making are of substantial clinical and scientific interest. Settings with high ecological validity might provide more informative insights into the central drivers of maladaptive behavior than laboratory-based studies (Anderson & Brown, 1984). If such trans-diagnostic traits are further exacerbated in e.g. addiction-related environments, this could constitute a mechanism underlying the maintenance and/or escalation of maladaptive behavior. Second, traits such as temporal discounting can be modulated (Bickel et al., 2011; Bickel et al., 2019; Lempert & Phelps, 2016) and could thus serve as a potential treatment target (Lempert et al. 2019).

The present pre-registered study thus had the following aims. First, we aimed to replicate the findings by Dixon et al. (Mark Dixon, Ghezzi, et al., 2006), who observed increased temporal discounting in gambling-related environments in regular gamblers, compared to neutral environments. Second, we extended their approach by including a modified version of the prominent 2-step sequential decision task (Daw et al., 2011) to test whether model-based control of behavior is likewise under contextual control. Reduced model-based

control has been linked to a range of psychiatric conditions (see above) including gambling disorder (Wyckmans et al., 2019). Third, we directly tested for associations of contextual effects with gambling severity and working memory capacity. Finally, our tasks allowed for comprehensive computational modelling of choices and response time (RT) distributions. Analyses of reinforcement learning and decision-making have recently been shown to substantially benefit from an incorporation of RTs (Fontanesi et al., 2019; Pedersen et al., 2017; Peters & D'Esposito, 2020; Shahar et al., 2019; Wagner et al., 2020) via the application of sequential sampling models such as the drift diffusion model (DDM) (Forstmann et al., 2016). Such analyses yield additional insights into the latent processes underlying decision-making (Wagner et al., 2020) and can improve parameter stability (Shahar et al., 2019). To account for these recent developments, we complemented our pre-registered analyses with additional analyses of temporal discounting and reinforcement learning drift diffusion models (RLDDM).

Methods

Preregistration

This study was preregistered via the open science framework (osf.io/ua42h). We deviated from the pre-registered study design in two ways. First, it was initially planned to use a lab-setting for the neutral (non-gambling) testing environment. However, this was changed following pre-registration to a café, which we felt was more similar to the gambling environment in terms of the presence of social cues and the overall level of distraction. Second, we initially aimed to include gamblers fulfilling at least one DSM-5 criterion for gambling disorder. This was adjusted to a stricter inclusion criterion of at least three DSM-5 criteria. Both of these changes were implemented before testing began.

To account for recent developments in computational modelling we also complemented the pre-registered computational modeling analyses with additional analyses of RT distributions via temporal discounting and reinforcement learning DDMs (Fontanesi et al., 2019; Pedersen et al., 2017; Peters & D'Esposito, 2020; Wagner et al., 2020). As a model-free measure of intertemporal choice we decided to simply use choice proportions of larger-but-later values instead of computing the Area under the empirical discounting curve (AUC).

A-priori sample size was calculated based on results by Dixon et al. (2006). Dixon et al. observed an effect size of d = .5 for the effect of gambling environments on temporal discounting in regular gamblers. Power analysis (Faul et al., 2007) yielded a minimum sample

size of n = 26 with alpha error probability of .05 and power of .80. We then pre-registered a target sample size of n = 30.

Participants

Participants were recruited via advertisements posted online and in local gambling venues. First, they were screened via a telephone interview to verify that they show evidence for problematic gambling behavior, with a primary gambling mode of electronic slot machines. Further inclusion criteria were age in the range of nineteen to fifty, no illegal drug use, and no history of neuropsychiatric disorders, current medication or a history of cardiovascular disease. The ethics committee of the University of Cologne Medical Center approved all study procedures.

Forty-two participants were then invited to a first appointment, were they provided written informed consent and completed a questionnaire assessment and a set of working memory tasks (see section on *background screening* below). Five participants dropped out during or after the first appointment. Four additional participants were excluded after the first appointment because they fulfilled less than three DSM-V criteria for gambling disorder. Two participants dropped out after the first experimental testing session, and one participant was excluded because he fell asleep twice during one testing session. Due to technical problems, we obtained complete datasets for thirty participants for the intertemporal choice task and twenty-nine participants for the two-step task, with twenty-eight participants overlapping.

Overall procedure

Participants were invited to three appointments. At the first appointment (*baseline screening*; see below) participants were invited to our lab and performed a questionnaire assessment and four working memory tasks. Participants were randomly assigned to one of the two locations (café vs. casino) on the first experimental appointment (pseudorandomized location [first session neutral or gambling] and task-version; see section on tasks below). We label the café environment as neutral because no gambling associated cues were present. In both locations, the delay discounting task was completed first, followed by the 2-step task. Appointments were made on an individual basis but spaced within 7+-2 days and around the same time of day +- 2 hours. The café environment was an ordinary café serving non-alcoholic drinks and snacks and furnished with 10 tables and approximately 50 m^2 of size. Testing occurred while the café was in business as usual and experimenter and participant sat at a table next to a wall to assure some privacy. The café was usually moderately attended and testing occurred at the same spot for all participants, with only a few exceptions when this seat was taken. The gambling environment was a common slot-machine venue operated by a German gambling conglomerate. The experimenter and participant were seated at a table placed next to a wall in sight of the electronic gaming machines (EGMs). In total there were four EGMs in direct sight of the participant and a total of ten in the room (hidden by eye protection walls). The density of gambling related cues varied as a function of people playing at EGMs, background sounds e.g. sounds of winning or money dropping were all depended on regularly customers. However, in nearly all cases other people were playing EGMs in direct sight of the participants. The experimenter was granted permission to conduct research in two local gambling venues. Two chairs and a table to use for the experimental session were provided. In both locations, subjects were placed in such a way that neither experimenter nor customers could view their screen. Both tasks ran on a 15inch Laptop using the Psychopysics toolbox running in Matlab (The MathWorks ©).

Background screening

Participants filled out a battery of questionnaires regarding gambling related cognition (GRCS) (Raylu & Oei, 2004) and symptom severity (DSM-5;KFG,SOGS) (Falkai, 2015; Lesieur & Blume, 1987; J. Petry & Baulig, 1996), demographic evaluation and standard psychiatric diagnostic tools (see Supplemental Table S1 and S2).

We assessed working memory capacity using a set of four working memory paradigms. First, in an Operation Span Task (Redick et al., 2012) subjects were required to memorize a sequence of letters while being distracted by math-operations. Second, in a Listening Span Task (adapted from the German version of the Reading Span Test developed by van den Noort et al. (van den Noort et al., 2008) subjects were required to listen to a series of sentences and had to recall the last word of each sentence. Last, subjects performed two different versions of a Digit Span Task (forward/backward) that were adopted from the Wechsler Adult Intelligence Scale (Wechsler, 2008). Here, participants listened to a series of numerical digits which they had to recall as a series in regular or reverse order. All working memory scores were *z*-transformed and averaged to obtain a single compound working memory score (*z*-score).

Temporal discounting task

Participants performed 140 trials of a temporal discounting task where on each trial they made a choice between a smaller-but-sooner (SS) immediate reward, and a larger-but-later (LL) reward delivered after a specific delay. SS and LL rewards were randomly displayed on the left and right sides of the screen, and participants were free to make their choice at any time. While SS rewards were held constant at 20€f LL rewards were computed as multiples thereof. In one version these amounted to 1.05, 1.055, 1.15, 1.25, 1.35, 1.45, 1.55, 1.65, 1.85, 2.05, 2.25, 2.55, 2.85, 3.05, 3.45, 3.85, and in the other version they were 1.025, 1.08, 1.2, 1.20, 1.33, 1.47, 1.5, 1.70, 1.83, 2.07, 2.3, 2.5, 2.80, 3.10, 3.5, 3.80. Each LL reward from one version was then combined with each delay option for this version (in days): (either: 1, 7, 13, 31, 58, 122, or v: 2, 6 15, 29, 62, 118) yielding 140 trials in total. The mean larger later reward magnitude was the same across versions and the order was counterbalanced across subjects and session (neutral/gambling).

At the end of each session one decision was randomly selected and paid out in the form of a gift certificate for a large online store, either immediately (in the case of a smaller-sooner choice) or via email/text message after the respective delay (in the case of a larger-later choice).

2-step task

Participants performed a slightly modified version of the 2-step task, a sequential reinforcement-learning task developed by Daw et al. (Daw et al., 2011). Based on suggestions by Kool et al. (Kool et al., 2016) we modified the outcome stage by replacing the fluctuating reward probabilities (reward/ no reward) with fluctuating reward magnitudes (Gaussian random walks with reflecting boundaries at 0 and 100, and standard deviation of 2.5). In total the task contained 300 trials. Each trial consisted of two successive stages: In the 1st stage (S1), participants chose between two fractals embedded in grey boxes. After taking an S1 action, participants transitioned to one of two possible 2nd stages (S2) with fixed transition probabilities of 70% and 30%. In S2, participants chose between two new fractals each providing a reward outcome in points (between 0-100) that fluctuated over time. To achieve optimal performance, participants had to learn two aspects of the task. They had to learn the transition structure, that is, which S1 stimulus preferentially (70%) leads to which pair of S2 stimuli. Further, they had to infer the fluctuating reward magnitudes associated with each S2 stimulus. .

In both versions, the tasks differed in the S1 and S2 stimuli, and in the fluctuating rewards in S2. However both task versions reward walks were equal in variance and mean and were presented in counterbalanced order per session (neutral/gambling). Participants were instructed about the task structure and performed 40 practice trials (with different random walks and symbols) at the first appointment (*Baseline screening*). Following task completion, points (*0.25) were converted to € and participants could win a bonus of up to 4.50€ that was added to the baseline reimbursement of 10€/h.

Computational modeling and Statistical Analysis

Temporal discounting model

We applied a single-parameter hyperbolic discounting model to describe how subjective value changes as a function of LL reward height and delay (Mazur, 1987; Green and Myerson, 2004):

$$SV(LL_t) = \frac{A_t}{1 + \exp(k + s_k * I_t) * D_t} \quad (1)$$

Here, A_t is the reward height of the LL option on trial t, D_t is the LL delay in days on trial t and I_t is an indicator variable that takes on a value of 1 for trials from the gambling context and 0 for trials from the neutral condition. The model has two free parameters: k is the hyperbolic

discounting rate (modeled in log-space) and s_k is a weighting parameter that models the degree of change in discounting in the gambling compared with the neutral context condition.

Softmax action selection

Softmax action selection models choice probabilities as a sigmoid function of value differences (Sutton and Barto, 1998):

$$P(LL)_t = \frac{\exp\left(\left(\beta + s_\beta * I_t\right) * SV(LL_t)\right)}{\exp\left(\left(\beta + s_\beta * I_t\right) * SV(SS_t)\right) + \exp\left(\left(\beta + s_\beta * I_t\right) * SV(LL_t)\right)} \quad (2)$$

Here, *SV* is the subjective value of the larger but later reward according to Eq. 1 and β is an inverse temperature parameter, modeling choice stochasticity (for $\beta = 0$, choices are random and as β increases, choices become more dependent on the option values). $SV(SS_t)$ was fixed at at 20 and I_t is again the dummy-coded context regressor, and s_β models the context effect on β.

Temporal discounting drift diffusion models

To more comprehensively examine environmental effects on choice dynamics, we additionally replaced softmax action selection with a series of drift diffusion model (DDM)-based choice rules. In the DDM, choices arise from a noisy evidence accumulation process that terminates as soon as the accumulated evidence exceeds one of two response boundaries. In the present setting, the upper boundary was defined as selection of the LL option, whereas the lower boundary was defined as selection of the SS option.

RTs for choices of the SS option were multiplied by -1 prior to model fitting. We furthermore used a percentile-based cut-off, such that for each participant the fastest and slowest 2.5 percent of trials were excluded from the analysis. We then first examined a null model (DDM_0) without any value modulation. Here, the RT on each trial t ($t \in 1{:}140$) is distributed according to the Wiener First Passage Time (*wfpt*):

$$RT_t \sim wfpt(\alpha + s_\alpha * I_t, \tau + s_\tau * I_t, z + s_z * I_t, v + s_v * I_t) \quad (3)$$

The parameter α models the boundary separation (i.e. the amount of evidence required before committing to a decision), τ models the non-decision time (i.e., components of the RT related to motor preparation and stimulus processing), z models the starting point of the evidence accumulation process (i.e., a bias towards one of the response boundaries, with z>.5 reflecting

a bias towards the LL boundary, and z<.5 reflecting a bias towards the SS boundary) and ν models the rate of evidence accumulation. Note that for each parameter x, we also include a parameter s_x that models the change in that parameter from the neutral context to the gambling context (coded via the dummy-coded condition regressor I_t).

As in previous work (Pedersen et al., 2017; Fontanesi et al., 2019; Peters and D'Esposito, 2020, Wagner et al. 2020), we then set up temporal discounting diffusion models with modulation of drift rates by the difference in subjective values between choice options. First, we set up a version with linear modulation of drift-rates (DDM_{lin}) (Pedersen et al., 2017):

$$v_t = \left(v_{coeff} + s_{v_{coeff}} * I_t\right) * \left(SV(LL_t) - SV(SS_t)\right) \quad (4)$$

Here, the drift rate on trial t is calculated as the scaled value difference between the subjective LL and SS rewards. As noted above, RTs for SS options were multiplied by -1 prior to model estimation, such that this formulation predicts SS choices whenever SV(SS)>SV(LL) (the trial-wise drift rate is negative), and predicts longest RTs for trials with the highest decision-conflict (i.e., in the case of SV(SS)= SV(LL) the trial-wise drift rate is zero). We next examined a DDM with non-linear trial-wise drift rate scaling (DDM_S) that has recently been reported to account for the value-dependency of RTs better than the DDM_{lin} (Fontanesi et al., 2019; Peters & D'Esposito, 2020; Wagner et al., 2020). In this model, the scaled value difference from Eq. 4 is additionally passed through a sigmoid function with asymptote v_{max}:

$$v_t = S\left[\left(v_{coeff} + s_{v_{coeff}} * I_t\right) * \left(SV(LL_t) - SV(SS_t)\right)\right] \quad (5)$$

$$S(m) = \frac{2 * \left(v_{max} + s_{v_{max}} * I_t\right)}{1 + \exp\left(-m\right)} - \left(v_{max} + s_{v_{max}} * I_t\right) \quad (6)$$

All parameters including v_{coeff} and v_{max} were again allowed to vary according to the context, such that we included s_x parameters for each parameter x that were multiplied with the dummy-coded condition predictor I_t.

Reinforcement Learning model

Hybrid model

We first applied a slightly modified version of the hybrid RL model (Daw et al., 2011) to analyze the strength of model-free and model-based RL strategies. The model updates MF state-action values (Q_{MF}-values, Eq. 7, 8) in both stages through prediction errors (Eq. 9, 10). In stage 1, MB state-action values (Q_{MB}) are then computed from the transition and reward estimates using the Bellman Equation (Eq. 11).

$$Q_{MF,S1}(a_{j,t}) = Q_{MF,S1}(a_{j,t}) + (\eta_1 + s_{\eta 1} * I_t)\delta_{s1,t} + (\eta_2 + s_{\eta 2} * I_t)\delta_{S2,t} \quad (7)$$

$$Q_{MF,S2}(s_{2i,t}, a_{j,t}) = Q_{MF,S2}(s_{2i,t}, a_{j,t}) + (\eta_2 + s_{\eta 2} * I_t)\delta_{S2,t} \quad (8)$$

$$\delta_{S1,t} = Q_{MF,S2}(s_{2i,t}, a_{j,t}) - Q_{MF,S1}(a_{j,t-1}) \quad (9)$$

$$\delta_{S2,t} = r_{2\,t} - Q_{MF,S2}(s_{2i,t-1}, a_{j,t-1}) \quad (10)$$

$$Q_{MB}(a_{j,t}) = P(s_{21}|s_1, a_j) \max_{a\in\{a_1,a_2\}} Q_{MF,S2}(s_{21}, a) + P(s_{22}|s_1, a_j) \max_{a\in\{a_1,a_2\}} Q_{MF,S2}(s_{22}, a) \quad (11)$$

Here, i indexes the two different second stages (S_{21}, S_{22}), j indexes actions a (a_1, a_2) and t indexes the trials. Further, η_1 and η_2 denote the learning rate for S1 and S2, respectively. S2 MF Q-values are updated by means of reward ($r_{2,t}$) prediction errors ($\delta_{S2,t}$) (Eq. 8, 10). To model S1 MF Q-values we allow for reward prediction errors at the 2nd-stage to influence 1st-stage Q-values (Eq. 7, 9).

In addition, as proposed by Toyama et al. (Toyama et al., 2017, 2019) Q-values of all unchosen stimuli were assumed to decay with decay-rate η_{decay} and centered to the mean of reward walks (0.5). A decay of Q-values over time accounts for the fact that participants know that reward walks fluctuate over time. The decay was implemented according to Eq. 12 and 13:

$$Q_{unchosen}(s_{k,t}, a_{j,t}) = Q_{unchosen}(s_{k,t-1}, a_{j,t-1}) * (\eta decay_S) + (1 - (\eta decay_S) * 0.5 \quad (12)$$

where

$$\eta decay_S = \eta_{decay} + s_{\eta decay} * I_t \quad (13)$$

and $k \in \{1, 21, 22\}$, that is, k indexes the three task stages.

S1 action selection is then modelled via weighting S1 MF and MB Q-values through a softmax action-selection. S2 stage action selection is likewise modelled as a function of MF Q-value

differences. Separate 'inverse temperature' parameters β model subjects' weights of MF and MB *Q*-Values (Eq. 14 and Eq. 15). The additional parameter ρ captures 1st-stage choice perseveration, and is set to 1 if the previous S1 choice was the same and is zero otherwise.

$$p(a_{j,t} = a|s_{1,t}) = \frac{\exp\left(\beta_{MB_s} * Q_{MB}(a) + \beta_{MF_s} * Q_{MF,S1}(a) + \rho_s * rep(a)\right)}{\sum_{a'} exp(\beta_{MB_s} * Q_{MB}(a') + \beta_{MF_s} * Q_{MF,S1}(a') + \rho_s * rep(a'))}, \quad (14)$$

$$p(a_{j,t} = a|s_{2,t}) = \frac{\exp\left(\beta_{2_s} * Q_{MF,S2}(a)\right)}{\sum_{a'} \exp\left(\beta_{2_s} * Q_{MF,S2}(a')\right)}, \quad (15)$$

where:

$\beta_{MB_s} = \beta_{MB} + s_{\beta_{MB}} * I_t$

$\beta_{MF_s} = \beta_{MF} + s_{\beta_{MF}} * I_t$

$\rho_s = \rho + s_\rho * I_t$

$\beta_{2_s} = \beta_2 + s_{\beta_2} * I_t$

Hybrid model with drift diffusion action selection

As in our analysis of temporal discounting we replaced softmax action selection with a DDM choice rule (Shahar et al., 2019), leaving the reinforcement learning equations unchanged. For each stage of the task, the upper boundary was defined as selection of one stimulus, whereas the lower boundary was defined as selection of the other stimulus. We modelled each stage of the task using separate non-decision time (τ), boundary separation (α) and drift- rate (v) parameters. The bias (z) was fixed to 0.5. All parameters including $vcoeff_{MF}$, $vcoeff_{MB}$ and v_{max} were again allowed to vary according to the context, such that we included s_x parameters for each parameter x that were multiplied with the dummy-coded condition predictor I_t (see above). Data were filtered using a percentile-based cut-off, such that for each participant the fastest and slowest 2.5 percent of RTs/trials were excluded from further analysis. In addition, an absolute cutoff of > 150ms was applied. We then first examined a null model (DDM_0; Eq. 3) without any value modulation followed by two value-informed models where the drift-rate (v) is a linear (Eq. 16 and 17) or sigmoid (Eq. 18) function of MF and MB *Q*-value weights. For the linear version, the drift rate in S1 is

$$v_{S1,t} = vcoeff_{MB_s} * (Q_{MB[2]} - Q_{MB[1]}) + vcoeff_{MF_s} * (Q_{MF[2]} - Q_{MF[1]}) + p_s * rep(a') \quad (16)$$

and the drift rate in S2 is calculated as

$$v_{S2,t} = vcoeff_{S2} * (Q_{MF_{S2}[2]} - Q_{MF_{S2}[1]}) \quad (17)$$

For the non-linear version, the linear drift rate from equations 16 and 17 are additionally passed through a sigmoid:

$$v_{Si,t} = \frac{2 * v_{max_{Si_S}}}{1 + exp\,(-m)} - v_{max_{Si_S}} \quad (18)$$

where

$$vcoeff_{MB_s} = vcoeff_{MB} + s_{v_{MB}} * I_t$$

$$vcoeff_{MF_s} = vcoeff_{MF} + s_{v_{MF}} * I_t$$

$$vcoeff_{S2_s} = vcoeff_{S2} + s_{S2} * I_t$$

$$vmax_{Si_s} = vmax_{Si} + s_{Si} * I_t$$

Hierarchical Bayesian models

Softmax models were fit to all trials from all participants using a hierarchical Bayesian modeling approach with separate group-level distributions for all baseline parameters for the neutral context and shift parameters (s_x) for the gambling context.

For the intertemporal choice model fitting was performed using Markov Chain Monte Carlo (MCMC) sampling as implemented in the JAGS (Version 4.3) software package (Martyn Plummer, 2003) in combination with the Wiener module (Wabersich and Vandekerckhove, 2014). Model estimation was done in R (Version 4.0.3) using the corresponding R2Jags package (Version 0.6-1). For baseline group-level means, we used uniform priors defined over numerically plausible parameter ranges (see code and data availability section for details). For all s_x parameters modeling context effects on model parameters, we used Gaussian priors with means of 0. For group-level precisions, we used gamma distributed priors (.001, .001). We initially ran 2 chains with a varying burn-in period and thinning of two until convergence. Chain convergence was then assessed via the Gelman-Rubinstein convergence diagnostic $\hat{R}$ and sampling was continued until $1 \leq \hat{R} \leq 1.02$ for all group-level and individual-subject parameters. 20k additional samples were then retained for further analysis.

For reinforcement learning model fitting was performed using MCMC sampling as implemented in and STAN (Stan Development Team, 2020) via R (Version 4.0.3) and the rSTAN package (Version 2.21.0).

For baseline group-level means, we used uniform and normal priors defined over numerically plausible parameter ranges (see code and data availability section for details). For all s_x parameters modeling context effects on model parameters, we used Gaussian priors with means of 0. For group-level standard deviations we used cauchy (0, 2.5) distributed priors. We initially ran 2 chains with a burn-in period of 1000 and retained 2000 samples for further analysis. Chain convergence was then assessed via the Gelman-Rubinstein convergence diagnostic $\hat{R}$ and sampling was continued until $1 \leq \hat{R} \leq 1.02$. This threshold was not met for one participant ($\hat{R}$ < 1.4).

For both tasks, relative model comparison was performed via the *loo*-package in R (Version 2.4.1) using the Widely-Applicable Information Criterion (WAIC) where lower values reflect a superior fit of the model (Vehtari et al., 2017). We then show posterior group distributions for all parameters of interest as well as their 85% and 95% highest density intervals. For group comparisons we report Bayes Factors for directional effects for s_x hyperparameter distributions of $s_x > 0$ (gambling context > neutral context), estimated via kernel density estimation using R via the RStudio (Version 1.3) interface. These are computed as the ratio of the integral of the posterior difference distribution from 0 to +∞ vs. the integral from 0 to -∞. Using common criteria (Beard et al. 2016), we considered Bayes Factors between 1 and 3 as anecdotal evidence, Bayes Factors above 3 as moderate evidence and Bayes Factors above 10 as strong evidence. Bayes Factors above 30 and 100 were considered as very strong and extreme evidence respectively, whereas the inverse of these reflect evidence in favor of the opposite hypothesis.

Posterior Predictive checks

We carried out posterior predictive checks to examine whether models could reproduce key patterns in the data, in particular the value-dependency of RTs (Peters & D'Esposito, 2020; Wagner et al., 2020) and participant's choices. For the intertemporal choice task, we binned trials of each individual participant into five bins, according to the absolute difference in subjective larger-later vs. smaller-sooner value ("decision conflict", computed according to each participant's median posterior log(k) parameter from the DDM_S, and separately for the neutral and gambling context. For each participant and context, we then plotted the mean observed RTs as a function of decision conflict, as well as the mean RTs across 10k data sets simulated from the posterior distributions of the DDM_0, DDM_{lin} and DDM_S. For the two-step task, we extracted mean posterior parameter estimates and simulated 1000 datasets in R (Version 4.0.3) using the Rwiener package (Version 1.3.3). We then show RTs as a function of S2 reward difference of observed data and the mean RTs across 1000 simulated datasets for of all DDMs. We further show that our models capture the relationship of S2 reward difference and optimal (max[reward]) choices.

Model free analysis

As a model-agnostic measure of temporal discounting, we examined arcsine-square-root transformed proportions of LL choices as a function context (neutral vs. gambling) with order (neutral vs. gambling session first) as fixed and subject as random effect using a hierarchical generalized linear model (HGLM). For the 2-step task we likewise use a HGLM approach and modeled 2nd-stage RTs as a function of transition (common vs. rare) and context (neutral vs. gambling) as fixed and subject as random effect. In line with our modelling analyses, data were filtered so that unplausible fast RTs were excluded (see methods). A standard analysis of stay probabilities (Daw et al., 2011) adapted to our task version is reported in the supplement.

Subjective Craving Rating

On each testing day, participants rated their subjective craving ("How much do you desire to gamble right now?") on a visual-analogue scale ranging from 0 to 100, both at the beginning of the testing session, and at the end following task completion. We then used paired t-tests to examine whether subjective craving differed between the testing environments (neutral vs. gambling).

Results

Subjective craving

Craving was assessed on a visual-analogue-scale before and after task performance. Subjective craving was substantially higher in the gambling-related environment compared to the neutral environment (paired t-test pre-task: t_{23} = -3.13; p = 0.0048, Cohen's *d*: 0.75; post-task: t_{21} = -4.32, p = 0.0003, Cohen's *d* = 0.68; Figure 1).

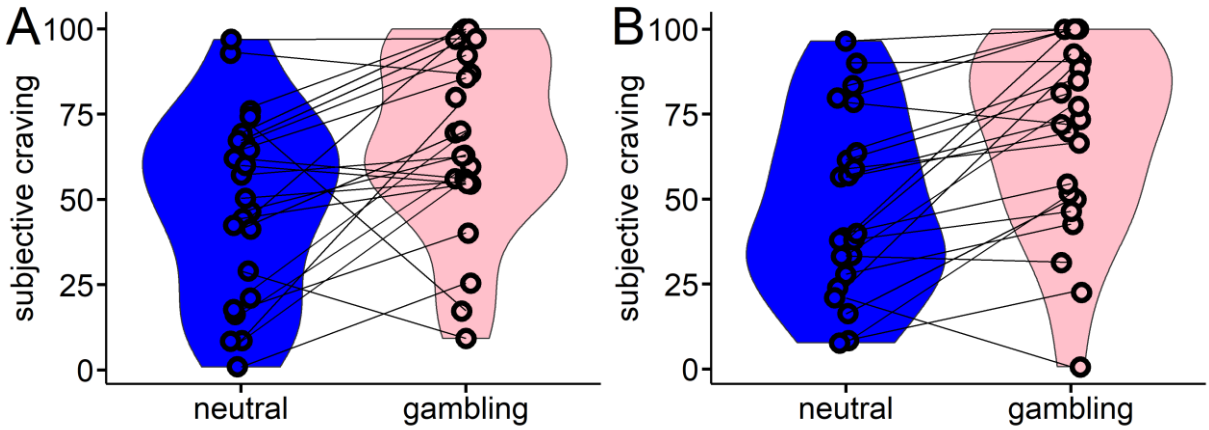


Figure 1. Subjective craving was assessed at the beginning (A) and at the end (B) of each testing session via a visual-analogue scale rating. Craving was significantly higher in the gambling environment, both at the start of the session (p = 0.0048) and at the end of the session (p = 0.0003). Due to technical problems, craving ratings of the first eight participants were lost. Another two participants missed the ratings after task performance.

Model-agnostic analysis temporal discounting task

Raw proportions of larger-but-later (LL) choices are plotted in Figure 2A for each context. A hierarchical linear model on arcsine-square-root transformed proportion values with context and order (gambling vs. neutral first) as fixed effects and subject as random effect confirmed a significant main effect of context (F_{28} = 13.33, p = 0.01) such that participants made more LL selections in the neutral vs. the gambling-related environment. There was no effect of order on choice proportions. Overall response time (RT) distributions are plotted in Figure 2B with choices of the LL option coded as positive RTs and choices of the smaller-sooner option coded as negative RTs.

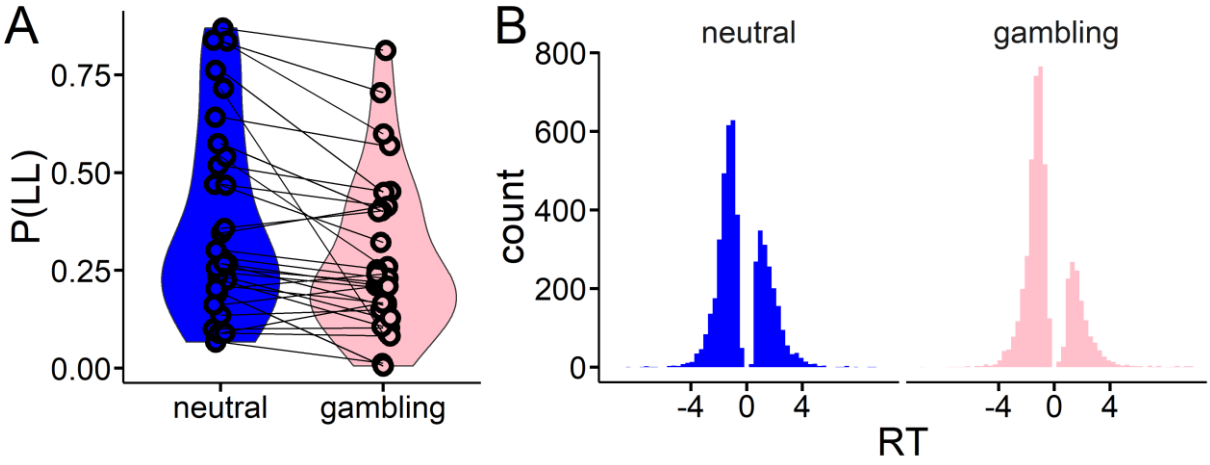


Figure 2. Behavioral data from the temporal discounting task. A: raw proportions of larger-later (LL) choices in each context. B: Overall response time distributions with choices of the LL option coded as positive RTs and choices of the smaller-sooner option coded as negative RTs.

Softmax choice rule

We first modeled our data using standard softmax-action selection. This analysis revealed an overall context effect on log(k), such that discounting was substantially steeper in the gambling context compared to the neutral context (Figure 3B, 95% HDI > 0). Examination of Bayes Factors indicated that an increase in log(k) in the gambling context (s_k) was about 116 times more likely than a decrease (see Figure 3 and Table 3).

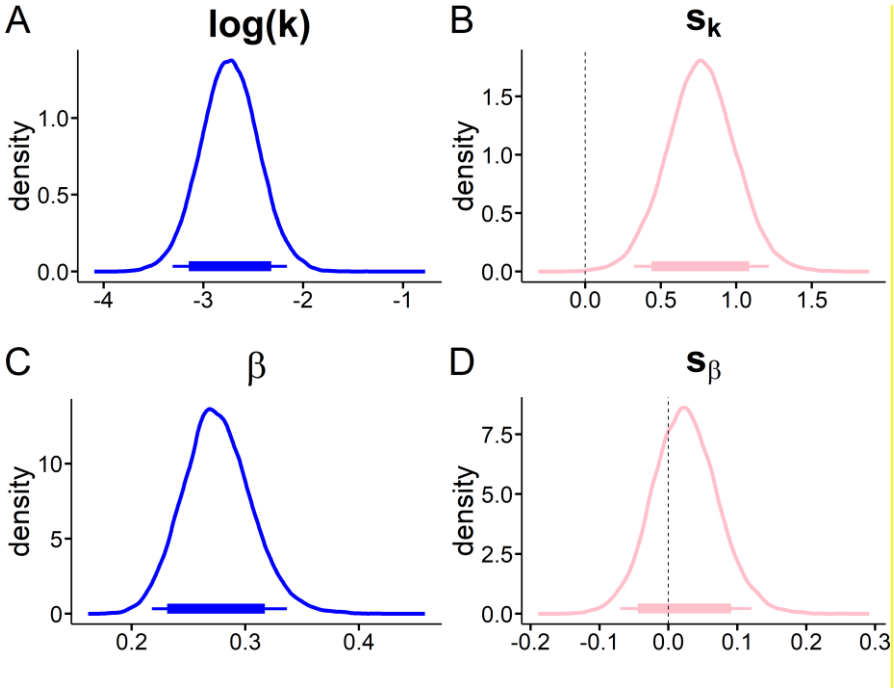


Figure 3. Softmax model; Posterior distributions of mean hyperparameter distributions for the neutral baseline context (blue) and the corresponding shift in the gambling context (pink). A, discount-rate log (k); B, shift in discount-rate (s_k); C, softmax β; D, shift in softmax β; Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals

Temporal discounting drift diffusion models (DDMs)

We next compared three versions of the drift diffusion model (DDM) that varied in the way that they accounted for the influence of value differences on trial-wise drift rates, based on model-fit (WAIC). To verify comparable model ranking across conditions, we first carried out a model comparison separately for each environment (see Table 1). In both environments, a DDM with nonlinear drift-rate scaling (DDM_s) (Fontanesi et al., 2019; Peters & D'Esposito, 2020; Wagner et al., 2020) accounted for the data best when compared to a DDM with linear scaling (DDM_{lin}) (Pedersen et al., 2017) and a null model without value modulation (DDM_0).

We then build a full model with group level distributions for the baseline condition (neutral context) and s_x parameters for each model parameter x, modeling the change from the neutral to the gambling context. S_x parameters where modeled with Gaussian priors with means of zero (see methods section). Model ranking was confirmed for the full model (Table 1).

We next compared the DDMs and the softmax model with respect to the proportion of binary choices (LL vs. SS selections) that they correctly accounted for. As can be seen from Table 2, the DDM_S and DDM_{lin} performed numerically on par with the softmax model, whereas the DDM_0 performed substantially worse (see Supplemental Figure S1). Posterior predictive checks for the winning model showed that it accurately captured the effect of decision conflict

(value difference) on RTs (see *Posterior Predictive Checks* in the supplement; Supplemental Figure S2). Parameter recovery for this model was reported in our prior papers (Peters & D'Esposito, 2020; Wagner et al., 2020).

Table 1. Temporal discounting DDM model comparison using the Watanabe-Akaike Information Criterion (WAIC) revealed the same model ranking for each context (neutral vs. gambling) and the full model. Scores are WAIC (SE).

	Neutral	**Gambling**	**Full model**
DDM_0	12037.7 (150.1)	11754.9 (157.7)	23792.4 (217.4)
DDM_{lin}	9304.6 (155.5)	9174.7 (158.5)	18949.6 (219.0)
DDM_s	8982.3 (155.6)	8744.6 (157.6)	17656.0 (220.8)

Table 2 Proportions of correctly predicted binary choices (mean [range]) for the temporal discounting models (neutral vs. gambling context; see Supplemental Figure S1).

	Neutral	**Gambling**
Softmax	0.89 [0.57-0.98]	0.90 [0.66-1.00]
DDM_0	0.74 [0.52-0.93]	0.76 [0.55-0.99]
DDM_{lin}	0.89 [0.58-0.98]	0.90 [0.65-0.99]
DDM_s	0.90 [0.60-0.99]	0.91 [0.70-1.00]

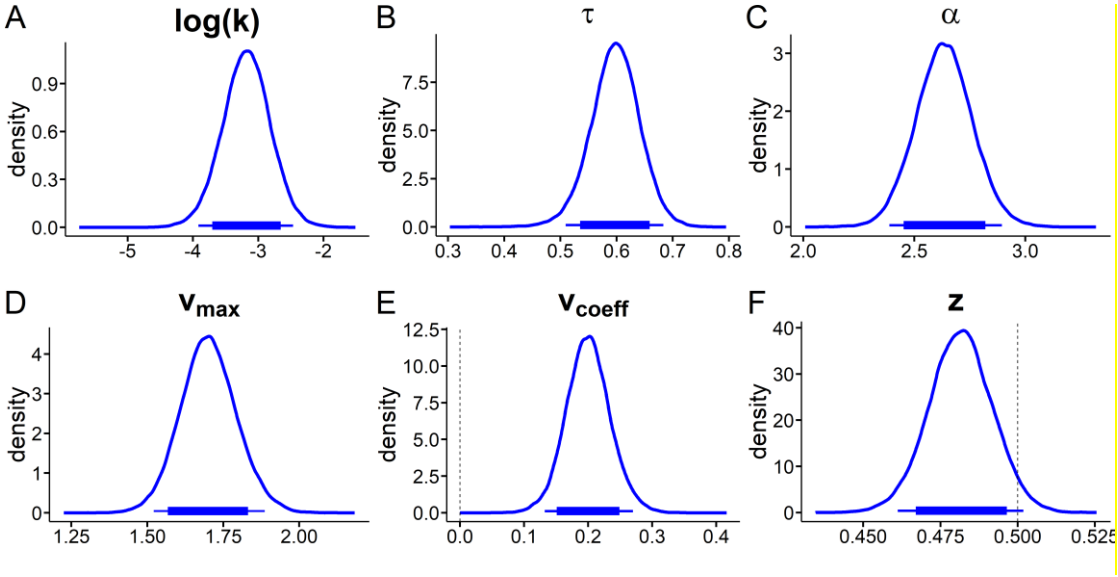


Figure 4. Temporal discounting drift diffusion model results: posterior distributions for hyperparameter means from the neutral context. A: discount-rate log(k), B: non-decision time τ, C: boundary separation α, D: maximum drift-rate v_{max}, E: drift-rate coefficient v_{coeff}, F: starting-point *z*. Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals

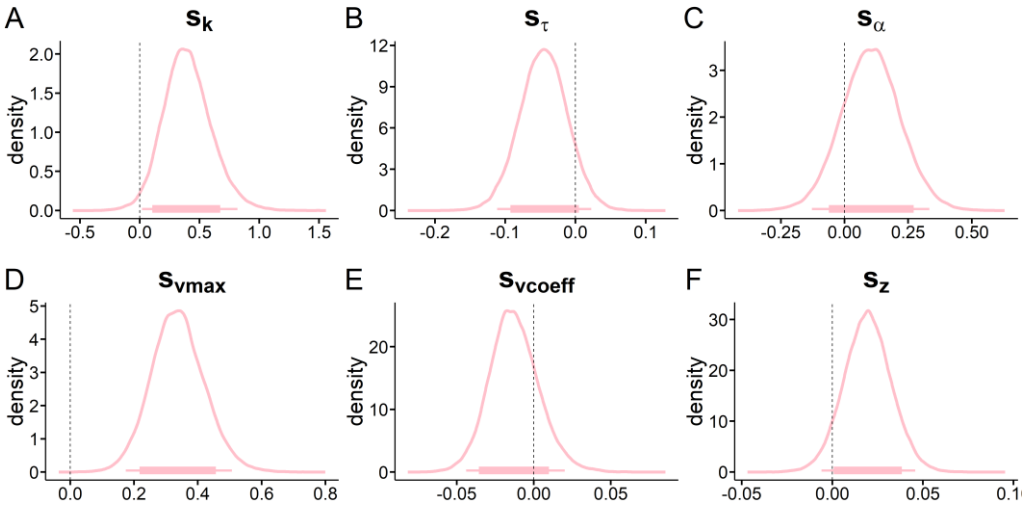


Figure 5. Temporal discounting drift diffusion model results: posterior distributions for hyperparameter means for context shift (s_x) parameters modeling changes from the neutral to the gambling context. A: shift in discount-rate (s_k), B: shift in non-decision time s_τ, C: shift in boundary separation s_α, D: shift in maximum drift-rate v_{max}, E: shift in drift-rate coefficient v_{coeff}, F: shift in starting-point s_z. Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals.

Overall context differences

We next examined the posterior distributions of model parameters of the best-fitting TD-DDM model (DDM_S). Results are plotted in Figure 4 and Figure 5 and Bayes Factors for all context-effects are listed in Table 3. There was a consistent positive association between trial-wise drift rates and value differences in the neutral context (Figure 4E, the 95% HDI for the drift rate coefficient parameter did not include 0). Likewise, there was a numerical bias towards the smaller-sooner option in the baseline condition (85% HDI < 0.5, see Figure 4F). The non-decision time was numerically smaller in the gambling context (85 % HDI < 0, Figure 5B, Table 3), amounting to on average a 50ms faster non-decision time. The maximum drift-rate was substantially higher in the gambling context (95% HDI > 0, Figure 5D).

As in the softmax model (Figure 3), we observed a substantial increase in the discount rates log(k) in the gambling context (95% HDI > 0, see Figure 5A, Table 3).

Table 3. Overview of overall context differences. For group comparisons we report Bayes Factors for directional effects for s_x hyperparameter distributions of $s_x > 0$ (gambling context > neutral context).

Model parameter (change in gambling context)	**Softmax Model**		**DDM_S**	
	Mean	***dBF***	***Mean***	***dBF***
s_k (discount-rate)	0.77	1688.53	0.40	54.20
s_β (softmax beta)	0.025	2.27	-	-
s_{vcoeff} (drift-rate coeff.)	-	-	-0.012	0.25
s_τ (non-decision time)	-	-	-0.05	0.10
s_α (boundary separation)	-	-	0.10	4.40
s_z (starting point bias)	-	-	0.02	13.64
s_{vmax} (max drift-rate)	-	-	0.33	39490.71

Temporal discounting and gambling-related questionnaire data

As preregistered, we next examined whether the increased in discount-rate s_k in the gambling context was associated with symptom severity or gambling related cognition. We therefore computed a compound symptom severity *z*-score of DSM-5 (Falkai, 2015), SOGS (Lesieur & Blume, 1987) and KFG (J. Petry & Baulig, 1996) questionnaire data. Gambling context related shifts in impulsive choice i.e. a positive shift in the discount-rate parameter was not related to

this score (rho = -0.05, p = 0.78). Increases in discounting (s_k) were positively associated with Gambling Related Cognitions Scale (Raylu & Oei, 2004) (rho = 0.39; p = 0.03); see Figure 6). Further analyses of temporal discounting and working memory are reported in the supplement.

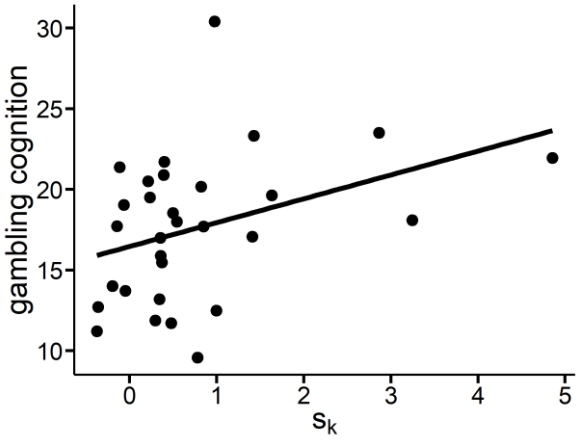


Figure 6. Relationship of total scores from the gambling-related cognition scale (GRCS) (Raylu & Oei, 2004) and changes in discount-rate from neutral to gambling environment (s_k)[softmax model].

Model-agnostic analysis 2-step task

Participants earned significantly more points in the gambling context (t-test: t_{28} = -2.44, p = 0.02, Cohen's d = 0.22). As a second model agnostic performance measure, we analyzed the effect of transition (common vs. rare) on S2 RTs. Longer RTs following rare transitions are an indirect measure for MB control (Otto et al., 2015; Shahar et al., 2019). We used an HGLM approach to model S2 RTs as a function of transition (rare vs. common) and context (neutral vs. gambling) allowing for interactions including subject as random effect. We observed a significant main effect of transition (see Table 4 and Figure 7) and a trend (p = 0.07; see Table 4) for a transition x context interaction. The latter reflected a tendency for greater RT increases following rare-transitions in the gambling context (see Figure 7). As a model-agnostic performance measure, the probability of choosing the same S1 option as in the previous trial (stay-probability) is typically analyzed as a function of reward, transition, and their interaction (Daw et al., 2011). Since the 2-step task version employed here utilized continuous payoffs, every trial was rewarded. The "reward" in S2 can thus not be used to directly predict stay probabilities, as done in previous work. In Table S3, we present an analogous regression model for stay probabilities using a moving average of recent rewards as a reference.

Table 4. Model agnostic analysis of S2 RTs. HGLM with transition and context as fixed effects and subject as random effect.

S2 RT Model			
	Estimate	**T-statistic**	p
Transition	0.13	17.227	< 2e-16***
Context	0.0004	0.07	0.94
Transition*Context	0.02	1.804	0.07

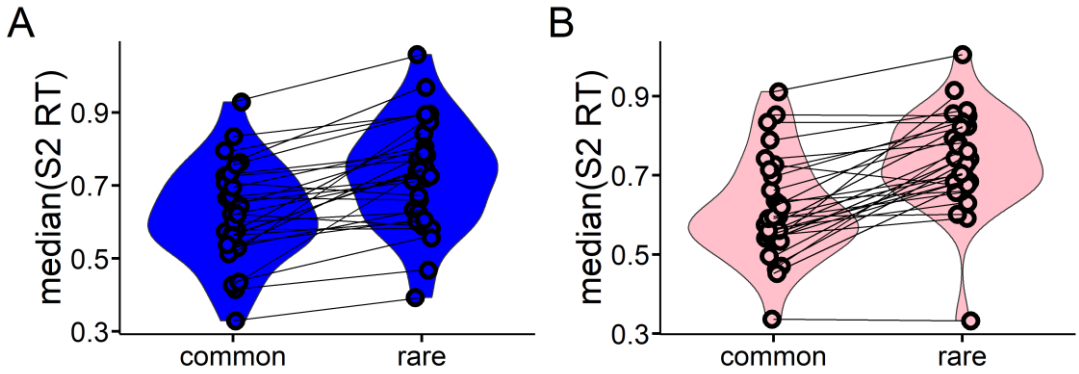


Figure 7. Model free analysis of S2 RTs. RTs were substantially slower following rare transitions, both in the neutral (A) and the gambling context (B), see also Table 4.

Hybrid model with softmax choice rule

We first examined a hybrid model as proposed by Daw et al. (Daw et al., 2011) with extensions by ourselves and Otto et al. (Otto et al., 2015) using a standard softmax choice rule (for details see methods; Figure 8). This model included separate parameters for S1 and S2 learning rates, model-free and model-based β weights for S1 and a β weight for S2 *Q*-value differences. We confirmed substantial contributions of both MB and MF values to S1 choices (Figure 8 B,C). There was an increase in the S2 learning-rate η (95% HDI > 0, Figure 8F) in the gambling context. Furthermore, there was a strong decrease in MF β weights (95% HDI < 0, Figure 8H) such that participants showed substantially less MF behavior in the gambling environment compared to the neutral environment. BFs for directional effects indicate that an increase in MB reinforcement learning is 4 times more likely than a decrease. For examination of Bayes Factors see Table 6.

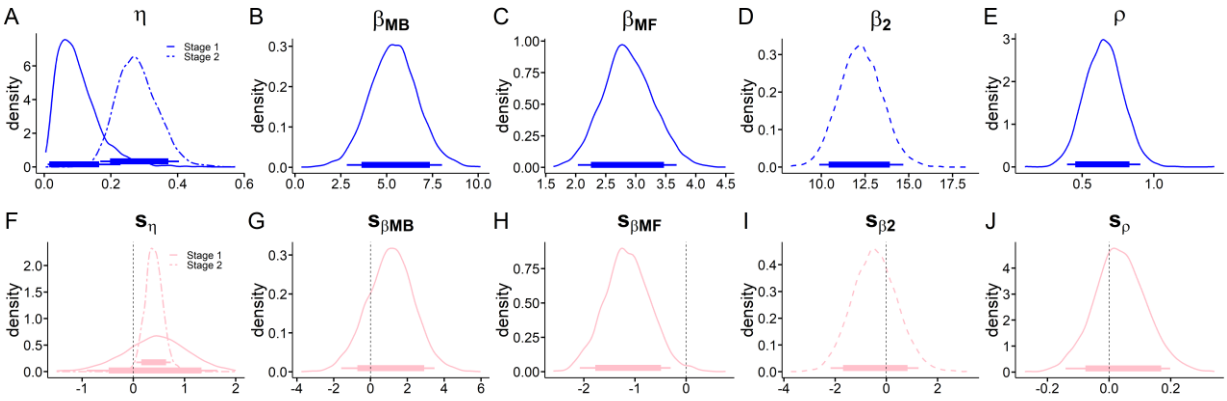


Figure 8. Hybrid model with softmax choice rule posterior distributions (top row: neutral context, bottom row: parameter changes in gambling context) of all group level means. A, S1 and S2 learning-rates. B, MB β weight. C, MF β weight. D, S2 β weight. E, perseveration parameter ρ. G, shift in MB β. H, shift in MF β. I, shift in S2 β. J, shift in stickiness parameter ρ. Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals.

Hybrid model with drift diffusion choice rule

We next combined the hybrid model with a DDM choice-rule (Shahar et al., 2019) and likewise compared DDMs that varied in the way that they accounted for the influence of *Q*-value differences on trial-wise drift rates in both task stages. Model comparison based on the WAIC (Vehtari et al., 2017) (see Table 5) revealed that in the neutral context, a DDM with nonlinear drift-rate scaling DDM_s (Fontanesi et al., 2019; Peters & D'Esposito, 2020; Wagner et al., 2020) accounted for the data best when compared to a DDM with linear drift rate scaling (DDM_{lin}) (Pedersen et al., 2017) and a null-model without learning (DDM0) (see Table 5). The same ranking held for the gambling context.

We next build a full model with group level distributions for the baseline condition (neutral context) and additional s_x parameters for each model parameter x, modeling the change in from the neutral to the gambling context. These s_x parameters where modeled with Gaussian priors with means of zero (see methods section). The full model reproduced the model ranking (see Table 5). We then compared the three DDMs and the softmax model with respect to the proportion of binary choices that they correctly accounted for. As can be seen from Table 7, the DDM_S and DDM_{lin} performed numerically on par with the softmax model, whereas the DDM_0 performed substantially worse. Posterior predictive checks showed that the final model accurately captured the effect of reward differences on second stage RTs and reproduced choice behavior (see Supplemental Figure S3 and S4).

Table 5. Reinforcement learning DDM model comparison using the Widely-Applicable Information Criterion (WAIC) revealed the same model ranking for each condition (neutral or gambling context) as well as for the full model. Scores are WAIC (SE).

	Neutral	**Gambling**	**Full model**
DDM_0	11808.5 (1549.3)	10942.1 (1569.2)	22764.0 (2193.7)
DDM_{lin}	4616.2 (1897.0)	4429.0 (1729.9)	15519.7 (3984.4)
DDM_s	4357.2 (1935.1)	4197.2 (1749.7)	8800.3 (2670.0)

Table 6. Overview of overall context differences. For group comparisons we report Bayes Factors or directional effects for s_x hyperparameter distributions of $s_x > 0$ (gambling context > neutral context).

Model parameter (shift)	**Softmax Model**		**DDMs**	
	Mean	***dBF***	***Mean***	***dBF***
$s_{\eta S1}$ (learning-rate S1)	0.44	3.29	0.0801	1.186
$s_{\eta S2}$ (learning-rate S2)	0.40	92.3	0.280	14.658
$s_{\tau S1}$ (non-decision times S1)	-	-	0.001	0.8454
$s_{\tau S2}$ (non-decision times S2)	-	-	0.001	1.161
s_{ρ} (Stickiness S1)	0.04	1.946	0.05	2.365
$s_{\alpha S1}$ (boundary separation S1)	-	-	-0.002	0.9354
$s_{\alpha S2}$ (boundary separation S2)	-	-	0.0149	2.026
β_{MF} / $S_{vcoeffMF}$ (MF beta/ drift-rate coeff.)	-1.14	0.010	-0.93	0.083
β_{MB} / $S_{vcoeffMB}$ (MB beta/ drift-rate coeff.)	1.08	4.00	4.01	169.62
β_{S2} / $S_{vcoeffS2}$ (S2 beta / drift-rate coeff.)	-0.44	0.428	-0.64	0.271
s_{vmaxS1} (max drift-rate S1)	-	-	-0.19	0.296
s_{vmaxS2} (max drift-rate S2)	-	-	0.41	15.83

Table 7. 2-step task models. Proportions of correctly predicted binary choices (mean [range]) for all models.

	Neutral		**Gambling**	
	Stage 1	Stage 2	Stage 1	Stage 2
DDM_0	0.63 [0.49-1.00]	0.62 [0.50-0.78]	0.56 [0.46-0.99]	0.63 [0.51-0.79]
DDM_{lin}	0.74 [0.51-1.00]	0.80 [0.53-0.96]	0.74 [0.50-0.99]	0.81 [0.59-0.95]
DDM_s	0.74 [0.49-1.00]	0.80 [0.55-0.96]	0.72 [0.47-0.99]	0.81 [0.59-0.95]
Softmax	0.72 [0.42-1.00]	0.79 [0.49-0.96]	0.72 [0.45-0.99]	0.81 [0.56-0.96]

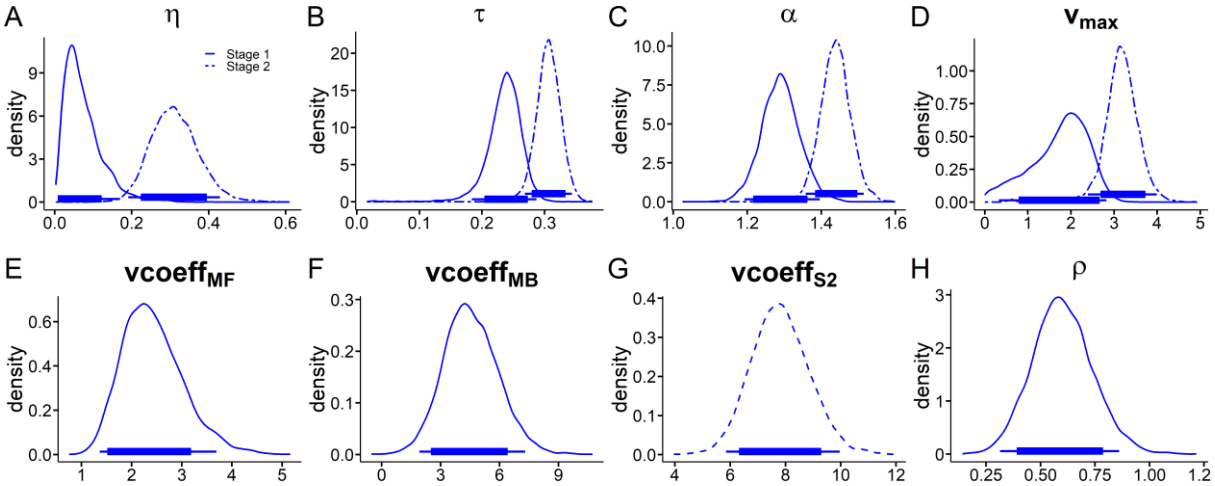


Figure 9. RL-DDM. Posterior distributions of all hyperparameters for the neutral baseline condition. A: S1 and S2 learning rates η. B: S1 and S2 non-decision time τ. C: S1 and S2 boundary separation α. D: S1 and S2 drift-rate maximum v_{max}. E: MF drift-rate coefficient $vcoeff_{MF}$. F: MB drift-rate coefficient $vcoeff_{MB}$. G: S2 drift-rate coefficient $vcoeff_{S2}$. H: stickiness parameter ρ. Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals.

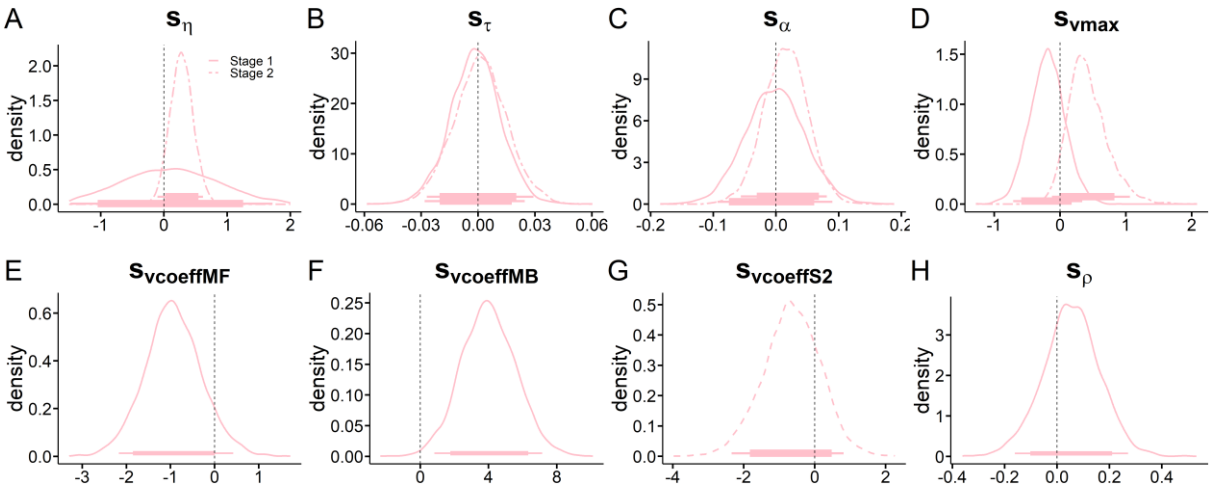


Figure 10. RL-DDM. Posterior distributions of all hyperparameters shift-parameters modelling the change from neutral to gambling condition. A, shift in Stage 1 and Stage 2 learning rates η. B, shift in S1 and S2 non-decision time τ. C, shift in S1 and S2 boundary separation α. D, shift in S1 and S2 drift-rate maximum v_{max}. E, shift in S1 MF drift-rate coefficient $vcoeff_{MF}$. F, shift in S1 MB drift-rate coefficient $vcoeff_{MB}$. G, shift in S2 drift-rate coefficient $vcoeff_{S2}$. H, shift in stickiness parameter ρ. Thin (thick) horizontal line denote 95% (85%) highest posterior density intervals.

Overall Context Differences

Posterior distributions for the best-fitting RL-DDM are shown in Figure 9 (neutral context parameters) and Figure 10 (gambling context changes). In general there was a positive association between trial-wise drift rates and *Q*-value differences (Figure 9F-J, all 95% HDIs above 0). Likewise, in the hybrid model with softmax choice-rule beta weights were positive indicating an effect of MB and MF *Q*-values on choice behavior (Fig9E-G, all 95% HDIs > 0). In terms of context related changes we observe a decrease in MF strategies and a robust increase in MB strategies. This is evident in terms of negative shifts in MF drift-rate coefficients (DDMs 85% HDI <= 0) and positive shifts in MB parameters, respectively (DDMs: 95% HDI > 0). We report BFs for directional effects in Table 6. Overall these results suggest a systematic association of gambling environments with decreased MF and increased MB reinforcement learning.

Reinforcement learning and gambling related questionnaire data

As preregistered, we examined associations between ρ (perseveration/stickyness) and gambling severity (average z-score across SOGS (Lesieur & Blume, 1987), KFG (J. Petry & Baulig, 1996) and DSM-5 criteria). The association was non-significant ρ (r = -0.10, p = 0.59). We next ran exploratory analyses to examine associations between MB/MF behavior ($S_{vcoeffMB}$ and $S_{vcoeffMF}$) and gambling severity (see above) and gambling-related cognition (GRCS) scores. Participants with higher gambling severity exhibited a greater reduction in MF learning in the gambling context (r = -0.48, p = 0.009; see Figure 11). The corresponding exploratory analyses on MB changes and gambling related cognitive distortions are provided in the Supplement (Reinforcement Learning section).

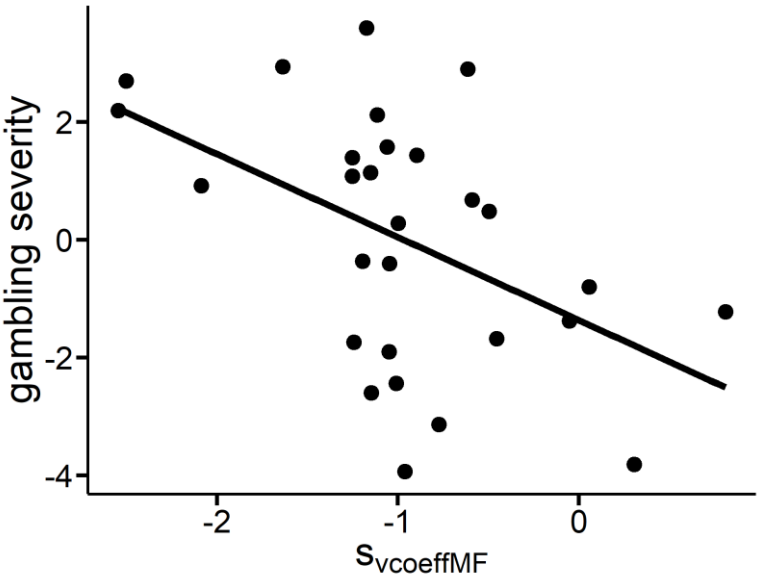


Figure 11. Gambling severity (y-axis; average z-score across DSM, KFG and SOGS) was associated with a greater gambling context related decrease in MF drift-rate weights ($S_{vcoeffMF}$, r = -0.48, p = 0.009).

Discussion

Here we comprehensively examined the contextual modulation of two putatively trans-diagnostic markers implicated in addiction, temporal discounting (Bickel et al., 2019; Lempert et al., 2019) and model-based control (Gillan et al., 2016; Gillan et al., 2020) in a pre-registered study. We studied regular slot machine gamblers, a group previously characterized by high levels of temporal discounting (Wiehler & Peters, 2015) and reduced model-based control (Wyckmans et al., 2019). Following a seminal study by Dixon et al. (Mark. Dixon, Jacobs, & Sanders, 2006), regular gamblers were tested in gambling environments (slot-machine venues) and neutral control environments. Gambling cue exposure modulated temporal discounting and

718 model-based control in gamblers in opposite ways: replicating Dixon et al., (2006), discounting
719 substantially increased in a gambling context. In contrast, model-based (MB) control improved
720 (increased). This differential modulation of two prominent trans-diagnostic traits in
721 (behavioral) addiction has important theoretical and clinical implications.

722 Theoretical accounts highlight the central role of addiction-related cues and
723 environments in drug addiction (T. Robinson & Berridge, 1993). Similar mechanisms have
724 been suggested to underlie gambling disorder (M. J. F. Robinson et al., 2016). Because
725 terrestrial slot machine gambling is directly linked to specific locations, gambling disorder is
726 uniquely suited to investigate the impact of cue exposure on behavior. We replicated the finding
727 of Dixon et al. (2006) of steeper discounting in gambling vs. neutral environments in gamblers.
728 This effect was observed across model agnostic analyses (proportion of LL choices) and
729 computational modeling (softmax, drift diffusion models [DDM]). We additionally extended
730 these earlier results in the following ways. First, we observed an association of this effect with
731 maladaptive control beliefs (GRCS) (Raylu & Oei, 2004) suggesting that such beliefs contribute
732 to increased temporal discounting in gambling environments. Second, in a subset of
733 participants, we confirmed that exposure to gambling environments substantially increases
734 subjective craving. Third, comprehensive modeling via DDMs revealed additional effects on
735 latent decision processes. The gambling context-related attenuation in non-decision time
736 mirrors previous effects of pharmacological enhancement of dopamine transmission (Wagner
737 et al. 2020). In contrast to these earlier pharmacological results, we observed a substantial
738 *increase* in maximum drift rate (V_{max}) in the gambling context, reflecting increased value
739 sensitivity of RTs. Lastly, our results complement cue-reactivity designs showing increased
740 impulsive and/or risky choice in gamblers during exposure to gambling cues in laboratory
741 studies (Dale et al., 2019; Genauck et al., 2020; Miedl et al., 2014). However, effect sizes during
742 naturalistic cue exposure (e.g. the present study and Dixon et al., 2006) were substantially larger
743 than during lab-based exposure in these previous studies.

744 In addition to temporal discounting, we included a 2-step sequential decision-making
745 task designed to dissociate model-based (MB) from model-free (MF) contributions to behavior
746 (Daw et al., 2011). Reductions in MB control are associated with compulsivity-related disorders
747 (Gillan et al., 2016; Gillan et al., 2020; V. Voon et al., 2015a). We observed increased MB
748 learning and reduced MF learning in gamblers in the gambling context, a pattern of results
749 consistent between softmax and DDM models. These findings were again corroborated my
750 model-agnostic analyses. First, participants earned more points in the gambling context, an
751 effect linked to MB learning (Kool et al., 2016). Second, the slowing of RTs following rare

752 transitions, an indirect measure for MB learning (Otto et al., 2015) tended to be more
753 pronounced in the gambling vs. neutral context. The MF effect correlated with gambling
754 severity in an exploratory analysis, such that higher gambling severity was associated with a
755 greater reduction in MF reinforcement learning in the gambling context. Together, these
756 findings converge on the picture of *decreased* MF and *increased* MB control in gamblers when
757 tested in gambling-related environments.

758 The latter result contrast with our pre-registered hypothesis of *reduced* MB control,
759 which was based on findings of reduced MB control in populations with extensive habit
760 formation (Gillan et al., 2016; Gillan et al., 2020; V. Voon et al., 2015b). Addiction is likewise
761 thought to be inherently associated with pathological habits (Barry J Everitt & Trevor W
762 Robbins, 2005; Robbins & Everitt, 1999) which are thought to be triggered by exposure to
763 environmental cues (Antons et al., 2020). We thus hypothesized gambling environments would
764 likewise trigger increased MF behavior and reduced MB behavior on the 2-step task. However,
765 critics of habit theory have emphasized that addiction might in contrast be associated with
766 excessive goal-directed behavior, in particular in the presence of addiction-related cues
767 (Hogarth 2020). Our findings are more in line with this latter perspective. This interpretation is
768 compatible with incentive sensitization theory (T. Robinson & Berridge, 1993; Terry E.
769 Robinson & Berridge, 2008), which proposes that addiction-related environments exert their
770 influence on behavior in part via a potentiation in dopamine release (Anselme & Robinson,
771 2013; Berridge, 2016; T. E. Robinson & Berridge, 2001). Earlier studies observed increased
772 MB control following increases in DA neurotransmission (Sharp et al., 2016; Wunderlich et
773 al., 2012), which could contribute to the present findings regarding 2-step task performance.
774 Furthermore, our results are compatible with decreased MF control under L-Dopa (Kroemer et
775 al., 2019). The gambling context might thus enhance goal-directed control via an improved
776 construction and/or utilization of the task transition structure. This interpretation further
777 resonates with other perspectives on DA function including a regulation of outcome sensitivity
778 or precision (FitzGerald et al., 2015; Shiner et al., 2012), or the general motivation to exert
779 (cognitive) effort (Berke, 2018). All of these perspectives are compatible with the idea that 2-
780 step task performance might improve with enhanced DA neurotransmission. The observed
781 increase in S2 learning rates could likewise be mediated in part by increases in DA transmission
782 (Frank & O'Reilly, 2006).

783 If the effects of gambling environments on 2-step task performance are (at least in part)
784 driven by increases in DA, then the question arises why gamblers at the same time exhibited
785 substantially increased temporal discounting. The literature on DA effects on temporal

discounting is a mixed bag (D'Amour-Horvat & Leyton, 2014) with some studies showing reduced discounting (van Gaalen et al., 2006; Wagner et al., 2020), some increased discounting (Pine et al., 2010) and others suggesting baseline-dependent effects (Petzold et al., 2019). Although we cannot conclusively address these questions in our study, it is plausible that DA might be context-dependent. In some situations DA might facilitate cognitive control (Ott & Nieder, 2019) whereas in addiction-related contexts it might facilitate the contrary (Terry E. Robinson & Berridge, 2008) i.e. approach and consumption. Given that DA facilitates execution of action-plans originating elsewhere in cortex (Frank & O'Reilly, 2006) it is thus theoretically plausible that it facilitates impulsive choice in the setting of addiction-related cues (Antons et al., 2020), when short-sighted cortical action schemas are activated. A further mechanism known to modulate temporal discounting is episodic future thinking or future prospection (Gershman & Bhui, 2020; Peters & Büchel, 2010). Future prospection has been shown to attenuate temporal discounting in a range of settings (Rösch et al., 2021) but might be attenuated at gambling venues. Participants might be focused on the present in the presence of cues or contexts endowed with high levels of incentive salience (Flagel et al., 2009).

Our results show that two prominent (potentially trans-diagnostic) computational processes, temporal discounting and MB control, are differentially modulated by addiction-related environments in regular slot machine gamblers. This provides a computational psychiatry perspective on factors that contribute to the understanding of this disorder. The substantial contextual effects on temporal discounting further highlight the potential clinical relevance of this process (Amlung et al., 2019; Lempert et al., 2019). Gambling disorder is reliably associated with increased temporal discounting (Mark. Dixon et al., 2003; Mark. Dixon, Jacobs, & Sanders, 2006; MacKillop et al., 2011; Miedl et al., 2012; Wiehler & Peters, 2015) . This trait-like behavior then appears to be further exacerbated during exposure to gambling-related environments, potentially contributing to the maintenance of maladaptive behavior. In contrast, MB control improved (increased) in a gambling context, despite the fact that an earlier study reported reduced MB control in gamblers (Wyckmans et al., 2019). In general these findings are further compatible with a greater tendency for pattern matching (Wilke et al., 2014) or enhanced cause-effect associations that might translate into increased MB control (Joukhador et al., 2004). 2-step task transitions are not random, but can be learned and exploited. An increased tendency to seek for patterns during gambling context exposure might facilitate this behavior. Our findings suggest that gamblers might not generally exhibit MB control impairments (Wyckmans et al., 2019). This is supported by the robust RTs increases observed following rare transitions (Table 4, Figure 7) and the positive MB

parameters observed across models, somewhat contrasting with the findings of Wyckmans et al. (Wyckmans et al., 2019), although different 2-step task versions have been used in these studies.

We also extended previous studies on this topic via a recent class of value-based decision models based on the DDM (Fontanesi et al., 2019; Pedersen et al., 2017; Peters & D'Esposito, 2020; Shahar et al., 2019; Wagner et al., 2020). Comprehensive RT-based analysis revealed that standard DDM parameters were largely unaffected by context, suggesting that primarily MF and MB contributions to evidence accumulation were affected by gambling environments (Figure 10.). Posterior predictive checks showed that a DDM with non-linear trial-wise drift rate scaling captured the relationship of decision conflict (SS-LL value difference) and RTs, replicating prior findings (Peters & D'Esposito, 2020; Wagner et al., 2020). We previously reported good parameter recovery of such temporal discounting DDMs (Peters & D'Esposito, 2020; Wagner et al., 2020).

A number of limitations need to be acknowledged. First, as in the original study (Mark. Dixon, Jacobs, & Sanders, 2006) we did not test a non-gambling control group. However, the observed associations between experimental effects and gambling severity / gambling-related cognition suggests that these effects are at least in part driven by the underlying problem gambling symptoms. Second, MB and MF effects in the 2-step task might be affected by instructions (da Silva & Hare, 2019). Participants in our study were well instructed in written and verbal form and completed extensive training trials. Moreover, due to the counter-balanced within-subject design, the observed context-dependent changes in MB/MF behavior cannot be attributed to overall instruction effects. Third, MB control might more generally be related to attentional or motivational processes. For example, incentives boost 2-step task performance (Patzelt et al., 2019). However, we ensured that mean and variance of reward walks as well as incentives were identical in both contexts. Fourth, although participants were tested in the same venues, the number of customers present varied across participants, affecting e.g. noise levels and auditory gambling cues (slot machine sounds etc.). A trade-off between the control of such variables and ecological validity is unavoidable when testing in naturalistic settings. Finally, DA neurotransmission was obviously not assessed, rendering our interpretation of the effects in terms of the incentive sensitization theory speculative. But the substantial increase in subjective craving supports the idea that cue exposure had subjective effects predicted by incentive sensitization.

To conclude, here we show that two computational trans-diagnostic markers with high relevance for gambling disorder in particular and addiction more generally are modulated in

854 opposite ways by exposure to real gambling environments. Gamblers showed increased
855 temporal discounting in a gambling context, and this effect was modulated by maladaptive
856 control beliefs. In contrast, MB control improved, a finding that posits a challenge for
857 habit/compulsion theories of addiction. Ecologically valid testing settings such as those
858 investigated here can thus yield novel insights into environmental drivers of maladaptive
859 behavior underlying mental disorders.
860
861

Acknowledgements

This work was supported by Deutsche Forschungsgemeinschaft (PE1627/5-1 to JP)

Author contributions

JP conceived the idea and acquired the funding. JP and BJW designed the study. BJW acquired the data. BJW analyzed the data and performed the modeling. DM contributed analytical tools/software. BJW wrote the paper. JP and DM provided revisions. JP supervised the project.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

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), 1176–1186. https://doi.org/10.1001/jamapsychiatry.2019.2102

Anderson, G., & Brown, R. I. (1984). Real and laboratory gambling, sensation-seeking and arousal. *British Journal of Psychology (London, England : 1953)*, *75 (Pt 3)*, 401–410. https://doi.org/10.1111/j.2044-8295.1984.tb01910.x

Anselme, P., & Robinson, M. J. F [Mike J. F.] (2013). What motivates gambling behavior? Insight into dopamine's role. *Frontiers in Behavioral Neuroscience*, *7*, 182. https://doi.org/10.3389/fnbeh.2013.00182

Antons, S., Brand, M., & Potenza, M. N. (2020). Neurobiology of cue-reactivity, craving, and inhibitory control in non-substance addictive behaviors. *Journal of the Neurological Sciences*, *415*, 116952. https://doi.org/10.1016/j.jns.2020.116952

887 Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American
888 Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

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

893 Balodis, I. M., & Potenza, M. N. (2020). Common neurobiological and psychological
894 underpinnings of gambling and substance-use disorders. *Progress in Neuro-*
895 *Psychopharmacology & Biological Psychiatry*, *99*, 109847.
896 https://doi.org/10.1016/j.pnpbp.2019.109847

897 Barry J Everitt, & Trevor W Robbins (2005). Neural systems of reinforcement for drug
898 addiction: From actions to habits to compulsion. *Nature Neuroscience*, *8*(11), 1481–1489.
899 https://doi.org/10.1038/nn1579

900 Beck, A., Steer, R., & Brown, G. (1996). Manual for the beck depression inventory-ii (bdi-ii).

901 Berke, J. D. (2018). What does dopamine mean? *Nature Neuroscience*, *21*(6), 787–793.
902 https://doi.org/10.1038/s41593-018-0152-y

903 Berridge, K. C [K. C.]. (2016). Incentive motivation and incentive salience☆. In J. Stein
904 (Ed.), *Reference module in neuroscience and biobehavioral psychology*. Elsevier.
905 https://doi.org/10.1016/B978-0-12-809324-5.00342-4

906 Bickel, W. K., Athamneh, L. N., Basso, J. C., Mellis, A. M., DeHart, W. B., Craft, W. H., &
907 Pope, D. (2019). Excessive discounting of delayed reinforcers as a trans-disease process:
908 Update on the state of the science. *Current Opinion in Psychology*, *30*, 59–64.
909 https://doi.org/10.1016/j.copsyc.2019.01.005

910 Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Koffarnus, M. N., & Gatchalian, K. M.
911 (2012). Excessive discounting of delayed reinforcers as a trans-disease process
912 contributing to addiction and other disease-related vulnerabilities: Emerging evidence.
913 *Pharmacology & Therapeutics*, *134*(3), 287–297.
914 https://doi.org/10.1016/j.pharmthera.2012.02.004

915 Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., & Baxter, C. (2011). Remember the future:
916 Working memory training decreases delay discounting among stimulant addicts.
917 *Biological Psychiatry*, *69*(3), 260–265. https://doi.org/10.1016/j.biopsych.2010.08.017

918 Blaszczynski, A., & Nower, L. (2002). A pathways model of problem and pathological
919 gambling. *Addiction (Abingdon, England)*, *97*(5), 487–499. https://doi.org/10.1046/j.1360-
920 0443.2002.00015.x

921 Clark, L., Boileau, I., & Zack, M. (2019). Neuroimaging of reward mechanisms in gambling
922 disorder: An integrative review. *Molecular Psychiatry*, *24*(5), 674–693.
923 https://doi.org/10.1038/s41380-018-0230-2

924 Courtney, K. E., Schacht, J. P., Hutchison, K., Roche, D. J. O., & Ray, L. A. (2016). Neural
925 substrates of cue reactivity: Association with treatment outcomes and relapse. *Addiction*
926 *Biology*, *21*(1), 3–22. https://doi.org/10.1111/adb.12314

927 Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M., & Barch, D. M. (2016). Reduced
928 model-based decision-making in schizophrenia. *Journal of Abnormal Psychology*, *125*(6),
929 777–787. https://doi.org/10.1037/abn0000164

930 da Silva, C. F., & Hare, T. A. (2019). *Humans are primarily model-based and not model-free*
931 *learners in the two-stage task* (Vol. 20). https://doi.org/10.1101/682922

932 Dale, G., Rock, A. J., & Clark, G. I. (2019). Cue-reactive imagery mediates the relationships
933 of reward responsiveness with both cue-reactive urge to gamble and positive affect in

poker-machine gamblers. *Journal of Gambling Studies.* Advance online publication. https://doi.org/10.1007/s10899-019-09864-x

D'Amour-Horvat, V., & Leyton, M. (2014). Impulsive actions and choices in laboratory animals and humans: Effects of high vs. Low dopamine states produced by systemic treatments given to neurologically intact subjects. *Frontiers in Behavioral Neuroscience*, *8*, 432. https://doi.org/10.3389/fnbeh.2014.00432

Daw, N., Gershman, S., Seymour, B., Dayan, P [Peter], & 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

Dixon, M [Mark], Ghezzi, P., Lyons, C., & Wilson, G. (2006). *Gambling: Behavior Theory, Research, and Application*. New Harbinger Publications.

Dixon, M [Mark.], Jacobs, E. A., & Sanders, S. (2006). Contextual control of delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, *39*(4), 413–422. https://doi.org/10.1901/jaba.2006.173-05

Dixon, M [Mark.], Marley, J., & Jacobs, E. A. (2003). Delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, *36*(4), 449–458. https://doi.org/10.1901/jaba.2003.36-449

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *22*(6), 1075–1081. https://doi.org/10.1016/j.conb.2012.08.003

Falkai, P. (2015). *Diagnostisches und statistisches Manual psychischer Störungen: DSM-5*. Hogrefe.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

959 Fauth-Bühler, M., Mann, K., & Potenza, M. N. (2017). Pathological gambling: A review of
960 the neurobiological evidence relevant for its classification as an addictive disorder.
961 *Addiction Biology*, *22*(4), 885–897. https://doi.org/10.1111/adb.12378

962 FitzGerald, T. H. B., Dolan, R. J., & Friston, K. (2015). Dopamine, reward learning, and
963 active inference. *Frontiers in Computational Neuroscience*, *9*, 136.
964 https://doi.org/10.3389/fncom.2015.00136

965 Flagel, S. B., Akil, H., & Robinson, T. E [Terry E.] (2009). Individual differences in the
966 attribution of incentive salience to reward-related cues: Implications for addiction.
967 *Neuropharmacology*, *56 Suppl 1*, 139–148.
968 https://doi.org/10.1016/j.neuropharm.2008.06.027

969 Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning
970 diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*,
971 *26*(4), 1099–1121. https://doi.org/10.3758/s13423-018-1554-2

972 Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in
973 cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of*
974 *Psychology*, *67*, 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

975 Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in
976 human cognition: Psychopharmacological studies with cabergoline and haloperidol.
977 *Behavioral Neuroscience*, *120*(3), 497–517. https://doi.org/10.1037/0735-7044.120.3.497

978 Genauck, A., Andrejevic, M., Brehm, K., Matthis, C., Heinz, A., Weinreich, A.,
979 Kathmann, N., & Romanczuk-Seiferth, N. (2020). Cue-induced effects on decision-making
980 distinguish subjects with gambling disorder from healthy controls. *Addiction Biology*,
981 *25*(6), e12841. https://doi.org/10.1111/adb.12841

982 Gershman, S. J., & Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature*
983 *Communications*, *11*(1), 3365. https://doi.org/10.1038/s41467-020-16852-y

984 Gillan, C. M., Kalanthroff, E., Evans, M., Weingarden, H. M., Jacoby, R. J., Gershkovich, M.,
985 Snorrason, I., Campeas, R., Cervoni, C., Crimarco, N. C., Sokol, Y., Garnaat, S. L.,
986 McLaughlin, N. C. R., Phelps, E. A., Pinto, A., Boisseau, C. L., Wilhelm, S., Daw, N. D.,
987 & Simpson, H. B [H. B.] (2020). Comparison of the association between goal-directed
988 planning and self-reported compulsivity vs obsessive-compulsive disorder diagnosis.
989 *JAMA Psychiatry*, *77*(1), 77–85. https://doi.org/10.1001/jamapsychiatry.2019.2998

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

993 Hautzinger, M., Keller, F., & Kühner, C. (2009). Bdi-ii. Beck-depressions-inventar. Revision.
994 2, auflage.

995 Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The
996 fagerström test for nicotine dependence: A revision of the fagerstrom tolerance
997 questionnaire. *British Journal of Addiction*, *86*(9), 1119–1127.

998 Joukhador, J., Blaszczynski, A., & Maccallum, F. (2004). Superstitious beliefs in gambling
999 among problem and non-problem gamblers: Preliminary data. *Journal of Gambling*
1000 *Studies*, *20*(2), 171–180. https://doi.org/10.1023/B:JOGS.0000022308.27774.2b

1001 Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When does model-based control pay
1002 off? *PLOS Computational Biology*, *12*(8), e1005090.
1003 https://doi.org/10.1371/journal.pcbi.1005090

1004 Kroemer, N. B., Lee, Y., Pooseh, S., Eppinger, B., Goschke, T., & Smolka, M. N. (2019). L-
1005 dopa reduces model-free control of behavior by attenuating the transfer of value to action.
1006 *NeuroImage*, *186*, 113–125. https://doi.org/10.1016/j.neuroimage.2018.10.075

1007 Leeman, R. F., & Potenza, M. N. (2012). Similarities and differences between pathological
1008 gambling and substance use disorders: a focus on impulsivity and compulsivity.
1009 *Psychopharmacology*, *219*(2), 469–490. https://doi.org/10.1007/s00213-011-2550-7

1010 Lempert, K. M., & Phelps, E. A. (2016). The malleability of intertemporal choice. *Trends in*
1011 *Cognitive Sciences*, *20*(1), 64–74. https://doi.org/10.1016/j.tics.2015.09.005

1012 Lempert, K. M., Steinglass, J. E., Pinto, A., Kable, J. W., & Simpson, H. B [Helen Blair]
1013 (2019). Can delay discounting deliver on the promise of rdoc? *Psychological Medicine*,
1014 *49*(2), 190–199. https://doi.org/10.1017/S0033291718001770

1015 Lesieur, H. R., & Blume, S. B. (1987). The south oaks gambling screen (sogs): A new
1016 instrument for the identification of pathological gamblers. *American Journal of Psychiatry*,
1017 *144*(9).

1018 Lobo, D. S. S., Aleksandrova, L., Knight, J., Casey, D. M., el-Guebaly, N., Nobrega, J. N., &
1019 Kennedy, J. L. (2015). Addiction-related genes in gambling disorders: new insights from
1020 parallel human and pre-clinical models. *Molecular Psychiatry*, *20*(8), 1002.
1021 https://doi.org/10.1038/mp.2014.113

1022 MacKillop, J., Amlung, M. T., Few, L. R., Ray, L. A., Sweet, L. H., & Munafò, M. R. (2011).
1023 Delayed reward discounting and addictive behavior: A meta-analysis.
1024 *Psychopharmacology*, *216*(3), 305–321. https://doi.org/10.1007/s00213-011-2229-0

1025 Martyn Plummer. (2003). Jags: A program for analysis of bayesian graphical models using
1026 gibbs sampling. In M. Plummer (Ed.), *Jags: a program for analysis of bayesian graphical*
1027 *models using gibbs sampling.: proceedings of the 3rd international workshop on*
1028 *distributed statistical computing (dsc 2003)*. 2003.

1029 Miedl, S. F., Büchel, C., & Peters, J. (2014). Cue-induced craving increases impulsivity via
1030 changes in striatal value signals in problem gamblers. *The Journal of Neuroscience : The*

Official Journal of the Society for Neuroscience, *34*(13), 4750–4755. https://doi.org/10.1523/JNEUROSCI.5020-13.2014

Miedl, S. F., Peters, J., & Büchel, C. (2012). Altered neural reward representations in pathological gamblers revealed by delay and probability discounting. *Archives of General Psychiatry*, *69*(2), 177–186. https://doi.org/10.1001/archgenpsychiatry.2011.1552

Moeller, S. J., & Paulus, M. P. (2018). Toward biomarkers of the addicted human brain: Using neuroimaging to predict relapse and sustained abstinence in substance use disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *80*(Pt B), 143–154. https://doi.org/10.1016/j.pnpbp.2017.03.003

Muggleton, N., Parpart, P., Newall, P., Leake, D., Gathergood, J., & Stewart, N. (2021). The association between gambling and financial, social and health outcomes in big financial data. *Nature Human Behaviour.* Advance online publication. https://doi.org/10.1038/s41562-020-01045-w

Nelson, L. D., Strickland, C., Krueger, R. F., Arbisi, P. A., & Patrick, C. J. (2016). Neurobehavioral traits as transdiagnostic predictors of clinical problems. *Assessment*, *23*(1), 75–85. https://doi.org/10.1177/1073191115570110

Ott, T., & Nieder, A. (2019). Dopamine and cognitive control in prefrontal cortex. *Trends in Cognitive Sciences*, *23*(3), 213–234. https://doi.org/10.1016/j.tics.2018.12.006

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

Patzelt, E. H., Kool, W., Millner, A. J., & Gershman, S. J. (2019). Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biological Psychiatry*, *85*(5), 425–433. https://doi.org/10.1016/j.biopsych.2018.06.018

1055 Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule
1056 in reinforcement learning. *Psychonomic Bulletin & Review*, *24*(4), 1234–1251.
1057 https://doi.org/10.3758/s13423-016-1199-y

1058 Peters, J., & Büchel, C. (2010). Episodic future thinking reduces reward delay discounting
1059 through an enhancement of prefrontal-mediotemporal interactions. *Neuron*, *66*(1), 138–
1060 148. https://doi.org/10.1016/j.neuron.2010.03.026

1061 Peters, J., & Büchel, C. (2011). The neural mechanisms of inter-temporal decision-making:
1062 understanding variability. *Trends in Cognitive Sciences*, *15*(5), 227–239.
1063 https://doi.org/10.1016/j.tics.2011.03.002

1064 Peters, J., & D'Esposito, M. (2020). The drift diffusion model as the choice rule in inter-
1065 temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and
1066 controls. *PLOS Computational Biology*, *16*(4), e1007615.
1067 https://doi.org/10.1371/journal.pcbi.1007615

1068 Petry, J., & Baulig, T. (1996). Kfg: Kurzfragebogen zum glücksspielverhalten.
1069 *Psychotherapie Der Glücksspielsucht*, 300–302.

1070 Petry, N. M. (2010). Pathological gambling and the dsm-v. *International Gambling Studies*,
1071 *10*(2), 113–115. https://doi.org/10.1080/14459795.2010.501086

1072 Petzold, J., Kienast, A., Lee, Y., Pooseh, S., London, E. D., Goschke, T., & Smolka, M. N.
1073 (2019). Baseline impulsivity May moderate l-dopa effects on value-based decision-
1074 making. *Scientific Reports*, *9*(1), 5652. https://doi.org/10.1038/s41598-019-42124-x

1075 Pine, A., Shiner, T., Seymour, B., & Dolan, R. J. (2010). Dopamine, time, and impulsivity in
1076 humans. *The Journal of Neuroscience : The Official Journal of the Society for*
1077 *Neuroscience*, *30*(26), 8888–8896. https://doi.org/10.1523/JNEUROSCI.6028-09.2010

1078 Raylu, N., & Oei, T. P. S. (2004). The gambling related cognitions scale (grcs): Development,
1079 confirmatory factor validation and psychometric properties. *Addiction (Abingdon,*
1080 *England)*, *99*(6), 757–769.

1081 Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., &
1082 Engle, R. W. (2012). Measuring working memory capacity with automated complex span
1083 tasks. *European Journal of Psychological Assessment*, *28*(3), 164–171.
1084 https://doi.org/10.1027/1015-5759/a000123

1085 Reynolds, B. (2006). A review of delay-discounting research with humans: Relations to drug
1086 use and gambling. *Behavioural Pharmacology*, *17*(8), 651–667.
1087 https://doi.org/10.1097/FBP.0b013e3280115f99

1088 Robbins, T. W., & Everitt, B. J. (1999). Drug addiction: Bad habits add up. *Nature*,
1089 *398*(6728), 567–570. https://doi.org/10.1038/19208

1090 Robinson, M. J. F [M. J. F.], Fischer, A. M., Ahuja, A., Lesser, E. N., & Maniates, H. (2016).
1091 Roles of "wanting" and "liking" in motivating behavior: Gambling, food, and drug
1092 addictions. *Current Topics in Behavioral Neurosciences*, *27*, 105–136.
1093 https://doi.org/10.1007/7854_2015_387

1094 Robinson, T., & Berridge, K. C [Kent C.] (1993). The neural basis of drug craving: An
1095 incentive-sensitization theory of addiction. *Brain Research Reviews*, *18*(3), 247–291.
1096 https://doi.org/10.1016/0165-0173(93)90013-P

1097 Robinson, T. E., & Berridge, K. C [K. C.] (2001). Incentive-sensitization and addiction.
1098 *Addiction (Abingdon, England)*, *96*(1), 103–114. https://doi.org/10.1046/j.1360-
1099 0443.2001.9611038.x

1100 Robinson, T. E [Terry E.], & Berridge, K. C [Kent C.] (2008). Review. The incentive
1101 sensitization theory of addiction: Some current issues. *Philosophical Transactions of the*

Royal Society B: Biological Sciences, *363*(1507), 3137–3146. https://doi.org/10.1098/rstb.2008.0093

Rösch, S. A., Stramaccia, D. F., & Benoit, R. G. (2021). *Promoting farsighted decisions via episodic future thinking: A meta-analysis.* https://doi.org/10.31234/osf.io/53ju2

Saunders, J. B., Aasland, O. G., Babor, T. F., La Fuente, J. R. de, & Grant, M. (1993). Development of the alcohol use disorders identification test (audit): Who collaborative project on early detection of persons with harmful alcohol consumption-ii. *Addiction (Abingdon, England)*, *88*(6), 791–804.

Sebold, M., Deserno, L., Nebe, S [Stephan], Nebe, S [Stefan], Schad, D. J., Garbusow, M., Hägele, C., Keller, J., Jünger, E., Kathmann, N., Smolka, M. N., Smolka, M., Rapp, M. A., Schlagenhauf, F., Heinz, A., & Huys, Q. J. M. (2014). Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*, *70*(2), 122–131. https://doi.org/10.1159/000362840

Shahar, N., Hauser, T. U., Moutoussis, M., Moran, R., Keramati, M., & Dolan, R. J. (2019). Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Computational Biology*, *15*(2), e1006803. https://doi.org/10.1371/journal.pcbi.1006803

Sharp, M. E., Foerde, K., Daw, N. D., & Shohamy, D. (2016). Dopamine selectively remediates 'model-based' reward learning: A computational approach. *Brain : A Journal of Neurology*, *139*(Pt 2), 355–364. https://doi.org/10.1093/brain/awv347

Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K. P., Dayan, P [Peter], & Dolan, R. J. (2012). Dopamine and performance in a reinforcement learning task: Evidence from parkinson's disease. *Brain : A Journal of Neurology*, *135*(Pt 6), 1871–1883. https://doi.org/10.1093/brain/aws083

1126 Singer, B. F., Anselme, P., Robinson, M. J. F [Mike J. F.], & Vezina, P. (2020). An overview
1127 of commonalities in the mechanisms underlying gambling and substance use disorders.
1128 *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *101*, 109944.
1129 https://doi.org/10.1016/j.pnpbp.2020.109944

1130 Stan Development Team. (2020). *Stan* (Version 2.26) [Computer software]. https://mc-
1131 stan.org

1132 Toyama, A., Katahira, K., & Ohira, H. (2017). A simple computational algorithm of model-
1133 based choice preference. *Cognitive, Affective & Behavioral Neuroscience*, *17*(4), 764–783.
1134 https://doi.org/10.3758/s13415-017-0511-2

1135 Toyama, A., Katahira, K., & Ohira, H. (2019). Reinforcement learning with parsimonious
1136 computation and a forgetting process. *Frontiers in Human Neuroscience*, *13*, 153.
1137 https://doi.org/10.3389/fnhum.2019.00153

1138 van den Noort, M., Bosch, P., Haverkort, M., & Hugdahl, K. (2008, January 21). *A standard*
1139 *computerized version of the reading span test in different languages.* Hogrefe & Huber
1140 Publishers. https://econtent.hogrefe.com/doi/abs/10.1027/1015-5759.24.1.35

1141 van Gaalen, M. M., van Koten, R., Schoffelmeer, A. N. M., & Vanderschuren, L. J. M. J.
1142 (2006). Critical involvement of dopaminergic neurotransmission in impulsive decision
1143 making. *Biological Psychiatry*, *60*(1), 66–73.
1144 https://doi.org/10.1016/j.biopsych.2005.06.005

1145 Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-
1146 one-out cross-validation and waic. *Statistics and Computing*, *27*(5), 1413–1432.
1147 https://doi.org/10.1007/s11222-016-9696-4

1148 Voon, V [V.], Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J.,
1149 Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W.,
1150 Harrison, N. A., Wood, J., Daw, N. D [N. D.], Dayan, P [P.], Grant, J. E., &

1151 Bullmore, E. T. (2015a). Disorders of compulsivity: A common bias towards learning
1152 habits. *Molecular Psychiatry*, *20*(3), 345–352. https://doi.org/10.1038/mp.2014.44

1153 Voon, V [V.], Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J.,
1154 Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W.,
1155 Harrison, N. A., Wood, J., Daw, N. D [N. D.], Dayan, P [P.], Grant, J. E., &
1156 Bullmore, E. T. (2015b). Disorders of compulsivity: A common bias towards learning
1157 habits. *Molecular Psychiatry*, *20*(3), 345–352. https://doi.org/10.1038/mp.2014.44

1158 Voon, V [Valerie], Reiter, A., Sebold, M., & Groman, S. (2017). Model-based control in
1159 dimensional psychiatry. *Biological Psychiatry*, *82*(6), 391–400.
1160 https://doi.org/10.1016/j.biopsych.2017.04.006

1161 Wagner, B., Clos, M., Sommer, T., & Peters, J. (2020). Dopaminergic modulation of human
1162 intertemporal choice: A diffusion model analysis using the d2-receptor antagonist
1163 haloperidol. *The Journal of Neuroscience : The Official Journal of the Society for*
1164 *Neuroscience*, *40*(41), 7936–7948. https://doi.org/10.1523/JNEUROSCI.0592-20.2020

1165 Wiehler, A., & Peters, J. (2015). Reward-based decision making in pathological gambling:
1166 The roles of risk and delay. *Neuroscience Research*, *90*, 3–14.
1167 https://doi.org/10.1016/j.neures.2014.09.008

1168 Wilke, A., Scheibehenne, B., Gaissmaier, W., McCanney, P., & Barrett, H. C. (2014).
1169 Illusionary pattern detection in habitual gamblers. *Evolution and Human Behavior*, *35*(4),
1170 291–297. https://doi.org/10.1016/j.evolhumbehav.2014.02.010

1171 Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over
1172 model-free choice behavior. *Neuron*, *75*(3), 418–424.
1173 https://doi.org/10.1016/j.neuron.2012.03.042

1174 Wyckmans, F., Otto, A. R., Sebold, M., Daw, N., Bechara, A., Saeremans, M., Kornreich, C.,
1175 Chatard, A., Jaafari, N., & Noël, X. (2019). Reduced model-based decision-making in

gambling disorder. *Scientific Reports*, *9*(1), 19625. https://doi.org/10.1038/s41598-019-56161-z

1185 **Supplemental Information**

1186 **Supplemental Table S1.** Baseline screening questionnaires.

	Reference	Measure
AUDIT	(Saunders et al., 1993)	Alcohol-Use-Disorders Identification Text
BDI	(Beck et al., 1996; Hautzinger et al., 2009)	Beck Depression Inventory II
DSM-5	(American Psychiatric Association, 2013; Falkai, 2015)	Diagnostic criteria for gambling disorder
FTND	(Heatherton et al., 1991)	Fagerström Test for Nicotine Dependence
GRCS	(Raylu & Oei, 2004)	Gambling-related cognition scale
KFG	(J. Petry & Baulig, 1996)	Kurzfragebogen zum Glücksspielverhalten
SOGS	(Lesieur & Blume, 1987)	South Oaks Gambling Screen

1187

1188

1189

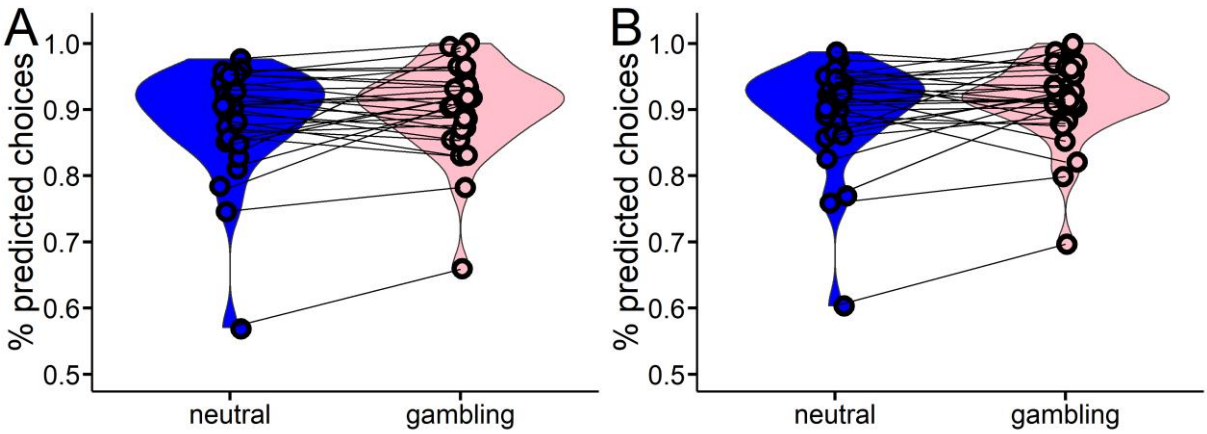
1190

1191

1192 **Supplemental Table S2.** Summary of demographics and clinical information (n = 30).

	Mean	SD	Range
Age	30.87	7.94	20-47
School years	11.83	1.93	8-17
Income	1284.73	590.71	300-2500
FTND	2.5	3.01	0-9
AUDIT	7.6	7.10	0-26
DSM-5	5.9	2.02	3-9
KFG	27.54	10.31	7-54
SOGS	9.5	3.87	3-16
BDI	17.27	10.13	0-45
GRCS	17.60	4.54	9.57-30.4

1193



1194

1195 **Supplemental Figure S1.** Proportions of correctly predicted binary choices for the softmax
1196 model (A) and the drift diffusion model with non-linear drift rate scaling (B, DDMs) in both
1197 contexts (neutral [blue], gambling [pink]).

1198

Model-free analysis 2-step task

We here present a model-agnostic analysis of stay probability in the present 2-step task version with continuous rewards at S2. First, we categorized a reward R_{t-1} as positive "R+" if R_{t-1} was higher than the mean of last 7 rewards ($R_t > \text{mean}[R_{t-1:\,t-7}]$) and as negative "R-" if $R_t < \text{mean}(R_{t-1:t-7})$. Stay probabilities were constructed in analogy with the original task (Daw et al., 2011), that is stay was set to 1 if the subject choose the same S1 option as before and 0 otherwise. We then fitted separate hierarchical general linear models (HGLMs) per condition, with stay probability as dependent variable and reward and transition (common vs. rare) as fixed effects and subject as random effect (see Supplemental Table S3), as well as a full model that additionally included a context factor as fixed effect.

The full model confirmed the expected main effect of reward (z = 4.34, p = 1.36e-05), transition (z = 3.91 p = 7.04e-05) and the reward*transition interaction (z = -4.95 , p =4.48e-07) . Using this analysis the reward*transition*context interaction was not statistically significant, even though it numerically suggests increased probability to switch after a "R+" ($R_t > \text{mean}[R_{t-1:\,t-7}]$) and a rare transition in the gambling context when compared to the neutral context. (see Supplemental Table S3).

Supplemental Table S3. Model agnostic analysis of stay probability using an hierarchical general linear model (HGLM). HGLMs were estimated for each context separately using reward and transition as fixed and subject as random effects. The fill model model with stay probability as dependent variable includeed the predictors reward, transition (rare vs. common) and context (gambling vs. neutral) as fixed effects and subject as random effect.

	Estimate	z-Value	p
Neutral Context			
Reward	0.27350	4.127	3.68e-05***
Transition	0.29908	3.455	0.00055***
Reward*Transition	-0.51326	-4.190	2.79e-05***
Gambling Context			
Reward	0.38419	5.889	3.89e-09***
Transition	0.39439	4.653	3.27e-06***
Reward*Transition	-0.71324	-5.927	3.08e-09***
Full Model			
Reward	0.27361	4.131	3.60e-05***
Transition	0.29924	3.460	0.00054***
Context	-0.12701	-0.426	0.67015
Reward*Transition	-0.51377	-4.199	2.68e-05***
Reward*Context	0.11057	1.190	0.23423
Transition*Context	0.09502	0.785	0.43249
Reward*Transition*Context	-0.19892	-1.160	0.24614

Model free analysis of Stage 1 RTs

S1 RTs were modeled as a function of categorized reward in the previous trial (see previous section for how this was defined) and context as fixed effects and trial and subject as random effects. Previous reward significantly increased RTs (t = -2.431, p = 0.015, see Supplemental Table S2). We also observed a reward * context interaction (see Supplemental Table S2), reflecting Also reaction times were slower, when previously rewarded in the gambling context when contrasted to the neutral context indicating an increased effect of reward on S1 response caution in the gambling context.

Supplemental Table S4. Hierarchical general linear model results of S1 RTs with reward and context as fixed effects and subject as random effect.

S1 RT Model			
	Estimate	t-Value	p
Reward	-0.025	-2.431	0.015*
Context	-0.0005	-0.044	0.97
Reward*Context	0.029	2.050	0.04*

Working memory assessment

Temporal discounting

Preregistered analysis:

We hypothesized a positive relationship of decision noise parameter and working memory z-score. Our data in fact confirms this hypothesis of a positive relationship between softmax beta and working memory (neutral condition: r = 0.42, p = 0.012); gambling condition: r = 0.44 p = p = 0.016).

Exploratory analysis:

There was relationship between discount-rate and working memory (r = -0.03, p = 0.85). One further exploratory analysis revealed an association of working memory and drift-rate coefficient. Here, higher working memory capacity was associated with higher drift rate coefficients (r = 0.42, p = 0.02).

Reinforcement learning

Preregistered analysis:

WM z-score was positively related to MB RL in the neutral condition, but nonsignificant (r = 0.27, p = 0.16). In the gambling context this association reached trend-level significance (r = 0.31, p = 0.10). The shift in MB RL (individual difference score) from neutral to gambling context was again only descriptively associated with WM (r = 0.27, p = 0.16).

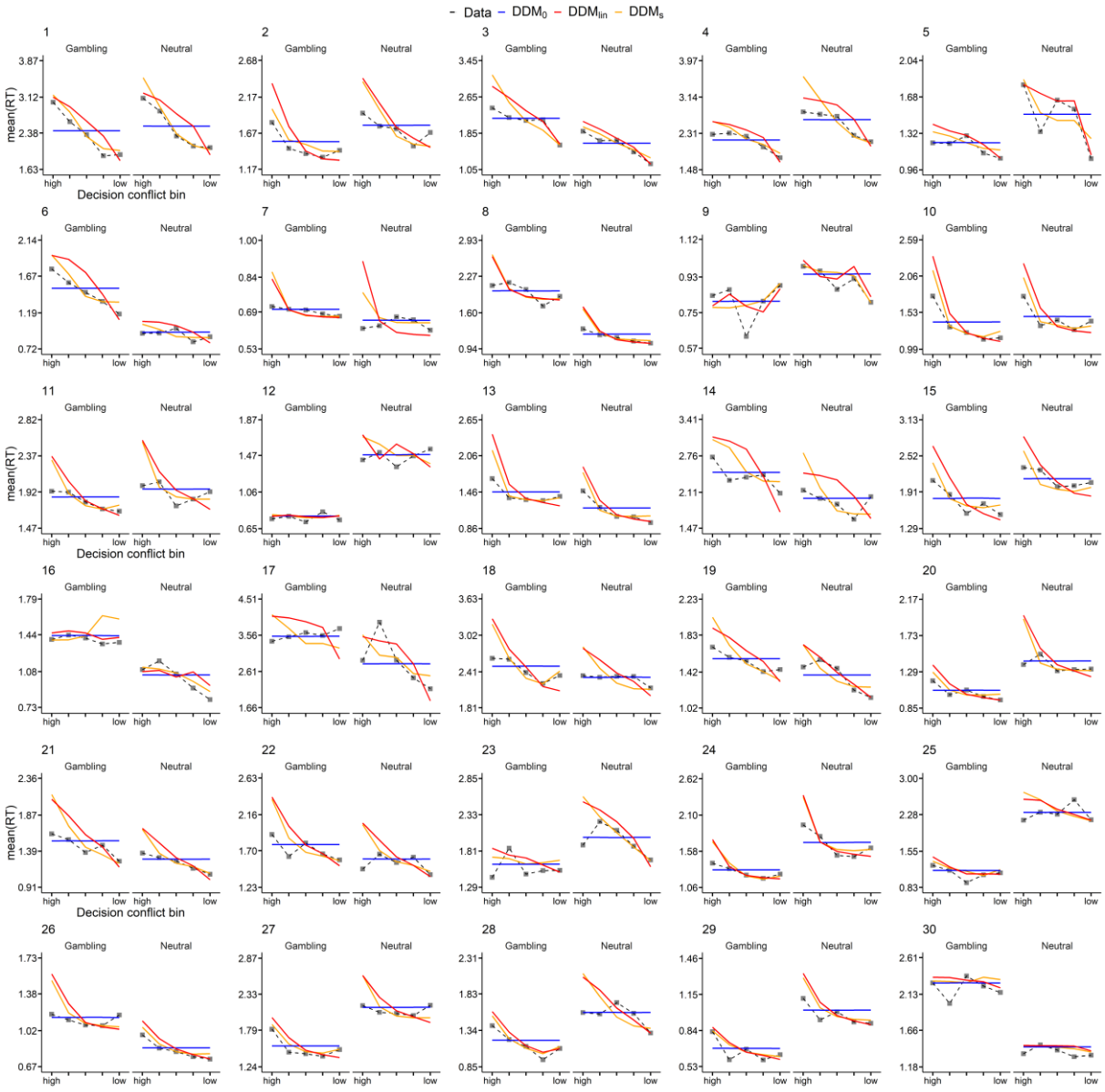
Exploratory analysis:

MF gambling related shifts were unrelated to WM capacity. We further explored the association of WM capacity and S2 learning rates. This analysis revealed that overall WM capacity was positively associated with baseline (neutral context) (r = 0.57, p = 0.001) and gambling context (r = 0.37, p = 0.048) S2 learning rates.

Posterior predictive checks

Temporal discounting

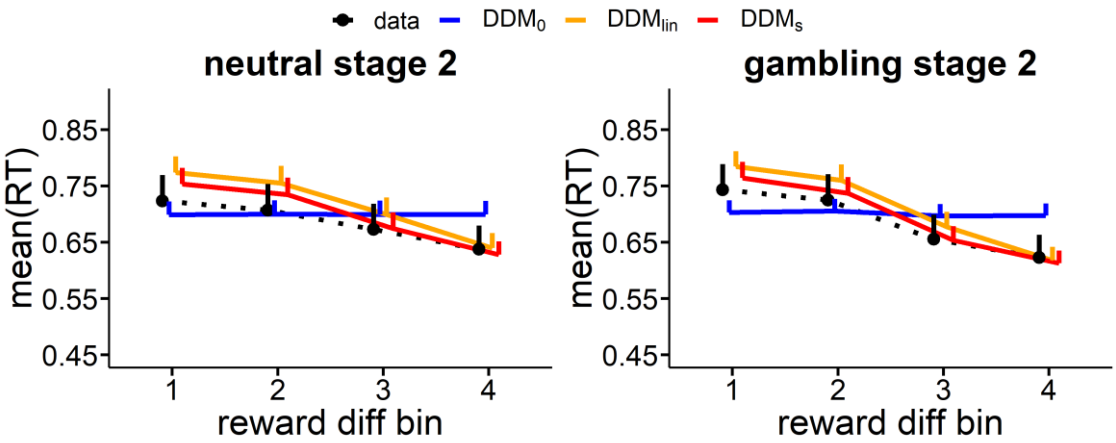
We carried out posterior predictive checks to visualize if our computational analysis captures key aspect in the data, in particular the value-dependency of RTs (Peters & D'Esposito, 2020; Wagner et al., 2020) . For the temporal discounting task, we binned trials per participant into five bins according to the absolute difference in larger-later vs. smaller-sooner value ("decision conflict", computed according to each participant's median posterior log(k) parameter from the DDM_S, and separately for the neutral and gambling context conditions). We then plotted the mean observed RTs as a function of decision conflict per participant and context, as well as the mean RTs across 10.000 data sets simulated from the posterior distributions of the DDM_0, DDM_{lin} and DDM_S (see Figure S3).



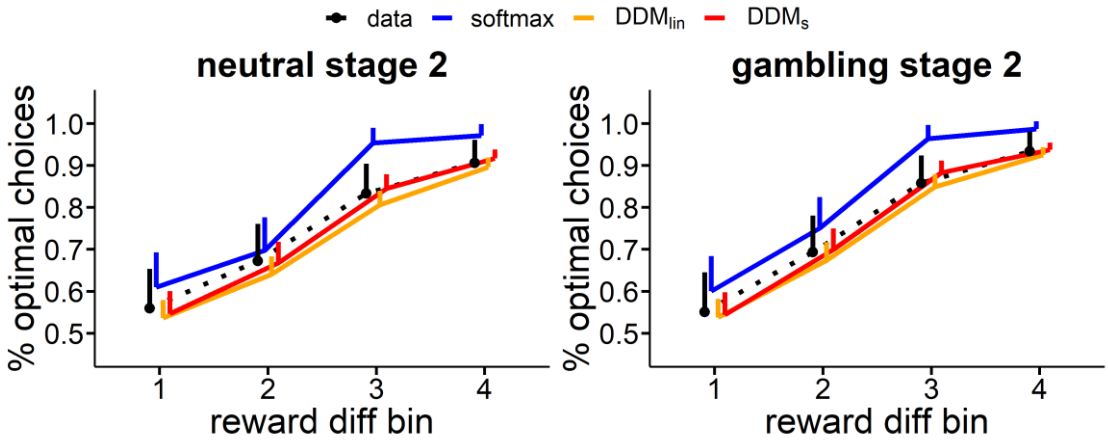
Supplemental Figure S2. Posterior predictive checks for temporal discounting drift diffusion models. For each participant and condition (Gambling vs. Neutral), trials were binned into five equal sized bins according to the absolute difference between subjective LL and SS option values (decision conflict bin). Plotted are mean observed RTs per bin (data) as well model-generated RTs (blue: DDM_0, red: DDM_{lin}, orange: DDM_S) averaged over 10,000 datasets simulated from the respective posterior distributions of the hierarchical models.

2-step task

We conducted posterior predictive checks to evaluate if our different hierarchical models capture both the relationship of RTs and reward differences and the relationship of reward differences and optimal choices. An optimal choice here is defined as a choice for the random walk with highest payout. To this end, we binned all trials into four bins, according to the absolute max(reward) differences in stage 2. For each reward difference bin we then plot the mean observed RTs, as well as the mean simulated RTs across 1000 datasets simulated using our mean parameter estimates for the posterior distributions of the DDM_0, DDM_{lin}, and DDM_S. We further show the mean observed optimal choices (max[reward]) vs. the mean simulated optimal choices given our mean parameter estimates for the posterior distribution of each model. These results are shown in Supplemental Figures S3 and S4. As can be seen, the DDM_s provided the best account of how RTs vary as a function of reward differences. This model outperformed the other models in capturing the relationship of reward differences and optimal choices (Supplemental Figure S4).



Supplemental Figure S3. Group level posterior predictive checks. Trials were binned into four equal sized bins according to the absolute difference in *reward* values given S2 reward walks. Plotted are mean observed RTs per bin (data; dashed line) as well model-generated RTs (blue represents DDM_0; orange represents DDM_{lin}; red represents DDM_S) averaged 1000 simulated datasets simulated from the mean parameter estimates the posterior distribution of each hierarchical model.



Supplemental Figure S4. Group level posterior predictive checks. Trials were binned into four equal sized bins according to the absolute difference in *reward* values given S2 reward walks. Plotted are mean optimal choices in % of all choices per bin (data; dashed line) as well model-generated RTs (blue represents DDM_0; orange represents DDM_{lin}; red represents DDM_S) averaged 1000 simulated datasets simulated from the mean parameter estimates for the posterior distribution of each hierarchical model.

PLOS ONE

RESEARCH ARTICLE

Temporal discounting in adolescents and adults with Tourette syndrome

Canan Beate Schüller[1‡]*, Ben Jonathan Wagner[2‡], Thomas Schüller[1], Juan Carlos Baldermann[1,3], Daniel Huys[1], Julia Kerner auch Koerner[4,5], Eva Niessen[6], Alexander Münchau[7], Valerie Brandt[8], Jan Peters[2‡], Jens Kuhn[1,9‡]

1 Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany, **2** Department of Biology Psychology, University of Cologne, Cologne, Germany, **3** Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany, **4** Educational Psychology, Helmut-Schmidt-University, Hamburg, Germany, **5** Center for Individual Development and Adaptive Education of Children at Risk, Frankfurt am Main, Germany, **6** Department of Individual Differences and Psychological Assessment, University of Cologne, Cologne, Germany, **7** Institute of Systems Motor Science, University of Lübeck, Lübeck, Germany, **8** Center for Innovation in Mental Health, School of Psychology, University of Southampton, Southampton, Hampshire, United Kingdom, **9** Department of Psychiatry, Psychotherapy and Psychosomatic, Johanniter Hospital Oberhausen, EVKLN, Oberhausen, Germany

‡ CBS and BJW share first authorship on this work. JP and JK share senior authorship on this work.
* canan.schueller1@uk-koeln.de

Citation: Schüller CB, Wagner BJ, Schüller T, Baldermann JC, Huys D, Kerner auch Koerner J, et al. (2021) Temporal discounting in adolescents and adults with Tourette syndrome. PLoS ONE 16(6): e0253620. https://doi.org/10.1371/journal.pone.0253620

Editor: Darrell A. Worthy, Texas A&M University, UNITED STATES

Received: February 1, 2021

Accepted: June 8, 2021

Published: June 18, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0253620

Data Availability Statement: We have now made the data publicly available on the platform: center for open science (Schüller C. Temporal discounting in adolescents and adults with Tourette syndrome

Abstract

Tourette syndrome is a neurodevelopmental disorder associated with hyperactivity in dopaminergic networks. Dopaminergic hyperactivity in the basal ganglia has previously been linked to increased sensitivity to positive reinforcement and increases in choice impulsivity. In this study, we examine whether this extends to changes in temporal discounting, where impulsivity is operationalized as an increased preference for smaller-but-sooner over larger-but-later rewards. We assessed intertemporal choice in two studies including nineteen adolescents (age: mean[sd] = 14.21[±2.37], 13 male subjects) and twenty-five adult patients (age: mean[sd] = 29.88 [±9.03]; 19 male subjects) with Tourette syndrome and healthy age- and education matched controls. Computational modeling using exponential and hyperbolic discounting models via hierarchical Bayesian parameter estimation revealed reduced temporal discounting in adolescent patients, and no evidence for differences in adult patients. Results are discussed with respect to neural models of temporal discounting, dopaminergic alterations in Tourette syndrome and the developmental trajectory of temporal discounting. Specifically, adolescents might show attenuated discounting due to improved inhibitory functions that also affect choice impulsivity and/or the developmental trajectory of executive control functions. Future studies would benefit from a longitudinal approach to further elucidate the developmental trajectory of these effects.

Introduction

Tourette syndrome (TS) is a childhood onset neuropsychiatric disorder characterized by motor and phonic tics that wax and wane in their severity with an estimated prevalence of

[Internet]. OSF; 10 Apr 2021. Available: osf.io/nw9pc).

Funding: C.B.S. and T.S. were funded by the Walter and Marga Boll Foundation (210-06-16, URL: https://www.bollstiftung.de/). T.S. and J.C.B. were supported by Deutsche Forschungsgemeinschaft (DFG: 431549029 – SFB 1451; URL: https://www.dfg.de/). J.P. was supported by Deutsche Forschungsgemeinschaft (DFG: PE 1627/5-1; URL: https://www.dfg.de/). A. M. was supported by the DFG (FOR 2698; URL: https://www.dfg.de/). V.B. was supported by the Academy of Medical Sciences (URL: https://acmedsci.ac.uk/). C.B.S., T.S., A.M., V.B, J.K., and J.P. conceived the idea. C.B.S., T.S., E.N., A.M., V. B., J.K.K., and J.P. conceived and designed the experiments; C.B.S., E.N., and V.B. performed the experiments; C.B.S., B.W., and J.P. analyzed the data; J.P. and B.W contributed analysis tools; B.W. performed the modelling. C.B.S. and B.W. wrote the paper. All authors reviewed and approved the final manuscript. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

around 1% [1]. Motor tics are repetitive, sudden movements such as eye blinking or facial muscle contractions and phonic tics are repetitive sounds such as throat clearing or verbal utterances [1, 2]. TS onset occurs predominantly in early childhood with a peak of symptom severity between the age of 10 and 12 years. Thereafter, tics improve in around 80% of children until the end of adolescence [3].

Both clinical and neuroscientific research have highlighted possible developmental dysfunctions in the cortico-striatal-thalamo-cortical (CSTC) loops [4–6] especially with respect to dopamine (DA) that strongly modulates these circuits [7, 8]. The striatum, a main gateway in these loops [9] plays a key role in selectively amplifying converging sensory input to enable situation specific behavioral adaptations, such as the adequate control of voluntary movement [7]. Predictions (i.e. expectations) of reward, as well as the gating of specific motor responses, are under dopaminergic modulation. Theories about the developmental underpinnings of TS in terms of DA function range from striatal DA receptor super-sensitivity [10] over tonic-phasic or presynaptic DA dysfunction [11, 12] to DA hyper-innervation [11, 13]. The DA hyper innervation hypothesis unifies previous findings under a promising framework.

To date, several studies have investigated motor impulsivity in patients with TS with reference to DA´s role in reward and motor control [14, 15]. However, fewer studies have explored alterations in value-based decision-making in TS. This question is of particular interest because motor and choice impulsivity might at least in part be supported by common neural systems. First, DA in fronto-striatal circuits plays a role in both motor control [16, 17] and choice impulsivity [18–20]. Second, some studies have suggested that lateral prefrontal cortical (LPFC) regions might support impulse control functions, both in motor and non-motor domains [21–24]. Two studies [25, 26] examined impairments in value-based decision-making in TS in the context of reinforcement learning tasks. Palminteri and Pessiglione observed impaired learning from negative feedback in TS [25], which is consistent with the idea of a hyperdopaminergic state. Kéri and colleagues observed impaired probabilistic classification learning, especially in children with severe tics [26]. However, the degree to which choice impulsivity is impaired in TS remains unclear. To date, only one study examined temporal discounting in patients with TS via the Kirby Monetary Choice Questionnaire [27] and observed slightly increased discounting compared to healthy controls.

Another way to reliably assess this process is via intemporal choice tasks [28, 29]. Temporal discounting describes a general preference for smaller sooner (SS) over larger, but later rewards (LL) [30]. A relative preference for SS rewards (steep discounting of value over time) is associated with a range of problematic behaviors, including substance use disorders and overweight/obesity [31], the tendency to procrastinate investing in a pension [32], and to procrastinate saving up for future investments [33]. The rate of temporal discounting is subject to complex modulation by individual and contextual variables [34, 35], where striatal DA networks and prefrontal top down modulation seem to play crucial roles. However, the precise relationship between dopaminergic states and impulsive choice is complex and might be a function of age with a proposed u-shape association [36]. On the one hand, pharmacological elevation of DA levels decreases discounting [20, 37–39]. On the other hand, hyperdopaminergic states, e.g. due to administration of the dopamine precursor L-DOPA, are in some studies associated with increased discounting [18], and patients with Parkinson's disease can exhibit increased impulsive behavior following DA replacement therapy [19]. To sum up, DA modulation likely contributes to the modulation of intertemporal choice via its action on different fronto-striatal loops, but scientific evidence suggests that there is no clear and simple linear relationship between DA levels and choice impulsivity.

The picture is clearer with regard to top-down inhibitory mechanisms. The LPFC is assumed to modify choice impulsivity [40–42], that is, inhibition of the selection of tempting

SS choices in this model depends on prefrontal inhibitory regulation of subcortical or ventromedial prefrontal value representations. Changes in structural and functional connectivity within this network are linked to the development of self-control (in this study the term 'self-control' generally refers to far-sighted behavior in value based decision making) from adolescence to early adulthood [43, 44]. Furthermore, inhibition and top-down control play a central role in motor impulsivity and are believed to modulate TS pathophysiology, e.g. in the context of suppressing urges and tics [14].

Studies have shown that motor and cognitive impulsive actions might require different forms of self-control and these can be differentiated [45]. To sum up, there is extensive evidence for regional overlap between inhibitory mechanisms in terms of motor impulsivity, choice impulsivity, and other forms of impulsivity, such as emotion regulation [22–24]. Training in one domain might affect performance in other domains [46]. Regarding choice and motor impulsivity, the dorsal striatum might be a key region of interest where top down inhibitory processes (originating in the PFC) modulate the execution or the re-evaluation of choice outcomes [47]. These anatomical regions and attributed functions might be affected in TS pathophysiology [48]. Even though temporal discounting has been proposed as an transdiagnostic trait [49] with valuable diagnostic potential [50] it is still an open question whether patients with TS show aberrations in the domain of intertemporal choice. In the present study, we compared adolescents (Study 1, Hamburg) and adults (Study 2, Cologne) with TS to controls, using two modified temporal discounting tasks to broaden the understanding of value based decisions in TS on one operational measure of choice impulsivity [32, 33].

Materials and methods

Ethics

The ethics committee of the University Hospital Hamburg approved the first study. Adolescent patients with TS provided written assent and their parents provided written consent (PV4439). Patients with TS were recruited in the University Hospital of Hamburg, whereas controls were recruited by advertisement. The second study was carried out in accordance with institutional guidelines and was approved from the ethics committee of the University of Cologne (protocol ID: DRKS00011748). All participants provided written consent. Patients were recruited at the University Hospital of Cologne whereas controls were recruited by advertisement.

Study 1 specific methods

Participants. We included nineteen adolescents with TS (age: mean[sd] = 14.21[±2.37], 13 male subjects, range: 10–17) and nineteen age, education and gender-matched controls (age: mean[sd] = 14.21[±2.53], 15 male subjects, range: 10–18). Adolescents with TS were recruited from a specialist Tourette syndrome clinic in Hamburg. All patients had been diagnosed with Tourette syndrome, some had been treated for their tics. Healthy controls were partly recruited from a pool of healthy participants who had participated in a previous study, partly via public advertisement. All participants underwent a clinical assessment and performed a modified delay discounting paradigm. Two adolescents with TS were taking antidopaminergic drugs (Tiaprid) as prescription medication.

Clinical assessment. Adolescents were assessed with the "Yale Global Tic Severity Scale" (YGTSS) [51], the "Premonitory Urge for Tic Disorders Scale" (PUTS), a self-report scale to identify premonitory urges [52], and the "Children's Yale-Brown Obsessive Compulsive Scale" (CY-BOCS), a semi structured interview to evaluate OCD severity. CY-BOCS data were available from all adolescents with TS and 13 controls; in total, three adolescents with TS had a

score above 12, the cut-off for clinically relevant OCD symptoms [53]. The "Parent-rated and Self-rated Questionnaires for Attention Deficit Hyperactivity Disorder" (German: "Fremdbeurteilungsbogen /Selbstbeurteilungsbogen für Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen") FBB-ADHD and SBB-ADHD are diagnostic instruments to identify ADHD [54]. FBB-ADHD data was available for all adolescents with TS and 16 controls. SBB-ADHS data was available for 18 adolescents with TS and 17 controls. All adolescents also filled out a questionnaire on demographic measurements (age, gender, medication).

Task. Participants performed a experiential delay discounting task based on prior procedure [55] where they chose between varying smaller sooner (SS ϵ [0, 1, 2, 3 or 4 cents]) or larger but later (LL [5 cents]) rewards. LL options were available after a specific waiting period of 10, 20, 30, 40 or 60 seconds. Each SS-option was paired twice with each LL-option resulting in 50 trials per participant. A progress bar indicated the number trials past. Position of the LL option was counterbalanced to the left or right side of the screen. LL waiting-time was visualized by the number of horizontal lines (e.g. 2 horizontal lines = 20s waiting period). Following choice rewards were transferred into a virtual piggy bank either immediately (if SS was chosen) or after the appointed waiting period (if LL was chosen). Depending on choices, participants could gain between 0 € and 2.50 €. Following this time spent with task, i.e. delay to reward delivery was related to the proportion of SS choices. (see **S1 Fig**).

Study 2 specific methods

Participants. We recruited twenty-five patients with diagnosed TS according to DSM-5 criteria [56] from the psychiatric outpatient clinic of the University Hospital Cologne (age: mean[sd] = 29.88 [±9.03]; 19 male subjects, range: 19–53) and 25 age, education and gender-matched controls (age: mean[sd] = 29.40 [±9.28]; 17 male subjects, range:19–49) through public advertisement. All participants underwent a clinical assessment, performed a temporal discounting paradigm, including a pretest based on prior procedures [57, 58]. Nine patients were taking medication or cannabinoids. Five patients were treated with antidopaminergic drugs (Aripiprazole, risperidone, tiapride) as a monotherapy, one patient with an anticonvulsant (Valproate), one patient was taking a noradrenergic and specific serotonergic antidepressant (Mirtazapine), and one patient was medicated with a combination of two antidopaminergic drugs (Aripiprazole, risperidone) and a selective serotonin reuptake inhibitor (Citalopram). One patient regularly smoked medical cannabis.

Clinical assessment. All participants filled out the Obsessive Compulsive Inventory-Revised (OCI-R) [59] and the Beck Depression Inventory (BDI) [60]. The Wender Utah Rating Scale was used to assess ADHD symptoms [61]. Furthermore, they filled out a short intelligence test (Leitprüfsystem-3 (LPS 3)) [62], followed by a demographic questionnaire with information on age, gender, handedness, years of education and current drug or alcohol use. Further, patients with TS completed an assessment with the YGTTS [51], and the PUTS [52]. All questionnaires were in German.

Task. Prior to the first testing session, participants completed a short adaptive pretest to estimate the individual discount- rate (k). This discount rate was used to construct a set of 140 participant-specific trials using MATLAB (version 8.4.0. Natick, Massachusetts: The MathWorks Inc). The task consisted of choices between an immediate smaller-sooner reward of 20 € and participant specific larger-but-later (LL) rewards delivered after some delay (1, 2, 7, 14, 30, 90 or 180 days). In 70 trials, LL amounts were uniformly spaced between 20.5 € and 80 €, whereas in the remaining 70 trials LL amounts were uniformly spaced around each estimated indifference point per delay (based on the pre-test discount rate). If indifference points were larger than 80 €, only uniformly-spaced LL amounts were used. Trials were presented in a

pseudorandomized order. Participants were informed that after task completion, one trial would be randomly selected and paid immediately in cash (smaller-sooner choice) or via a timed bank transfer (larger-but-later choice).

Statistical analyses (both studies)

Model free analysis. Using model agnostic approaches can avoid possible caveats associated with model-based analysis, e.g., problems with parameter estimation or the choice for a theoretical framework. Due to task structure in study 1 (adolescents) we used the percentage of LL in contrast to SS choices as a model agnostic quantification of choice behavior. For comparison, we used a two-sided parametric test on the arc-sin-transformed values of SS vs. LL choices.

In study 2 (adults) we computed the area under the empirical discounting curve (*AUC*) (Note, due to the low number of varying rewards [only four different SS rewards], computing the area under the points of indifference would decrease variability and in consequence information when applied to the data in study 1). In detail, the *AUC* corresponds to the area under the connected data points that describe the decrease of the subjective value (y-axis) over time (x-axis). Each specific delay was expressed as a proportion of the maximum delay and plotted against the normalized subjective (discounted) value. We then computed the area of the resulting trapezoids using Eq 1.

$$\frac{x_2 - x_1}{\left(\frac{(y_1+y_2)}{2}\right)} \quad \text{(Eq 1)}$$

Smaller *AUC*-values indicate more discounting (more impulsive choices) and higher *AUC*-values indicate less discounting.

Computational modeling. Based on prior analysis and basic research in the field of temporal discounting we a-priori assumed a hyperbolic model [63, 64] to model the decrease in subjective value over time. Bayesian estimation methods have the advantage of estimating the entire posterior distribution of parameter values given the data. Furthermore, hierarchical Bayesian parameter estimation benefits from the fact that the entire data set is taken into account via the hierarchical structure of the model. Parameters from each participant thus mutually inform and constrain each other (partial pooling), such that meaningful estimates can be derived even with limited data per subject (for details on Bayesian group comparison see [65, 66]; or for an overview see [67]. Due to the different time-scales of both intertemporal choice tasks in adolescents and adults we decided to compare hyperbolic (Eq 2) and exponential discounting (Eq 3) models. Both models assume that the LL reward, delivered after a specific delay (D), is devaluated via a subject specific discount rate (k) that weights the influence of time on subjective value (SV). A lower k-parameter reflects a lower weight on delay (reduced discounting) whereas a higher k-parameter reflects steeper discounting. Both models differ in the way they model this weight. In hyperbolic discounting the near future is discounted more heavily than distant events. In exponential discounting the discount rate is constant.

$$SV = \frac{LL}{(1 + kD)} \quad \text{(Eq 2)}$$

$$SV = LL * \exp(-kD) \quad \text{(Eq 3)}$$

After devaluating the delayed option a sigmoid function (softmax choice rule; Eq 4) maps the comparison of both the devaluated LL and SS option to choice probability on a trial by trial basis. Here a free β inverse temperature parameter scales the influence of value differences on

choice. A high β value implies that participants decide purely on value differences whereas lower values indicates higher choice stochasticity. For limit of $\beta = 0$ choices are completely random.

$$P(LL) = \frac{\exp(\beta * SV(LL))}{\exp(\beta * SV(SS)) + \exp(\beta * SV(LL))} \quad \text{(Eq 4)}$$

Models were fit using a hierarchical Bayesian framework to estimate parameter distributions via Markov Chain Monte Carlo (MCMC) sampling with JAGS [68]. Single subject parameters were drawn from group-level normal distributions, with mean and variance hyper-parameters that were themselves estimated from the data. Model convergence was assessed via the $\hat{R}$ -statistic (Gelman-Rubinstein convergence diagnostic) where values < 1.01. (two chains) were considered acceptable. For information on prior specification see S1 Table.

Analyses of group differences. Group comparisons were conducted by examining the differences in posterior distributions per parameter of interest (discount-rate *k* and softmax β). For group comparisons, we report Bayes factors (directional Bayes Factor (dBF)) for directional effects for the hyperparameter difference distributions of patients with TS and controls. BFs were estimated via kernel density estimation using R (4.03) via the RStudio (1.3.1) interface. These are computed as the ratio of the integral of the posterior difference distribution from 0 to ∞ versus the integral from 0 to–∞. Using common criteria [69], we considered BFs between 1 and 3 as anecdotal evidence, BFs > 3 as moderate evidence, and BFs > 10 as strong evidence. BFs > 30 and > 100 were considered as very strong and extreme evidence, respectively, the inverse of these reflect evidence in favor of the opposite hypothesis.

Results

Study 1

Demographic characteristics and clinical assessment. Demographic and clinical characteristics between adolescents with TS and controls are shown in Table 1. For demographic,

Table 1. Demographic, clinical and neuropsychological characteristics of adolescents with TS and healthy controls.

	Adolescents with TS (*n* = 19)		Controls (*n* = 19)		*T/U/ X^2*	*p*
	Mean	SD	Mean	SD		
Age (Years) [a]	14.21	2.37	14.21	2.53	0.000	1.000
Male/Female [c]	13/6	-	78.9	-	0.543	0.467
Right-handed [c]	14/19	-	84.2	-	1.276	0.435
Current medication	2/19	-	-	-	-	-
YGTSS impairment	16.00	8.00	-	-	-	-
YGTSS	23.37	12.38	-	-	-	-
PUTS	19.53	5.61	-	-	-	-
FBB-ADHD[b]	0.38	0.26	0.82	0.48	-3.226	0.093
SBB-ADHD[a]	0.39	0.22	0.68	0.39	88.0	0.497
CY-BOCS [b]	6.84	6.31	0.08	0.277	21.50	<0.001

ADHD, attention deficit hyperactivity disorder; CY-BOCS, Children's Yale-Brown Obsessive-Compulsive Scale; (FBB)-ADHD/(SBB)-ADHD, Fremdbeurteilungsbogen/Selbstbeurteilungsbogen für Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen; PUTS, Premonitory Urge for Tics Scale; TS, Tourette syndrome; YGTSS, Yale Global Tic Severity Scale.

a. T-test was used because data was normally distributed.

b. Mann Whitney U test was used because data was not normally distributed.

c. X^2 square test.

https://doi.org/10.1371/journal.pone.0253620.t001

Table 2. Model comparison of two variants of intertemporal choice.

	Adolescents			Adults		
	Patients with TS	**Controls**	**Full model**	**Patients with TS**	**Controls**	**Full model**
Hyberbolic	791.5	878.8	1668.83	2538.4*	2156.4*	4701.7*
Exponential	686.5*	806.2*	1535.92*	2634.8	2297.8	4926.9

https://doi.org/10.1371/journal.pone.0253620.t002

clinical and neuropsychological characteristics of adolescents with TS and controls adjusted for multiple comparison see S2 Table.

Model free analysis. Controls chose the LL option in 48.3% of all cases whereas patients with TS chose that option around 10% more often in 58.4% of all cases (see S2 Fig). Before using a parametric-t-test, we applied an arcsin-transformation on all mean choice proportions per participant. The groups did not differ significantly in the frequency of LL choices ($t_{(35.83)}$ = 1.0646; p = 0.29).

Computational modeling. Model comparison via DIC [70] revealed a better fit of the exponential model (see Table 2). This holds when applying a full model including all participants from both groups or when modeling both groups separately (see Table 2). We next examined overall group differences for the discount-rate k (Fig 1A). Analyzing the posterior group difference distribution (Fig 1B) revealed that 93% of the posterior distribution of controls is below the distribution of patients with TS. We then computed a dBF(see methods section) which quantifies the relative evidence for increases vs. decreases in patients compared to controls. This yielded a dBF of 12.52, i.e. given the data and model, an increase in discounting on the group level in controls was 12.52 times more likely than a decrease. The corresponding analysis of choice stochasticity is provided in S3 Fig.

Model comparison was based on the Deviance Information Criterion (DIC) [66] where lower values indicate a better model fit. The adolescent data were better accounted by a model

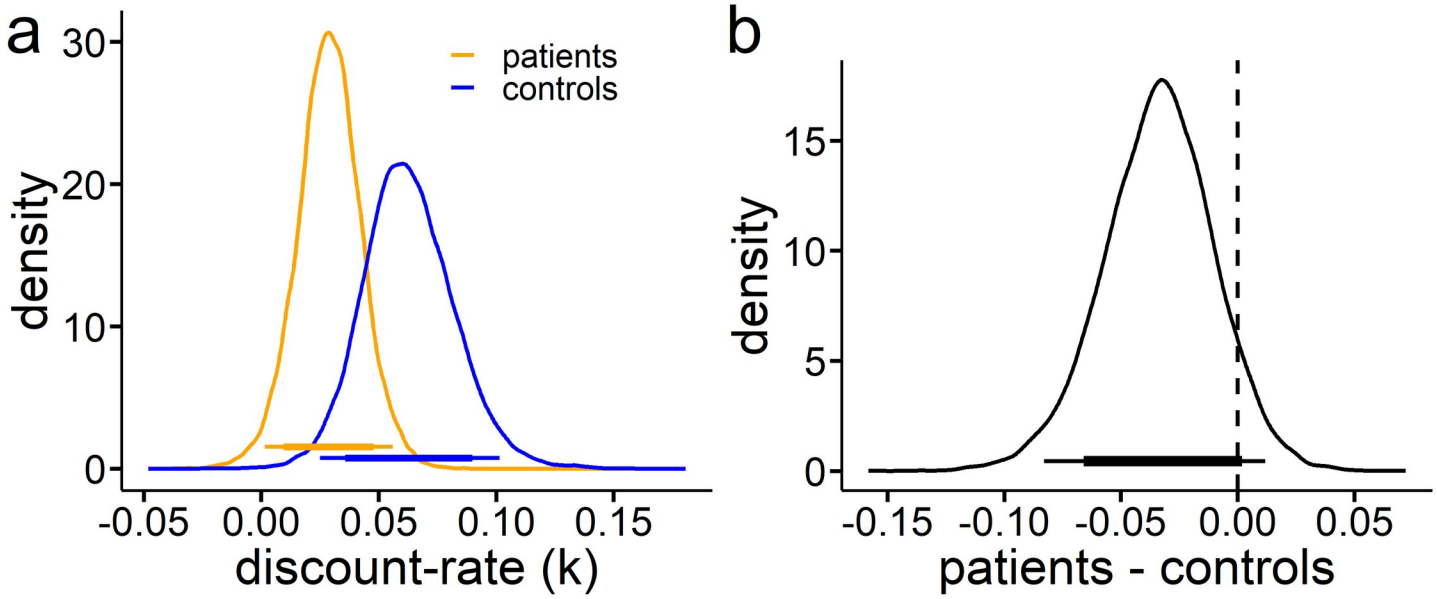


Fig 1. a, Group level hyperparameter distributions of the discount-rate parameter k revealed that discounting was lower in adolescents with TS (orange) when compared to controls (blue). b, Difference distribution of controls—adolescence with TS. Bayes factor for directional effects (dBF) indicated that a decrease in discount-rate (k) in patients was 12.52 times more likely than an increase. Thin and thick colored (a) and black (b) bars indicate the 95% and 85% highest density intervals respectively. TS, Tourette syndrome.

https://doi.org/10.1371/journal.pone.0253620.g001

with an exponential discount function and the adult data were generally better accounted for by a temporal discounting model with hyperbolic discounting whereas.

Study 2

Demographic characteristics and clinical assessment. Demographic and clinical characteristics of adult patients with TS and controls are shown in Table 3. Controls did not score in a clinically relevant range. Neither patients nor controls reported clinically relevant drug or alcohol abuse. We further conducted an analysis of correlations of discount-rate, age, compulsivity and symptom severity (S3 Table).

Model free analysis. Applying a parametric t-test on the integral of the area under the empirical discounting curve revealed no significant differences between patients with TS (mean[AUC] = 0.459) and controls (mean[AUC] = 0.511) ($t_{(46.1)}$ = -0.8791; df = 46.1; p = 0.38), see S4 Fig.

Computational modeling. Comparing hyperbolic and exponential discount functions based on the DIC [70] revealed a better fit of the hyperbolic model. This holds when applying a full model including all participants from both groups or when modeling both groups separately (see Table 2). In line with our model-agnostic approach, we did not find evidence for group differences when analyzing the posterior difference distribution of the discount-rate (k). Results are plotted in Fig 2. There was no evidence for consistent group differences (dBF = 0.38). The analysis was repeated excluding six patients with TS, that were taking antidopaminergic drugs at the time of the study. The exclusion of these patients only had a marginal effect and the result pattern did not change. For analysis of choice stochasticity see S5 Fig.

Table 3. Demographic, clinical and neuropsychological characteristics of patients with TS and healthy controls.

	Adult patients with TS (n = 25)		Controls (n = 25)		$T/U/X^2$	p
	Mean	SD	Mean	SD		
Age (Years)[a]	29.88	9.03	29.40	9.28	0.185	0.854
Male/Female [c]	19/6	-	68.00	-	0.397	0.529
Right-handed	22/25	-	88.00	-	0.000	1.000
Current medication	9/25	-	-	-	-	-
Years of education [b]	11.68	1.25	11.90	1.22	250.00	0.197
Tourette Onset	8.76	5.13	-	-	-	-
YGTSS motor	15.84	5.72	-	-	-	-
YGTSS verbal	12.32	6.36	-	-	-	-
YGTSS impairment	26.80	11.08	-	-	-	-
YGTSS	54.96	20.78	-	-	-	-
PUTS	30.02	4.22	-	-	-	-
BDI [b]	11.68	9.34	5.28	5.19	165.50	0.004
WURS-k [a]	26.12	11.60	16.04	9.55	3.36	0.002
OCI-R [b]	20.30	12.06	10.92	7.58	149.50	0.002
LPS-3 [b]	55.80	8.25	58.60	8.48	249.50	0.213

BDI, Becks depression inventory; LPS-3, Leistungsprüfsystem; OCI-R, Obsessive-Compulsive Inventory-Revised; PUTS, premonitory urge tic for scale; TS, Tourette syndrome; WURS-k, Wender-Utah-Rating-Scale; YGTSS, Yale Global Tic Severity Scale.

a. T-test was used because data was normally distributed.

b. Mann Whitney U test was used because data was not normally distributed.

c. X^2 square test.

https://doi.org/10.1371/journal.pone.0253620.t003

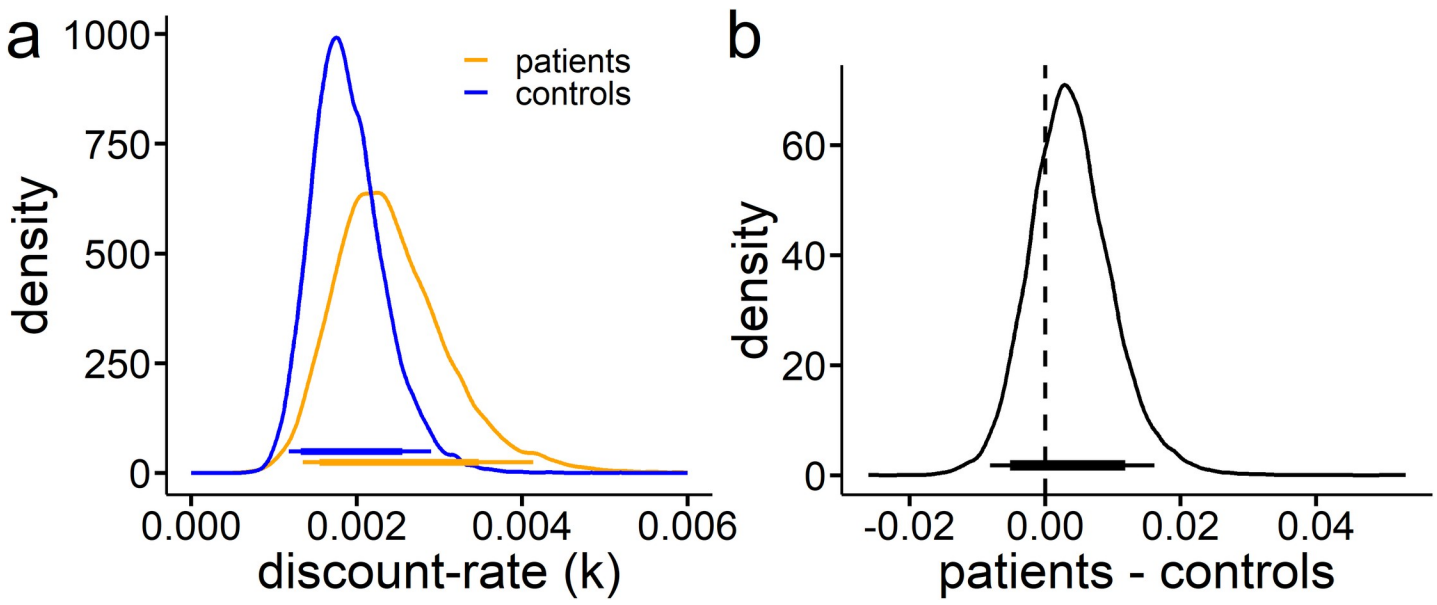


Fig 2. a, Group level hyperparameter distributions of discount-rate *k* for patients with TS (orange) and controls (blue); b Difference distribution of patients with TS—controls. The black bars indicate the 95% and 85% highest density interval respectively. Bayes factors for directional effects (dBF) of 0.36 patients > controls indicate no substantial difference between patients and controls. Thin and thick colored (a) and black (b) bars indicate the 95% and 85% highest density intervals respectively. TS, Tourette syndrome.

https://doi.org/10.1371/journal.pone.0253620.g002

Conclusions

The present study assessed temporal discounting in adolescent and adult patients with TS and matched healthy controls. Our data suggest reduced discounting (experiential task) in adolescent TS patients where in decrease in discounting was 12.52 times more likely than an increase when contrasted to controls. We did not find any difference in intertemporal choice in adults (hypothetical intertemporal choice task). TS is a complex neuropsychiatric disorder associated with developmental disturbances in dopaminergic transmission which possibly result in failure to control motor output [1, 2, 14, 15, 71]. These dopaminergic anomalies may either cause, enable or enhance tics via inadequate gating of information through the striatum [7]. Some studies point towards reductions in temporal discounting due to pharmacological elevation of DA levels, whereas others point to an increase [18]. Generally, the human literature on dopaminergic contributions to impulsivity is characterized by substantial heterogeneity [72]. A further complicating factor is that dopaminergic effects might be non-linear [73], as summarized in the inverted U-model of DA functioning [74]. However, acute dopaminergic modulation by pharmacological agents and long-term abnormal dopaminergic states such as in TS may effect behavior differently. In line with this distinction, our results suggest that the putative chronic hyperdopaminergic state of TS does not give rise to substantial changes in temporal discounting in adults.

However, we did find evidence for decrease in temporal discounting in adolescents with TS when compared to healthy controls, i.e. adolescents with TS preferred LL rewards.

Our analysis revealed that a decrease in temporal discounting in adolescents with TS was about 12 times more likely than an increase (dBF = 12.52). Adolescents typically show higher discount rates than adults [75]. This is thought to be attributable to functional and structural fronto-subcortical connectivity that undergoes maturation until early adulthood [15, 43, 44]. Adolescents with TS are constantly faced by tics and the need to control their motor output.

Even though these tics might emerge from complex neurophysiological interactions, i.e. hyperactive DA modulated striatal gating and reduced inhibition of GABAergic interneurons [76, 77], one could speculate that the ability to inhibit tics might foster the ability to inhibit other impulses, thereby strengthening cognitive control more generally [46]. However our results conflict with a recent study by Vicario and colleagues [27] who report increased discounting in adolescent patients with TS. We note that the task for the adolescent sample in our study differed distinctively, not only from the task of the adult sample but also from the Monetary Choice Questionnaire used by Vicario and colleagues [27]. Importantly, our task included a payout dependent on actual choice behavior (experiential task). Differences in the reported findings on impulsive choice of Vicario and colleagues and our findings might therefore be reflected in differences in task demands. In theory three complementary systems are thought to orchestrate intertemporal decisions: the valuation network, regions associated with cognitive control [40, 41], and a network associated with future prospection [29]. We therefore further propose that the weights between brain circuits involved in intertemporal choices might differ. That is the networks involved might depend on the temporal horizon of the task i.e. the need for future prospection might be less pronounced in the experiential task. However future studies are needed to clarify these issues.

The question then arises why such an effect would not likewise translate into greater self-control during temporal discounting in the adult TS patients. One possibility is that such a "training" account merely affects the developmental trajectory of self-control, such that adolescents with TS reach adult levels of self-control earlier than their healthy peers. Testing such a model would require longitudinal studies.

Additional clinical differences between adolescent and adult TS patients further complicate the interpretation of the differential effects in the two age groups. Adolescents and adults with TS exhibit different tic-phenomenology, for instance adolescents exhibit higher variability and/or fluctuations in tics. In consequence adolescents who successfully control their tics have a greater likelihood of eventual remission, likely due to better executive control capabilities [78]. In contrast, patients who still exhibit TS in adulthood exhibit attenuated inhibitory control [14]. In both samples, the discount rate (k) was not significantly correlated (corrected for multiple comparisons) with ADHD, OCD comorbid symptomatology or the YGTSS (see **S1** and **S2 Tables**). Interestingly, the data in adolescent patients with TS was best fit by an exponential function, while the data in adult patients with TS was best accounted for by a model with hyperbolic discount function, which is in line with most of the literature on intertemporal choice [63]. First, though speculative the function of temporal discounting, processed in CSTC-loops, might generally be sensitive to the time scale (seconds/minute in adolescents vs. days to weeks in adults) of the task (see discussion of task differences above). Second, there might be technical reasons for this finding so the differences in the relative model fit between tasks could be due to differences in the option sets.

The present study has several limitations. First, adolescents and adults performed different temporal discounting tasks with different reward magnitudes (0–4 cents vs. 20–80 €) on a different timescale. Reward magnitudes in the range of cents vs. tens of Euros may entail different valuation and/or control processes [79, 80]. This precludes direct comparisons in k between age groups. Importantly, both tasks experiential and hypothetical differ in what is known regarding their internal and external validity. While the hypothetical intertemporal choice task was proposed to constitute a transdiagnostic trait [81] less is known about the experiential task. Nonetheless, we note that the experiential task in study 1 is comparable with tasks like those used in the Marshmallow experiments by Mischel and Ebbesen [82] or other experiential adaptions [83, 84]. These experiential tasks have also shown some predictive value [85, 86] and successful treatment interventions in populations that are known for steep discounting [87].

Some studies do report correlation of experiential and hypothetical tasks (e.g. [88]). However, these findings are not always present [89, 90] and therefore represent a limitation of the current study.

Second, we draw theoretical conclusions from reward impulsivity to motor inhibition in patients with TS, even though motor inhibition was nor directly tested in the present studies. Further studies should further examine the developmental trajectories of both functions. Third, although only two adolescents with TS took medication, about a quarter of the adult patients ($n = 6$) were on antidopaminergic medication. An integrative review showed that most TS medication (i.e. D_2 antagonists) reduce phasic DA, tonic DA or both [71] such that processing in fronto-striatal circuits was likely affected by the medication. However, a control analysis, excluding participants on antidopaminergic medication yielded the same pattern of results. Fourth, the samples may not be representative of the true TS population. Generalizability is limited due to the respective age ranges, the exclusion of patients with severe comorbidities and the fact that all patients were seeking treatment in a specialized outpatient clinic. Fifth, another limitation is the relatively small sample size of both studies. This is especially relevant for the interpretation of study 2, were no significant between-group differences were observed. Importantly, the lack of difference should be interpreted carefully with further studies needed to verify this finding.

The present study assessed temporal discounting in adolescent and adult patients with TS and matched healthy controls. Our data suggest reduced discounting (via an experiential task) in adolescent TS patients. We speculate that this might be due to improved inhibitory functions that affect choice impulsivity and/or the developmental trajectory of executive control functions. Interestingly, adult patients with TS exhibited levels of discounting similar to controls. This might be due to higher disease severity in adult patients with TS (e.g., patients who acquired successful tic inhibition during adolescence might have gone into remission). Future studies would benefit from adopting a consistent longitudinal approach to further elucidate the developmental trajectory of neural correlates i.e. dopaminergic states and intertemporal preferences and further from directly examining effects of dopaminergic medication on these processes in TS.

Supporting information

S1 Fig. Example for two trials in the temporal discounting task adapted for children and adolescents. The blue circle depicts the LL reward (in cents) that participants will receive if they wait. How long they have to wait is indicated by blue lines, i.e. one blue line = 10s wait, six blue lines = 60s wait. The red circle indicates how much the participant will receive if they move on to the next trial immediately (0–4 cents). Participants received feedback about the amount earned after every trial (piggy bank). The green bar below the two circles indicates how many trials the participant has already finished. LL, larger but later.
(TIF)

S2 Fig. Percentage of larger, but later choices in adolescents with TS and controls. LL, larger but later.
(TIF)

S3 Fig. A: Softmax β in adolescent patients with TS vs. controls. Group level hyperparameter distributions of the inverse temperature parameter softmax β revealed no group differences between patients (orange) and controls (blue). **B: Difference distribution of controls—patients with TS.** Thin and thick colored (a) and black (b) bars indicate the 95% and 85%. TS, Tourette syndrome.
(TIF)

S4 Fig. Subject specific comparison of the integral under the empirical area under the curve in adults with TS and controls. TS, Tourette syndrome.
(TIF)

S5 Fig. A: Softmax β in adult patients with TS vs. controls. Group level hyperparameter distributions of the inverse temperature parameter softmax β revealed no group differences in the mean of the posterior or a shift in either direction between patients (orange) and controls (blue). However, variance was increased in controls indicating higher interindividual differences in decision noise. **B: Difference distribution of controls—patients with TS.** Thin and thick colored (a) and black (b) bars indicate the 95% and 85% highest density intervals respectively. TS, Tourette syndrome.
(TIF)

S1 Table. Prior specifications for group and subject level parameters. Discount-rate (k) parameters were estimated in logarithmic space due to parameter stability. Softmax β values were estimated in standard-normal space for the same reason.
(DOCX)

S2 Table. Correlation analysis of model parameters and subscale of the SBB-Questionnaire adjusted for multiple comparison. We report our exploratory analysis on discount-rate and questionnaire data. Scores are spearman correlation coefficients (p-value) not corrected for multiple comparisons. TS, Tourette syndrome.
(DOCX)

S3 Table. Correlation analysis of model parameters and questionnaire data in adult patients with TS and controls adjusted for multiple comparison. We report our exploratory analysis on discount-rate and questionnaire data. Scores are spearman correlation coefficients (p-value) not corrected for multiple comparisons. TS, Tourette syndrome; BDI, Becks depression inventory; OCI-R, Obsessive-Compulsive Inventory-Revised; TS, Tourette syndrome; WURS-k, Wender-Utah-Rating-Scale.
(DOCX)

Author Contributions

Conceptualization: Canan Beate Schüller, Thomas Schüller, Julia Kerner auch Koerner, Alexander Münchau, Valerie Brandt, Jan Peters, Jens Kuhn.

Data curation: Canan Beate Schüller, Ben Jonathan Wagner, Julia Kerner auch Koerner, Eva Niessen.

Formal analysis: Canan Beate Schüller, Ben Jonathan Wagner, Jan Peters.

Funding acquisition: Canan Beate Schüller, Alexander Münchau, Valerie Brandt, Jens Kuhn.

Investigation: Canan Beate Schüller, Ben Jonathan Wagner, Valerie Brandt.

Methodology: Canan Beate Schüller, Ben Jonathan Wagner, Thomas Schüller, Valerie Brandt, Jan Peters, Jens Kuhn.

Project administration: Canan Beate Schüller, Eva Niessen, Valerie Brandt, Jens Kuhn.

Software: Ben Jonathan Wagner, Jan Peters.

Supervision: Thomas Schüller, Juan Carlos Baldermann, Daniel Huys, Alexander Münchau, Valerie Brandt, Jan Peters, Jens Kuhn.

Writing – original draft: Canan Beate Schüller, Ben Jonathan Wagner.

Writing – review & editing: Canan Beate Schüller, Ben Jonathan Wagner, Thomas Schüller, Juan Carlos Baldermann, Daniel Huys, Julia Kerner auch Koerner, Eva Niessen, Alexander Münchau, Valerie Brandt, Jan Peters, Jens Kuhn.

References

1. Robertson MM. The Gilles de la Tourette syndrome: the current status. Arch Dis Child Educ Pract Ed. 2012; 97: 166–175. https://doi.org/10.1136/archdischild-2011-300585 PMID: 22440810
2. Leckman JF. Tourette ' s syndrome. 2002; 360: 1577–1586.
3. Bloch MH, Leckman JF. Clinical course of Tourette syndrome. J Psychosom Res. 2009; 67: 497–501. https://doi.org/10.1016/j.jpsychores.2009.09.002 PMID: 19913654
4. Albin RL, Mink JW. Recent advances in Tourette syndrome research. Trends Neurosci. 2006; 29: 175–182. https://doi.org/10.1016/j.tins.2006.01.001 PMID: 16430974
5. Dwyer JB. A Developmental Perspective of Dopaminergic Dysfunction in Tourette Syndrome. Biol Psychiatry. 2018; 84: e33–e35. https://doi.org/10.1016/j.biopsych.2018.07.008 PMID: 30115244
6. Kuhn J, Baldermann JC, Huys D. Dysregulation of the Reward and Learning Systems in Tourette Syndrome. JAMA Neurology. American Medical Association; 2019. p. 1124. https://doi.org/10.1001/jamaneurol.2019.2027 PMID: 31305877
7. Frank MJ. Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism. J Cogn Neurosci. 2005; 17: 51–72. https://doi.org/10.1162/0898929052880093 PMID: 15701239
8. Denys D, de Vries F, Cath D, Figee M, Vulink N, Veltman DJ, et al. Dopaminergic activity in Tourette syndrome and obsessive-compulsive disorder. Eur Neuropsychopharmacol. 2013; 23: 1423–1431. https://doi.org/10.1016/j.euroneuro.2013.05.012 PMID: 23876376
9. Haber SN, Knutson B. The reward circuit: Linking primate anatomy and human imaging. Neuropsychopharmacology. 2010; 35: 4–26. https://doi.org/10.1038/npp.2009.129 PMID: 19812543
10. Singer HS. The neurochemistry of Tourette syndrome. In: Martina D, Leckman JF, editors. Tourette Syndrome. New York: Oxford University Press; 2013. pp. 276–300.
11. Buse J, Schoenefeld K, Münchau A, Roessner V. Neuromodulation in Tourette syndrome: dopamine and beyond. Neurosci Biobehav Rev. 2013; 37: 1069–84. https://doi.org/10.1016/j.neubiorev.2012.10.004 PMID: 23085211
12. Singer HS, Szymanski S, Giuliano J, Yokoi F, D P, Dogan a S, et al. Elevated Intrasynaptic Dopamine Release in Tourette ' s Syndrome Measured by PET. Psychiatry Interpers Biol Process. 2002; 1329–1336.
13. Maia T V., Conceição VA. The Roles of Phasic and Tonic Dopamine in Tic Learning and Expression. Biol Psychiatry. 2017; 1–12. https://doi.org/10.1016/j.biopsych.2017.05.025 PMID: 28734459
14. Morand-beaulieu S, Grot S, Lavoie J, Leclerc JB. The puzzling question of inhibitory control in Tourette syndrome: A meta- analysis. Neurosci Biobehav Rev. 2017; 80: 240–262. https://doi.org/10.1016/j.neubiorev.2017.05.006 PMID: 28502600
15. Jackson GM, Draper A, Dyke K, Pépés SE, Jackson SR. Inhibition, Disinhibition, and the Control of Action in Tourette Syndrome. Trends Cogn Sci. 2015; 19: 655–665. https://doi.org/10.1016/j.tics.2015.08.006 PMID: 26440120
16. Smith CT, San Juan MD, Dang LC, Katz DT, Perkins SF, Burgess LL, et al. Ventral striatal dopamine transporter availability is associated with lower trait motor impulsivity in healthy adults. Transl Psychiatry. 2018; 8. https://doi.org/10.1038/s41398-018-0328-y PMID: 30531858
17. Canário N, Sousa M, Moreira F, Duarte IC, Oliveira F, Januário C, et al. Impulsivity across reactive, proactive and cognitive domains in Parkinson's disease on dopaminergic medication: Evidence for multiple domain impairment. PLoS One. 2019; 14: 1–18. https://doi.org/10.1371/journal.pone.0210880 PMID: 30759108
18. Pine A, Shiner T, Seymour B, Dolan RJ. Dopamine, time, and impulsivity in humans. J Neurosci. 2010; 30: 8888–8896. https://doi.org/10.1523/JNEUROSCI.6028-09.2010 PMID: 20592211
19. Voon V. Decision-making and impulse control disorders in parkinson's disease. Decision Neuroscience: An Integrative Perspective. Elsevier; 2016. https://doi.org/10.1016/B978-0-12-805308-9.00024–5
20. Freund N, Jordan CJ, Lukkes JL, Norman KJ, Andersen SL. Juvenile exposure to methylphenidate and guanfacine in rats: effects on early delay discounting and later cocaine-taking behavior. Psychopharmacology (Berl). 2019; 236: 685–698. https://doi.org/10.1007/s00213-018-5096-0 PMID: 30411140

21. Chamberlain SR, Sahakian BJ. The neuropsychiatry of impulsivity. Curr Opin Psychiatry. 2007; 20: 255–261. https://doi.org/10.1097/YCO.0b013e3280ba4989 PMID: 17415079

22. Nederkoorn C, Van Eijs Y, Jansen A. Restrained eaters act on impulse. Pers Individ Dif. 2004; 37: 1651–1658. https://doi.org/10.1016/j.paid.2004.02.020

23. Tabibnia G, Monterosso JR, Baicy K, Aron AR, Poldrack RA, Chakrapani S, et al. Different forms of self-control share a neurocognitive substrate. J Neurosci. 2011; 31: 4805–4810. https://doi.org/10.1523/JNEUROSCI.2859-10.2011 PMID: 21451018

24. Cohen JR, Berkman ET, Lieberman MD. Intentional and Incidental Self-Control in Ventrolateral Prefrontal Cortex. Principles of Frontal Lobe Function. Oxford University Press; 2014. pp. 417–440. https://doi.org/10.1093/med/9780199837755.003.0030

25. Palminteri S, Pessiglione M. Reinforcement Learning and Tourette Syndrome. International review of neurobiology. 2013. pp. 131–153. https://doi.org/10.1016/B978-0-12-411546-0.00005-6 PMID: 24295620

26. Kéri S, Szlobodnyik C, Benedek G, Janka Z, Gádoros J. Probabilistic classification learning in Tourette syndrome. Neuropsychologia. 2002; 40: 1356–1362. https://doi.org/10.1016/s0028-3932(01)00210-x PMID: 11931939

27. Vicario CM, Gulisano M, Maugeri N, Rizzo R. Delay Reward Discounting in Adolescents With Tourette's Syndrome. Mov Disord. 2020; 35: 1279–1280. https://doi.org/10.1002/mds.28096 PMID: 32392365

28. Vincent BT. Hierarchical Bayesian estimation and hypothesis testing for delay discounting tasks. Behav Res Methods. 2016; 48: 1608–1620. https://doi.org/10.3758/s13428-015-0672-2 PMID: 26542975

29. Peters J, Büchel C. The neural mechanisms of inter-temporal decision-making: understanding variability. Trends Cogn Sci. 2011; 15: 227–239. https://doi.org/10.1016/j.tics.2011.03.002 PMID: 21497544

30. Ainslie G. Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol Bull. 1975; 82: 463–496. https://doi.org/10.1037/h0076860 PMID: 1099599

31. Bickel WK, Jarmolowicz DP, Mueller ET, Koffarnus MN, Gatchalian KM. Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: Emerging evidence. Pharmacol Ther. 2012; 134: 287–297. https://doi.org/10.1016/j.pharmthera.2012.02.004 PMID: 22387232

32. Findley TS, Caliendo FN. Time inconsistency and retirement choice. Econ Lett. 2015; 129: 4–8. https://doi.org/10.1016/j.econlet.2015.01.027

33. Ersner-Hershfield H, Wimmer GE, Knutson B. Saving for the future self: Neural measures of future self-continuity predict temporal discounting. Soc Cogn Affect Neurosci. 2009; 4: 85–92. https://doi.org/10.1093/scan/nsn042 PMID: 19047075

34. Dalley JW, Robbins TW. Fractionating impulsivity: Neuropsychiatric implications. Nat Rev Neurosci. 2017; 18: 158–171. https://doi.org/10.1038/nrn.2017.8 PMID: 28209979

35. Peters J, Büchel C. Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions. Neuron. 2010; 66: 138–148. https://doi.org/10.1016/j.neuron.2010.03.026 PMID: 20399735

36. Dreher JC, Meyer-Lindenberg A, Kohn P, Berman KF. Age-related changes in midbrain dopaminergic regulation of the human reward system. Proc Natl Acad Sci U S A. 2008; 105: 15106–15111. https://doi.org/10.1073/pnas.0802127105 PMID: 18794529

37. Arrondo G, Aznárez-Sanado M, Fernández-Seara MA, Goñi J, Loayza FR, Salamon-Klobut E, et al. Dopaminergic modulation of the trade-off between probability and time in economic decision-making. Eur Neuropsychopharmacol. 2015; 25: 817–827. https://doi.org/10.1016/j.euroneuro.2015.02.011 PMID: 25840742

38. Kayser AS, Allen DC, Navarro-Cebrian A, Mitchell JM, Fields HL. Dopamine, corticostriatal connectivity, and intertemporal choice. J Neurosci. 2012; 32: 9402–9409. https://doi.org/10.1523/JNEUROSCI.1180-12.2012 PMID: 22764248

39. Wagner B, Clos M, Sommer T, Peters J. Dopaminergic modulation of human intertemporal choice: A diffusion model analysis using the D2-receptor antagonist haloperidol. J Neurosci. 2020; 40: 7936–7948. https://doi.org/10.1523/JNEUROSCI.0592-20.2020 PMID: 32948675

40. Figner B, Knoch D, Johnson EJ, Krosch AR, Lisanby SH, Fehr E, et al. Lateral prefrontal cortex and self-control in intertemporal choice. Nat Publ Gr. 2010; 13: 538–539. https://doi.org/10.1038/nn.2516 PMID: 20348919

41. Hare TA, Hakimi S, Rangel A. Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. Front Neurosci. 2014; 8: 1–15. https://doi.org/10.3389/fnins.2014.00001 PMID: 24478622

42. Hare TA, Camerer CF, Rangel A. Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. Science (80-). 2009; 324: 646–648. https://doi.org/10.1126/science.1168450 PMID: 19407204

43. van den Bos W, Rodriguez C a., Schweitzer JB, McClure SM. Adolescent impatience decreases with increased frontostriatal connectivity. Proc Natl Acad Sci. 2015; 201423095. https://doi.org/10.1073/pnas.1423095112 PMID: 26100897

44. Anandakumar J, Mills KL, Earl EA, Irwin L, Miranda-Dominguez O, Demeter D V., et al. Individual differences in functional brain connectivity predict temporal discounting preference in the transition to adolescence. Dev Cogn Neurosci. 2018; 34: 101–113. https://doi.org/10.1016/j.dcn.2018.07.003 PMID: 30121543

45. Caswell AJ, Morgan MJ, Duka T. Inhibitory control contributes to "motor"-but not "cognitive"-impulsivity. Exp Psychol. 2013; 60: 324–334. https://doi.org/10.1027/1618-3169/a000202 PMID: 23628696

46. Muraven M. Building Self-Control Strength: Practicing Self-Control Leads to Improved Self-Control Performance. J Exp Soc Psychol. 2010; 46: 465–468. https://doi.org/10.1016/j.jesp.2009.12.011 PMID: 20401323

47. Kim BS, Im HI. The role of the dorsal striatum in choice impulsivity. Ann N Y Acad Sci. 2018; 1451: 92–111. https://doi.org/10.1111/nyas.13961 PMID: 30277562

48. Wang Z, Maia T V., Marsh R, Colibazzi T, Gerber A, Peterson BS. The neural circuits that generate tics in Tourette's syndrome. Am J Psychiatry. 2011; 168: 1326–1337. https://doi.org/10.1176/appi.ajp.2011.09111692 PMID: 21955933

49. Amlung M, Marsden E, Holshausen K, Morris V, Patel H, Vedelago L, et al. Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders. JAMA Psychiatry. 2019. https://doi.org/10.1001/jamapsychiatry.2019.2102 PMID: 31461131

50. Lempert KM, Steinglass JE, Pinto A, Kable JW, Simpson HB. Can delay discounting deliver on the promise of RDoC? Psychological Medicine. Cambridge University Press; 2019. pp. 190–199. https://doi.org/10.1017/S0033291718001770 PMID: 30070191

51. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, et al. The Yale-Brown Obsessive Compulsive Scale. Arch Gen Psychiatry. 1989; 46: 1006. https://doi.org/10.1001/archpsyc.1989.01810110048007 PMID: 2684084

52. Woods DW, Piacentini J, Himle MB, Chang S. Premonitory Urge for Tics Scale (PUTS): initial psychometric results and examination of the premonitory urge phenomenon in youths with Tic disorders. J Dev Behav Pediatr. 2005; 26: 397–403. https://doi.org/10.1097/00004703-200512000-00001 PMID: 16344654

53. Lewin AB, Piacentini J. Evidence-Based Assessment of Child Obsessive Compulsive Disorder: Recommendations for Clinical Practice and Treatment Research. Child Youth Care Forum. 2010; 39: 73–89. https://doi.org/10.1007/s10566-009-9092-8 PMID: 20376181

54. Döpfner M, Görtz-Dorten A, Lehmkuhl G. DISYPS-II—Diagnostik-System für Psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV. Bern: Huber; 2008.

55. Janßen H. Vergleichende Untersuchung der Zukunftsorientierung des Wahlverhaltens von Kindern mit und ohne Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung in einem Belohnugsaufschub-Paradigma mit realen Temporal Discounting Design. 2011.

56. Robertson M, M., Eapen V, Singer HS, Martino D, Scharf JM, et al. Gilles de la Tourette syndrome. Nat Rev Dis Prim. 2017; 3. https://doi.org/10.1038/nrdp.2016.97 PMID: 28150698

57. Peisker CB, Schüller T, Peters J, Wagner BJ, Schilbach L, Müller UJ, et al. Nucleus Accumbens Deep Brain Stimulation in Patients with Substance Use Disorder Disorders and Delay Discounting. Brain Sci. 2018; 8: 1–15. https://doi.org/10.3390/brainsci8020021 PMID: 29382059

58. Peters J, Büchel C. Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. J Neurosci. 2009; 29: 15727–15734. https://doi.org/10.1523/JNEUROSCI.3489-09.2009 PMID: 20016088

59. Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, et al. The Obsessive-Compulsive Inventory: Development and validation of a short version. Psychol Assess. 2002; 14: 485–496. https://doi.org/10.1037/1040-3590.14.4.485 PMID: 12501574

60. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J, Comrey AL, et al. An Inventory for Measuring Depression. Arch Gen Psychiatry. 1961; 4: 561. https://doi.org/10.1001/archpsyc.1961.01710120031004 PMID: 13688369

61. Wender PH, Ward F, Reimherr FW. Scale: An Aid in the Retrospective Attention. Am J Psychiatry. 1993; 150: 885–890. https://doi.org/10.1176/ajp.150.6.885 PMID: 8494063

62. Horn W. Leistungsprüfsystem. Göttingen: Hogrefe; 1983.

63. McKerchar TL, Green L, Myerson J, Pickford TS, Hill JC, Stout SC. A comparison of four models of delay discounting in humans. Behav Processes. 2009; 81: 256–259. https://doi.org/10.1016/j.beproc.2008.12.017 PMID: 19150645

64. Mazur JE. An Adjusting Procedure for Studying Delayed Reinforcment. In: Commons ML, Mazur JE, Nevin JA, Rachlin H, editors. The Effect of Delay and of Intervening Events on Reinforcement Value: Quantitative Analyses of Behavior, Band 5. Hillsdale, NJ: Erlbaum; 1987. pp. 55–87.

65. Kruschke JK. Bayesian Assessment of Null Values Via Parameter Estimation and Model Comparison. 2011. https://doi.org/10.1177/1745691611406925

66. Baldwin SA, Fellingham GW. Bayesian methods for the analysis of small sample multilevel data with a complex variance structure. Psychol Methods. 2013; 18: 151–164. https://doi.org/10.1037/a0030642 PMID: 23148476

67. Wagenmakers EJ, Marsman M, Jamil T, Ly A, Verhagen J, Love J, et al. Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. Psychon Bull Rev. 2018; 25: 35–57. https://doi.org/10.3758/s13423-017-1343-3 PMID: 28779455

68. Plummer M. JAGS: A program for analysis of Bayesian graphical models using gibbs sampling JAGS: Just another gibbs sampler. Proc 3rd Int Work Distrib Stat Comput (DSC 2003). 2003; March 20–22, Vienna, Austria. ISSN 1609-395X. https://doi.org/10.1.1.13.3406

69. Beard E, Dienes Z, Muirhead C, West R. Using Bayes factors for testing hypotheses about intervention effectiveness in addictions research. Addiction. 2016; 111: 2230–2247. https://doi.org/10.1111/add.13501 PMID: 27347846

70. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol. 2002; 64: 583–616. https://doi.org/10.1111/1467-9868.00353

71. Maia T V., Conceição VA. Dopaminergic Disturbances in Tourette Syndrome: An Integrative Account. Biol Psychiatry. 2018; 1–13. https://doi.org/10.1016/j.biopsych.2018.02.1172 PMID: 29656800

72. D'Amour-Horvat V, Leyton M. Impulsive actions and choices in laboratory animals and humans: Effects of high vs. low dopamine states produced by systemic treatments given to neurologically intact subjects. Front Behav Neurosci. 2014; 8: 1–20. https://doi.org/10.3389/fnbeh.2014.00001 PMID: 24478648

73. Petzold J, Kienast A, Lee Y, Pooseh S, London ED, Goschke T, et al. Baseline impulsivity may moderate L-DOPA effects on value-based decision-making. Sci Rep. 2019; 9: 5652. https://doi.org/10.1038/s41598-019-42124-x PMID: 30948756

74. Cools R, D'Esposito M. Inverted-U–Shaped Dopamine Actions on Human Working Memory and Cognitive Control. Biol Psychiatry. 2011; 69: e113–e125. https://doi.org/10.1016/j.biopsych.2011.03.028 PMID: 21531388

75. Whelan R, McHugh LA. Temporal discounting of hypothetical monetary rewards by adolescents adults and older adults. Psychol Rec. 2009; 59: 247–258. https://doi.org/10.1007/BF03395661

76. Worbe Y, Baup N, Grabli D, Chaigneau M, Mounayar S, McCairn K, et al. Behavioral and movement disorders induced by local inhibitory dysfunction in primate striatum. Cereb Cortex. 2009; 19: 1844–1856. https://doi.org/10.1093/cercor/bhn214 PMID: 19068490

77. Puts NAJ, Harris AD, Crocetti D, Nettles C, Singer HS, Tommerdahl M, et al. Reduced GABAergic inhibition and abnormal sensory symptoms in children with Tourette syndrome. J Neurophysiol. 2015; 114: 808–817. https://doi.org/10.1152/jn.00060.2015 PMID: 26041822

78. Yaniv A, Benaroya-Milshtein N, Steinberg T, Ruhrman D, Apter A, Lavidor M. Executive control development in Tourette syndrome and its role in tic reduction. Psychiatry Res. 2018; 262: 527–535. https://doi.org/10.1016/j.psychres.2017.09.038 PMID: 28965812

79. Ballard IC, Kim B, Liatsis A, Aydogan G, Cohen JD, McClure SM. More Is Meaningful: The Magnitude Effect in Intertemporal Choice Depends on Self-Control. Psychol Sci. 2017; 28: 1443–1454. https://doi.org/10.1177/0956797617711455 PMID: 28858559

80. Green L, Myerson J, Ostaszewski P. Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. J Exp Psychol Learn Mem Cogn. 1999; 25: 418–427. https://doi.org/10.1037//0278-7393.25.2.418 PMID: 10093208

81. Lempert KM, Steinglass JE, Pinto A, Kable JW, Simpson HB. Can delay discounting deliver on the promise of RDoC? Psychol Med. 2019; 49: 190–199. https://doi.org/10.1017/S0033291718001770 PMID: 30070191

82. Mischel W, Ebbesen EB. Attention in delay of gratification. J Pers Soc Psychol. 1970; 16: 329–337. https://doi.org/10.1037/h0029815

83. Johnson MW. An efficient operant choice procedure for assessing delay discounting in humans: Initial validation in cocaine-dependent and control individuals. Exp Clin Psychopharmacol. 2012; 20: 191–204. https://doi.org/10.1037/a0027088 PMID: 22329554

84. Jimura K, Myerson J, Hilgard J, Braver TS, Green L. Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. Psychon Bull Rev. 2009; 16: 1071–1075. https://doi.org/10.3758/PBR.16.6.1071 PMID: 19966257

85. Mischel W, Shoda Y, Peake PK. The Nature of Adolescent Competencies Predicted by Preschool Delay of Gratification. J Pers Soc Psychol. 1988; 54: 687–696. https://doi.org/10.1037//0022-3514.54.4.687 PMID: 3367285

86. Mischel W, Shoda Y, Rodriguez MI. Delay of gratification in children. Science (80-). 1989/05/26. 1989; 244: 933–938. Available: http://www.ncbi.nlm.nih.gov/pubmed/2658056 https://doi.org/10.1126/science.2658056 PMID: 2658056

87. Gawrilow C, Gollwitzer PM, Oettingen G. If-then plans benefit delay of gratification performance in children with and without ADHD. Cognit Ther Res. 2011; 35: 442–455. https://doi.org/10.1007/s10608-010-9309-z

88. Steele CC, Gwinner M, Smith T, Young ME, Kirkpatrick K. Experience Matters: The Effects of Hypothetical versus Experiential Delays and Magnitudes on Impulsive Choice in Delay Discounting Tasks. Brain Sci. 2019; 9: 379. https://doi.org/10.3390/brainsci9120379 PMID: 31888218

89. Smits RR, Stein JS, Johnson PS, Odum AL, Madden GJ. Test-retest reliability and construct validity of the experiential discounting task. Exp Clin Psychopharmacol. 2013; 21: 155–163. https://doi.org/10.1037/a0031725 PMID: 23421359

90. Patt V, Hunsberger R, Jones D, Keane M, Verfaellie M. Temporal discounting when outcomes are experienced in the moment: Validation of a novel paradigm and comparison with a classic hypothetical intertemporal choice task. 2020 [cited 29 Apr 2021]. https://doi.org/10.31234/osf.io/5ajuk

Chronic deep brain stimulation of the human nucleus accumbens region disrupts the stability of inter-temporal preferences.

Ben J. Wagner[1§*], Canan B. Schüller[2§], Thomas Schüller [2], Juan C. Baldermann [2], Sina Kohl [2], Veerle Visser-Vandewalle [3], Daniel Huys [2], Jens Kuhn [2,4§] and Jan Peters [1§]

[1] Department of Psychology, Biological Psychology, University of Cologne, Germany.

[2] Department of Psychiatry and Psychotherapy, Medical Faculty, University of Cologne, Germany.

[3] Department of Stereotactic and Functional Neurosurgery, Faculty of Medicine and University Hospital Cologne, Germany.

[4] Department of Psychiatry, Psychotherapy and Psychosomatic, Johanniter Hospital Oberhausen, Germany.

[§] Shared first/senior authorship.

[*] Correspondence should be addressed to: B.W. (ben.jonathan.wagner@uni-koeln.de).

Abstract

When choosing between rewards that differ in temporal proximity (inter-temporal choice), human preferences are typically stable, constituting a clinically-relevant transdiagnostic trait. Here we show in patients undergoing deep brain stimulation (DBS) to the anterior limb of the internal capsule / nucleus accumbens for treatment-resistant obsessive-compulsive disorder, that chronic (but not acute) DBS disrupts inter-temporal preferences. Findings support a contribution of the human nucleus accumbens region to preference stability over time.

Main text

Humans continuously maneuver the world weighting the future against the present. The degree of temporal discounting of future rewards as assessed via inter-temporal choice tasks is a stable trait [1] with relevance for a range of psychiatric conditions [2]. For example, steep discounting of value over time is a hall-mark of addiction [3]. Multiple neural systems contribute to human self-control, including prefrontal cortex (PFC) regions involved in cognitive control, and regions of the mesolimbic and mesocortical dopamine system (e.g. ventral striatum and ventro-medial PFC) involved in reward valuation [4–6]. Earlier studies focused on characterizing potentially dissociable striatal and prefrontal value signals during inter-temporal choice [7–9]. This debate

has ultimately led to a revised view of self-control, according to which lateral PFC exerts top-down control over ventromedial PFC in support of self-controlled choices[10–12] . Despite the finding that both striatal reward responses[13] and cortico-striatal connectivity [14] are associated with temporal discounting in cross-sectional analyses, this model is largely silent with respect to the contribution of the ventral striatum.

Here we address this issue by capitalizing on the rare opportunity to longitudinally follow patients undergoing therapeutic deep brain stimulation (DBS) of the nucleus accumbens region for treatment-resistant obsessive-compulsive disorder (OCD). OCD is assumed to be associated with a dysregulation in fronto-striatal circuits [15], which can be normalized via anterior limb of the internal capsule / nucleus accumbens (ALIC/NAcc) DBS [16–19]. In the context of a DBS treatment-efficacy study [20] we examined acute and chronic effects of DBS on temporal discounting.

Patients with OCD and matched controls (for inclusion criteria and demographics see methods and Supplemental Table S1, for DBS stimulation details see methods and Supplemental Table S2) completed three separate testing sessions: one initial testing session at T1 (patients: pre-DBS, controls: session one) and two follow-up sessions at T2 (T1-T2 test-retest interval mean [range] in days for patients: 203 [155-260], controls: 206 [164-247]). The T2 sessions were spaced within a week (patients: DBS on vs. off in counterbalanced order with at least 24h wash-out; controls: sessions two and three). N=7 patients completed all three testing sessions. Two additional patients completed T1 but only one of the T2 sessions (total n=9 for pre vs. post DBS analyses). Eight additional patients were only tested at T2 (in total n=15 for on vs. off DBS analyses). N=28 controls participated, with one control missing the final testing session (yielding n=27 for the corresponding analysis). Due to the Covid-19 pandemic, ten of the controls completed testing at T2 online. On each testing day, participants completed a temporal discounting task involving 140 choices between smaller-immediate (20 €) and larger-but-later rewards (individualized amounts ranging between 20.5 – 80€, see methods). One trial per session was selected at random and paid out in cash or via timed bank transfer.

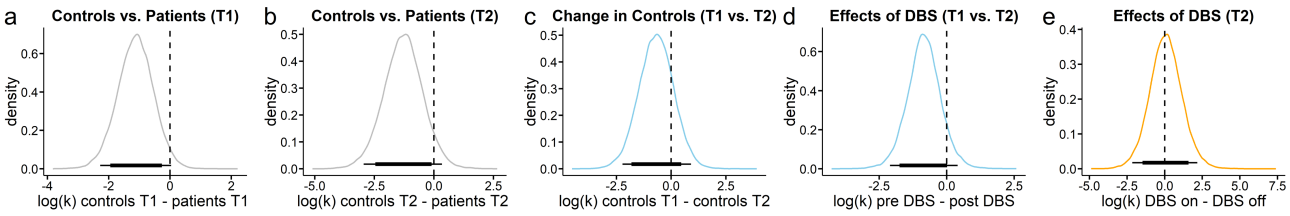


Figure 1. Group-level changes in inter-temporal choice. **a**, At T1 (pre DBS), partients (n = 9) discounted delayed rewards steeper compared to controls (n = 28) at the first session (directional Bayes Factor

(dBF) = 35.75). **b**, pooled second and third sessions in controls (n = 28) vs. pooled DBS on and DBS off in patients (n = 9). (T2; controls < patients; dBF = 15.85). **c** and **d**: Controls and patients tended to discount rewards steeper after six months of time (controls T1 < T2 (n = 28); dBF = 4.15) or 6 months of continuous stimulation (patients pre DBS < post DBS (n = 15); dBF = 10.87). **e**, Discounting on the group level did not reveal changes with consistent directionality following acute DBS (DBS on < DBS off (n = 15); dBF = 0.90). Thin (thick) horizontal lines denote the 95% (85%) highest density intervals.

Data were modeled for each time point separately using hierarchical Bayesian parameter estimation (see methods) and a hyperbolic discounting model with softmax action selection. In line with previous work [21], OCD patients exhibited increased discounting (a higher discount rate log[k]), both pre DBS at T1 (Figure 1a) and across on and off DBS sessions at T2 (Figure 1b). Controls showed no systematic change in discounting between the two T2 testing sessions (see Supplemental Table S3). There was no systematic effect of acute DBS on temporal discounting (n=15, Figure 1e). If anything, rewards where discounted somewhat steeper after six months (controls: Figure 1c) or six months of continuous DBS (patients: Figure 1d), a pattern observed previously in healthy participants [1]. An overview of all group comparisons is provided in Supplemental Table S3, and the corresponding analyses for decision noise (softmax β) are provided in Supplemental Figure S1.

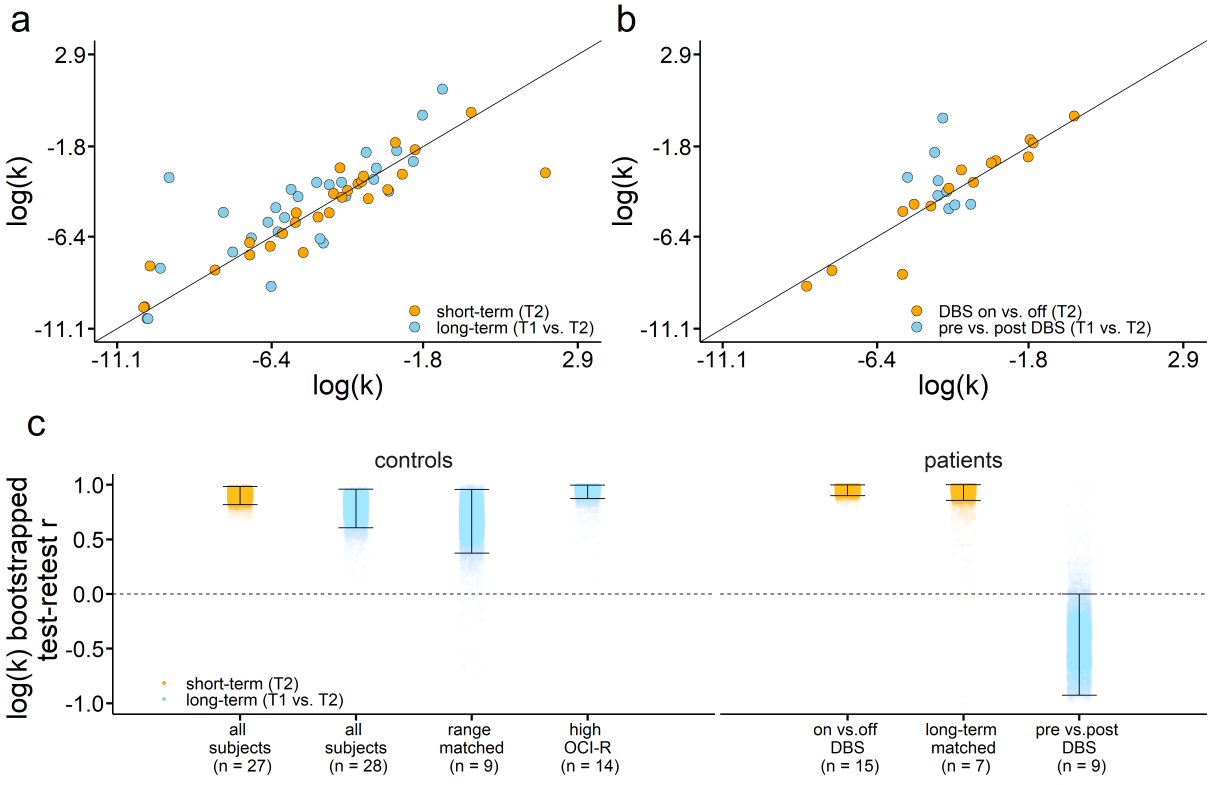


Figure 2. Reliability of inter-temporal choice (discount rate log[k]). **a**, and **b**, show individual discount rates *log(k)* at different points in time (task sessions at T1 and T2) in controls (a) and patients (b). The

x- and y axis with regard to the short-term data (orange) refer to the comparison within T2 (**a**: control data session 2 (x-axis) vs. session 3 (y-axis). With regard to the long-term data (blue) the axis refer to session 1 (x-axis) vs. pooled sessions 2 and 3 (y-axis). **b**: patient short-term data (orange) DBS on (x-axis) vs. DBS off (y-axis). In the long-term condition data (blue) the x-axis shows discount-rates in the first/prior to DBS task session and data on the y-axis refers to the pooled sessions at T2. **c**, bootstrapped correlation coefficients in controls and patients.

To study DBS effects on the stability of inter-temporal preferences, we applied both inter- and intra-individual analytical approaches. In controls, the discount rate log(k) exhibited the expected high test-retest reliability (Figure 2a), both between sessions two and three at T2 (one-week short-term stability, bootstrapped mean r = .90, 95% highest density interval [HDI] = .82 – .98 see Figure 2c) and between T1 and pooled T2 data (6-month long-term stability, bootstrapped mean r = .80, 95% HDI = .61 – .96, see Figure 2c). All bootstrapped correlation values are provided in Supplemental Table S4. In patients (Figure 2b), short-term test-retest reliability (DBS on vs. off, n=15) was comparable to controls (bootstrapped mean r = .96, 95% HDI = .90 – 1.0, see Figure 2c;). This was also the case when examining only patients who completed all three sessions (n=7, see Figure 2c, DBS on vs. off matched patients; bootstrapped mean r = .95, 95% HDI = .86 – 1.0).

In stark contrast, long-term stability in patients was completely disrupted (T1 pre-DBS vs. pooled T2 post-DBS, n=9, bootstrapped mean r = -.44, 95% HDI = -.93 – .00; see Figure 2c). This group difference was not due to range restriction in the patients (range-matched subset of n=9 controls: bootstrapped mean r = .68, 95% HDI = .37 – .96, Figure 2d). Likewise, it was unlikely attributable to the presence of OCD symptoms, as a subset of n=14 controls with high OCI-R scores (mean [range] = 23.57 [14 – 40]) still exhibited high long-term test-retest reliability (bootstrapped mean r = .95, 95% HDI = .87 – 1.0; see Figure 2d). Furthermore, the long-term test-retest reliability in DBS patients was lower than that of any n=9 sub-sample of controls with similarly narrow ranges of log(k) values (Supplemental Figure S2). Reliability in controls was similar for lab vs. online testing due to Covid-19 lockdown (Supplemental Figure S3).

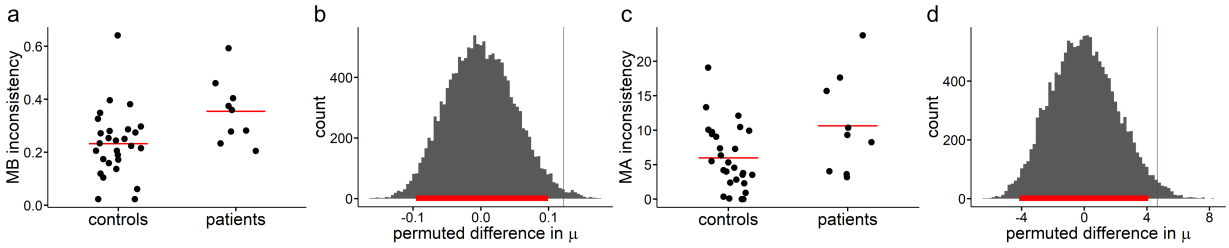


Figure 3. Within-subject changes in inter-temporal preferences. **a** and **c,** model-based (MB) T1-T2 choice inconsistency (deviation between T2 choices and model-predictions based on a decision model fitted to T1 data, see main text and methods) and model-agnostic (MA) T1-T2 choice inconsistency (mean absolute deviation of indifference points between T1 and T2, see main text and methods). **b** and **d** permutation test for model-based and model-agnostic inconsistency scores respectively. Histograms show null distributions of mean group differences across 10k randomly shuffled group labels; red vertical lines: observed group differences; red horizontal line: 95% highest density interval.

We next tested whether a disruption of preference stability would also manifest at the level of individual decisions, using both model-based and model-agnostic measures. First, we extracted individual-subject median discount-rates (log[k]) and decision noise parameters (β) from our hierarchical Bayesian model estimated on T1 data (see methods) to compute choice probabilities for each T2 decision (pooling across sessions). We then computed a model-based choice inconsistency score as the mean deviation of predicted and observed choices at T2 for each participant (higher values correspond to greater inconsistency). A permutation-based group comparison using 10k randomly shuffled group labels revealed a significant increase in patients (permutation test: p = 0.01, Figure 3a, b). This difference held when groups were matched on decision noise across a range of thresholds (see Supplemental Figure S4).

As a model-agnostic measure of within-participant changes in preferences, we computed the mean absolute change in indifference points from T1 to T2 (see methods and Supplemental Figures S5 and S6 for single-subject data). This confirmed a greater increase in patients vs. controls (permutation test, p = 0.018, Figure 3b, c). Inconsistency measures did not correlate with years of education, pre-post DBS symptom severity scores, overall duration of OCD or the T1-T2 interval (see Supplemental Table S5). These analyses suggest that choices at T2 deviated from T1 more after six months of continuous ALIC/NAcc DBS.

Taken together, we show using both inter- and intra-individual analyses that ALIC/NAcc DBS disrupts the long-term (but not short-term) stability of inter-temporal preferences in OCD patients undergoing DBS treatment. This suggests that in addition to short-term plasticity processes[22], long-term ALIC/NAcc DBS[23] can interfere with the expression of inter-temporal preferences that are thought to rely on the this same circuitry [4,5,7,13]. While earlier

reports noted effects of acute stimulation on risk-taking and impulsivity [24,25] (albeit with acute block-wise stimulation protocols), our longitudinal analysis revealed changes only following prolonged stimulation. Our data do not suggest a specific direction of change, nor do they reflect an association with a change in OCD symptoms. Rather, the data suggests a fundamental role of the nucleus accumbens region in maintaining preference stability over time. The exact cellular mechanisms underlying the DBS effects remain speculative [17,26] – potential effects range from DBS acting as an informational lesion, to changes in inter-regional functional connectivity[17] and a general modulation of oscillatory activity and in consequence pathological circuitry[26].

In summary, our data extend neural models of self-control[12] and inter-temporal choice[4,5] by revealing a contribution of the human nucleus accumbens region to the maintenance of preference stability over time. These findings reveal a case of subtle long-term modulation of higher cognitive function via DBS that further studies might elaborate on.

Acknowledgements

C.S. was funded under the Walter and Marga Boll Foundation and the Institutional Strategy of the University of Cologne within the German Excellence Initiative (ZUK 81/1). J.P. was supported by Deutsche Forschungsgemeinschaft (PE 1627/5-1). We thank Milena Marx for help with recruitment and testing of control participants.

Competing financial interests

J.K. has occasionally received honoraria from AstraZeneca, Lilly, Lundbeck, and Otsuka Pharma for lecturing at conferences and financial support to travel. He received financial support for investigator-initiated trials from Medtronic Europe SARL (Meerbusch, Germany). The remaining authors reported no biomedical financial interests or potential conflicts of interests.

Author contributions

J.K., J.P. C.S. and B.W. designed the experiment. C.S. and B.W. acquired the data. C.S. and B.W. analyzed the data. B.W. performed the modeling and statistical analyses. J.P. and J.K. supervised the project. B.W. and J.P. wrote the paper, and all authors provided revisions.

References

1. Kirby, K.N. One-year temporal stability of delay-discount rates. *Psychonomic bulletin & review* **16**, 457–462 (2009).

2. Amlung, M. *et al.* Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. *JAMA psychiatry* (2019).
3. Bickel, W.K., Koffarnus, M.N., Moody, L. & Wilson, A.G. The behavioral- and neuro-economic process of temporal discounting: A candidate behavioral marker of addiction. *Neuropharmacology* **76 Pt B**, 518–527 (2014).
4. Peters, J. & Büchel, C. The neural mechanisms of inter-temporal decision-making: understanding variability. *Trends in cognitive sciences* **15**, 227–239 (2011).
5. McClure, S.M. & Bickel, W.K. A dual-systems perspective on addiction: contributions from neuroimaging and cognitive training. *Annals of the New York Academy of Sciences* **1327**, 62–78 (2014).
6. Bartra, O., McGuire, J.T. & Kable, J.W. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* **76**, 412–427 (2013).
7. Kable, J.W. & Glimcher, P.W. The neural correlates of subjective value during intertemporal choice. *Nature neuroscience* **10**, 1625–1633 (2007).
8. Kohl, S. *et al.* Deep brain stimulation for treatment-refractory obsessive compulsive disorder: a systematic review. *BMC psychiatry* **14**, 214 (2014).
9. McClure, S.M., Laibson, D.I., Loewenstein, G. & Cohen, J.D. Separate neural systems value immediate and delayed monetary rewards. *Science (New York, N.Y.)* **306**, 503–507 (2004).
10. Peters, J. & D'Esposito, M. Effects of Medial Orbitofrontal Cortex Lesions on Self-Control in Intertemporal Choice. *Current biology : CB* **26**, 2625–2628 (2016).
11. Figner, B. *et al.* Lateral prefrontal cortex and self-control in intertemporal choice. *Nature neuroscience* **13**, 538–539 (2010).
12. Hare, T.A., Camerer, C.F. & Rangel, A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science (New York, N.Y.)* **324**, 646–648 (2009).
13. Hariri, A.R. *et al.* Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**, 13213–13217 (2006).
14. van den Bos, W., Rodriguez, C.A., Schweitzer, J.B. & McClure, S.M. Connectivity strength of dissociable striatal tracts predict individual differences in temporal discounting. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **34**, 10298–10310 (2014).

15. Robbins, T.W., Vaghi, M.M. & Banca, P. Obsessive-Compulsive Disorder: Puzzles and Prospects. *Neuron* **102**, 27–47 (2019).
16. Smolders, R. *et al.* Deep brain stimulation targeted at the nucleus accumbens decreases the potential for pathologic network communication. *Biological psychiatry* **74**, e27-8 (2013).
17. Figee, M. *et al.* Deep brain stimulation restores frontostriatal network activity in obsessive-compulsive disorder. *Nature neuroscience* **16**, 386–387 (2013).
18. Smith, E.E. *et al.* A brief demonstration of frontostriatal connectivity in OCD patients with intracranial electrodes. *NeuroImage* **220**, 117138 (2020).
19. Wu, H. *et al.* Deep brain stimulation for refractory obsessive-compulsive disorder (OCD): emerging or established therapy? *Molecular psychiatry* (2020).
20. Huys, D. *et al.* Open-label trial of anterior limb of internal capsule-nucleus accumbens deep brain stimulation for obsessive-compulsive disorder: insights gained. *Journal of neurology, neurosurgery, and psychiatry* **90**, 805–812 (2019).
21. Sohn, S.Y., Kang, J. in, Namkoong, K. & Kim, S.J. Multidimensional measures of impulsivity in obsessive-compulsive disorder: cannot wait and stop. *PloS one* **9**, e111739 (2014).
22. Kuhn, J. & Baldermann, J.C. Elucidating neural network changes induced by deep brain stimulation for OCD. *Brain : a journal of neurology* **143**, 1293–1296 (2020).
23. Denys, D. *et al.* Deep brain stimulation of the nucleus accumbens for treatment-refractory obsessive-compulsive disorder. *Archives of general psychiatry* **67**, 1061–1068 (2010).
24. Nachev, P. *et al.* Dynamic risk control by human nucleus accumbens. *Brain : a journal of neurology* **138**, 3496–3502 (2015).
25. Luigjes, J. *et al.* Deep brain stimulation increases impulsivity in two patients with obsessive-compulsive disorder. *International clinical psychopharmacology* **26**, 338–340 (2011).
26. Lozano, A.M. & Lipsman, N. Probing and regulating dysfunctional circuits using deep brain stimulation. *Neuron* **77**, 406–424 (2013).
27. Sutton, R.S. & Barto, A.G. *Reinforcement learning. An introduction to Reinforcement Learning* (MIT Press, Cambridge, Mass., 1998).
28. Beard, E., Dienes, Z., Muirhead, C. & West, R. Using Bayes factors for testing hypotheses about intervention effectiveness in addictions research. *Addiction (Abingdon, England)* **111**, 2230–2247 (2016).
29. Efron, B. Bootstrap Methods: Another Look at the Jackknife. *Ann. Statist.* **7**, 1–26 (1979).

Methods

Participants

All participants provided informed written consent prior to participation, and the study procedure was approved by the Ethics Committee of the Medical Faculty of the University of Cologne.

OCD Patients

OCD-Patients eligible for DBS had to meet the DSM-IV criteria for OCD, a Yale-Brown Obsessive Compulsive Scale (Y-BOCS) over 25, at least one cognitive-behavioral therapy (minimum of 45 sessions), at least two unsuccessful treatments with a serotonin reuptake inhibitor (SSRI) and one unsuccessful augmentation with either lithium, neuroleptics or buspirone. Patients were excluded due to drug, medication or alcohol abuse, suicidal ideation, mental retardation, pregnancy or breastfeeding and schizophrenia. Disease duration was on average 27.59 ± 13.02 years ranging from 6 to 48 years. The mean age at onset for OCD was 16.2 ± 9.25 years. For further details see Supplemental Table S1.

Controls

Exclusion criteria were drug, medication or alcohol abuse or the diagnosis of a psychiatric disorder. Controls were screened for OCD-symptoms via the OCI-R questionnaire. Here 9/28 subjects scored above the threshold (>21) for possibly obtaining OCD.

Sample size

N=9 patients and n=28 controls completed testing at T1 (session 1 and pre-DBS). N=15 patients completed DBS on and off sessions at T2. Out of the N=9 patients who completed pre DBS testing, N=7 completed both DBS on and DBS off testing at T2, whereas one patient missed the DBS off session, and one patient missed the DBS on session. N=27 controls completed both testing sessions at T2 (sessions 2 and 3).

Temporal discounting task

Prior to the first testing session, participants completed a short adaptive pretest to estimate the individual discount- rate (k). This discount rate was used to construct a set of 140 participant-specific trials using MATLAB (version 8.4.0. Natick, Massachusetts: The MathWorks Inc). The task consisted of choices between an immediate smaller-sooner reward of 20€ and participant specific larger-but-later (LL) rewards delivered after some delay (1, 2, 7, 14, 30, 90

or 180 days). In 70 trials, LL amounts were uniformly spaced between 20.5 € and 80 €, whereas in the remaining 70 trials LL amounts were uniformly spaced around each estimated indifference point per delay (based on the pre-test discount rate). If indifference points were larger than 80€, only uniformly-spaced LL amount were used. Trials were presented in a pseudorandomized order. Participants were informed that after task completion, one trial would be randomly selected and paid immediately in cash (smaller-sooner choice) or via a timed bank transfer (larger-but-later choice).

Deep brain stimulation

DBS was applied to the anterior limb of the internal capsule and nucleus accumbens region. Details on electrode placement (including reconstruction of electrode positions), surgical procedure and adjustment of stimulation parameters are available elsewhere[20]. Final stimulation amplitudes ranged from 2.6 to 4.8 volt and pulse-width was set between 60 and 150 µs (see Supplemental Table S2 for details). The frequency of DBS was 130 Hz except for two patients who received 150 Hz stimulation.

Computational Modeling

To assess inter-temporal preferences, applied a standard single-parameter hyperbolic discounting model:

$$SV(LL_t) = \frac{A}{1 + \exp(k) * D} \qquad (Eq. 1)$$

Here, A is the numerical reward amount of the LL option. The discount-rate (k) models the steepness of the hyperbolic discounting curve, with greater values corresponding to steeper discounting. Delay D of the LL option is expressed in days. To improve numerical stability of the estimation, k was estimated and is reported in logarithmic space. SV then corresponds to the subjective (discounted) value of the delayed option. We then used softmax action selection[27] (Eq. 2) to model the probability of selecting the LL option on trial t. Here, β is an inverse temperature parameter, modeling choice stochasticity. For $\beta = 0$, choices are random, and as β increases, choices become more dependent on option values:

$$P(LL_t) = \frac{\exp\big(\beta * SV(LL_t)\big)}{\exp\big(\beta * SV(SS)\big) + \exp\big(\beta * SV(LL_t)\big)} \qquad (Eq. 2)$$

Hierarchical Bayesian parameter estimation

Models were fit to all trials from all participants, separately per group and time point, using a hierarchical Bayesian modeling approach. We included separate group-level gaussian distributions for log(k) and β for patients and controls, and/or T1 and T2 time points. Parameter estimation was performed using Markov Chain Monte Carlo as implemented in the JAGS software package (Plummer, 2003) (Version 4.3) in combination with R (Version 3.4) and the R2Jags package. For group-level means, we used uniform priors defined over numerically plausible parameter ranges ([-20, 3] for log(k); [0, 10] for β). We initially ran 2 chains with a burn-in period of 400k samples and thinning of two. Chain convergence was then assessed via the Gelman-Rubinstein convergence diagnostic $\hat{R}$ and sampling was continued until $1 \leq \hat{R} \leq 1.01$ for all group-level and individual-subject parameters. 10k additional samples were then retained for further analysis. We then show posterior group distributions for all parameters of interest as well as their 85% and 95% highest density intervals (HDIs). For group comparisons (T1 vs. T2 or patients vs. controls) we report Bayes Factors for directional effects for the hyperparameter difference distributions, estimated via kernel density estimation using R (Version 4.01) via RStudio (Version 1.3) interface.

Analysis of group differences

To characterize differences between patients and controls, changes from T1 to T2 or within T2 (e.g. on/off DBS) we show posterior difference distributions and the corresponding 85 % and 95 % highest density intervals. We then report Bayes Factors for directional effects. These were computed as the ratio of the integral of the posterior difference distribution from 0 to +∞ vs. the integral from 0 to -∞. Using common criteria [28], we considered Bayes Factors between 1 and 3 as anecdotal evidence, Bayes Factors above 3 as moderate evidence and Bayes Factors above 10 as strong evidence. Bayes Factors above 30 and 100 were considered as very strong evidence.

Bootstrap analyses - test-retest reliability

We analyzed the group-level reliability of inter-temporal choice (log[k]) from T1 to T2 (*long-term stability*) and within a week at T2 (*short-term stability*). Distributions of test-retest correlation coefficients were estimated via bootstrapping [29]. To this end, pairs of individual-participant median log(k) values were sampled with replacement 15k times. We then report the mean and 95 % HDI of the resulting bootstrap-distribution of correlation coefficients.

Due to differences in group size and the relative and absolute range of log(k) values in patients we performed additional control analyses. Specifically, we repeated this bootstrap analysis for all sub-samples of N=9 controls with adjacent log(k) values, yielding twenty bootstrap correlations corresponding to sub-samples of the control group with maximally similar log(k) values. Results are shown in Supplemental Figure S2.

Model-based choice inconsistency

For both model-based and model-agnostic within-participant changes, we leveraged the fact that participants completed the exact same 140 choices at each testing session. To examine model-based changes in preferences, we extracted individual-participant median discount-rates log(k) and decision noise parameters (softmax β) from our hierarchical Bayesian model estimated on T1 data. We then used these parameters to compute choice probabilities for each T1 choice. To examine model-based preference changes from T1 to T2, we then subtracted the T1 choice probability from the corresponding observed choices at T2 (0 for smaller-sooner and 1 for larger-later choices). We then computed a choice inconsistency score as the mean of the absolute differences between T1 choice probabilities and T2 choices. Across the whole sample controls showed lower decision noise when compared to patients with OCD (see Supplemental Figure S1) which was also reflected in an overall reduced model fit in patients (Supplemental Table S6). To account for this in the model-based inconsistency analysis, we additionally matched groups on β. This eliminated group differences in model fit (Supplemental Table S6) but critically did not affect group differences in model-based inconsistency (Supplemental Figure S4).

Model free analysis of indifference points

Model-based analyses rely on specific mathematical assumptions regarding the shape of the discounting function. Furthermore, they can be affected by potential group differences in model fit. Therefore, we additionally examined a model-agnostic measure of within-participant changes in preferences. To this end, we fit sigmoid functions (see Eq. 3) to the choice data of each delay D per participant and time point T:

$$P\left(LL_{D,T}\right) = \frac{1}{1+\exp\left(-\left(A_{D,T}-c\right)*b\right)} \ (Eq.\,3).$$

That is, we modeled the probability to choose the delayed reward for delay D at time point T for each participant as a sigmoid function of the absolute LL reward amount A. Here, c is the inflection point of the sigmoid (corresponding to the indifference point, i.e. the point of

subjective equivalence between the delayed reward and the immediate reward at the respective delay D), and *b* is the slope.

For delays with only larger-later choices, the indifference point was set to the midpoint between the immediate reward (20€) and the smallest available LL reward. For delays with only smaller-sooner choices, the indifference point was conservatively set to max(LL). These rules were also applied in cases where there was only a single noisy LL or SS choice for a given delay. Using this procedure, we computed 196 indifference points in controls and 63 indifference points in patients. Six indifference points in controls and two in OCD patients could not be estimated. We then computed the mean absolute deviation in indifference points between T1 and T2 as a model-agnostic measure of preference consistency.

Individual-participant choice data for each session and estimated indifference points are plotted in Supplemental Figure S5 (patients) and S6 (controls).

Permutation-based group comparisons

Model-based and model-agnostic consistency measures (see previous sections) were compared between groups using permutation tests. To this end, we compared the observed group difference in preference consistency to a null-distribution of preference consistency that was obtained by randomly shuffling the group labels 10k times, and computing the group difference for these shuffled data. Significance was assessed using a two-tailed *p*-value of 0.05.

Supplemental Material

Supplemental Table S1. Demographic data. Scores are Mean (SD).

Long-term reliability	**Controls (n = 28)**	**Patients (n = 9)**	**Group Comparison**
Age (yrs)	40.2 (9.0)	41.4 (11.6)	$t_{(11.223)}$ =-0.30, p = 0.77
Sex (F/M)	14/14	4/5	$X^2_{(1)}$ < 0.001, p = 1
Education (yrs)	11.9(1.4)	10.7(1.4)	$t_{(13.325)}$ = 2.34, p = 0.04
Short-term reliability	**Controls (n = 27)**	**Patients (n = 15)**	
Age (yrs)	40.1 (9.1)	47.4 (11.3)	$t_{(21.944)=}$ -2.08, p = 0.05
Sex (F/M)	13/14	8/7	$X^2_{(1)}$ = 0, p = 1
Education (yr)	11.9 (1.4)	10.8 (1.5)	$t_{(24.4)}$ = 2.26, p = 0.03
Long-term (*β*-matched)	**Controls (n = 14)**	**Patients (n = 9)**	
Age (yrs)	41.6 (9.9)	41.4 (11.6)	$t_{(14.305)}$ =.044, p = 0.97
Sex (F/M)	9/8	5/4	$X^2_{(1)}$ < 0.001, p = 1
Education (yrs)	11.35	10.8	$T_{(27.577)}$ = 0.96, p = 0.34

Supplemental Table S2. Overview of sex, age, disease duration before surgery and stimulation parameters (monopolar, case anode, all bilateral, except for patient 13) of DBS patients with OCD. DBS, deep brain stimulation; F, Female; Hz, Hertz; L, Left; M, Male; µs, microsec; OCD, obsessive-compulsive disorder; R, Right.

ID	**Sex**	**Age**	**Years of OCD**	**Electrode contacts**	**Frequency**	**Amplitude**	**Pulse-width**	**Sessions**
1	W	26	21	L: 0- 1- R: 8- 9-	130 Hz	5.3 V	120 µs	Pre/On
2	M	25	14	L: 3- 1- R: 11-, 10-	150 Hz	2.6 V	120 µs	Pre/On/Off
3	F	50	25	L: 3- 2- R: 11-, 10-	130 Hz	4.7 V	150 µs	Pre/On/Off
4	F	34	18	L: 3- 2- R: 11- 10-	150 Hz	4.3 V	120 µs	Pre/On/Off
5	M	47	20	L: 3- 2- R: 11- 10-	130 Hz	3.3 V	120 µs	Pre/On/Off
6	M	45	37	L:2- 1- R: 10- 9-	130 Hz	4.8 V	150 µs	Pre/On/Off
7	F	54	47	L: 3- 2- R: 11-, 10-	130 Hz	4.6 V	90 µs	Pre/On/Off
8	W	36	10	L: 3-, 2- R: 11- 10-	130 Hz	4.7 V	150 µs	Pre/On
9	M	56	36	L: 3- 2- R: 11- 10-	130 Hz	3.5 V	60 µs	Pre/On/Off

10	M	48	6	L: 0- 1- R: 8- 9-	130 Hz	6.5 V	150 µs	On/Off
11	F	34	18	L: 3- 2- R: 11- 10-	90 Hz	2.5 V	180 µs	On/Off
12	F	36	27	L: 3- 2- R: 11- 10-	130 Hz	4.2 V	120 µs	On/Off
13	M	64	28	L: 3- 2- R: 11- 10-	130 Hz	5 V	120 µs	On/Off
14	F	57	37	L: 3- 2- R: 11- 10-	130 Hz	4.5 V	120 µs	On/Off
15	M	59	48	L: 0- 1- 2- R. 8- 9- 10-	130 Hz	6 V	150 µs	On/Off
16	F	54	48	L: 3- 2- R: 11- 10-	130 Hz	4 V	150 µs	On/Off

Supplemental Table S3. Posterior log(*k*) differences. We report mean posterior differences (M_{diff}) and Bayes factors for directional effects.

Group comparison	**M_{diff}**	**dBF**
Controls T1 (n = 28) < patients T1 (n = 9)	-1.13	35.75
Controls T2 (n = 28) < patients T2 (n = 9)	-1.29	15.85
Controls S2 < controls S3 (n = 27)	0.15	0.73
Controls T1 < controls T2 (n = 28)	-0.68	4.15
Patients DBS on < patients DBS off (n = 15)	0.08	0.90
Patients pre-DBS < patients post DBS (n = 9)	-0.84	10.87

Supplemental Table S 4. Bootstrapped correlation coefficients and highest density intervals

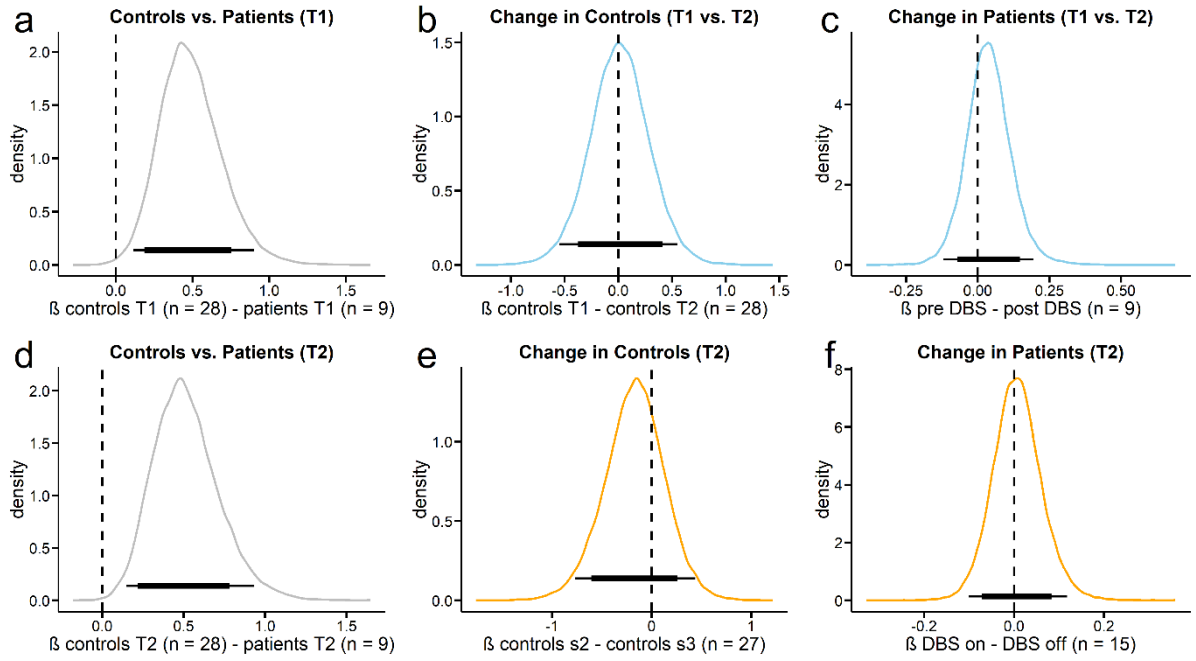
Group comparison	**Mean**	**95% HDI**
controls T1 - controls T2 (n = 28)	.80	.61 – .96
controls S2 - controls S3 (n = 27)	.90	.82 – .98
controls T1 – controls T2 range matched (n = 9)	.67	.37 – .96
controls T1 – controls T2 high OCI-R (n = 14)	.95	.87 – 1.0
patients DBS on - patients DBS off (n = 15)	.96	.90 – 1.0
patients DBS on vs. DBS off matched (n = 7)	.95	.86 – 1.0
patients pre-DBS - patients post DBS (n = 9)	-.44	-.92 – 0.0

Supplemental Table S5. Pearson correlations for model-based and model-agnostic choice inconsistency measures with demographic/clinical variables of interest.

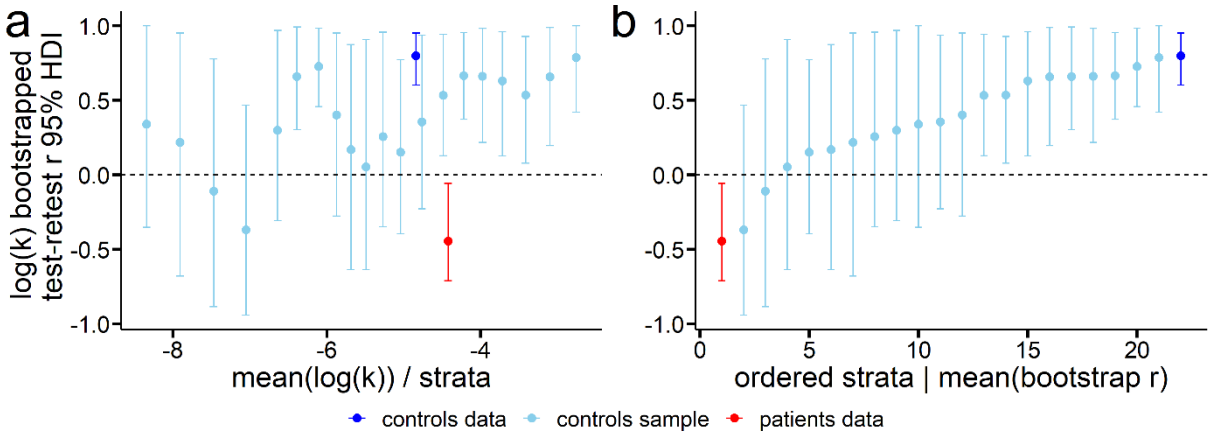
	Patients (n=9)	**Controls (n=28)**
MB inconsistency vs. years of education	r = -0.18, p = 0.63	r = -0.01, p = 0.94
MA inconsistency vs. years of education	r = -0.43, p = 0.24	r = -0.17, p = 0.37
MB inconsistency vs. pre-post YBOCS	r = -0.01, p = 0.99	-
MA inconsistency vs. pre-post YBOCS	r = -0.13, p = 0.73	-
MB inconsistency vs. OCD duration (yrs)	r = -0.25, p = 0.51	-
MA inconsistency vs. OCD duration (yrs)	r = -0.20, p = 0.60	-
MB inconsistency vs. T1-T2 interval (d)	r = -0.03, p = 0.94	r = -0.07, p = 0.74
MA inconsistency vs T1-T2 interval (d)	r = -0.36, p = 0.73	r = -0.02, p = 0.91

Supplemental Table S6. Mean (SD) proportion of correctly predicted choices of the hierarchical Bayesian model (hyperbolic discounting + softmax). At T1 the model performed better in controls than in patients. This difference was not significant at T2. This difference also trivially disappeared when groups were matched on decision noise (softmax β).

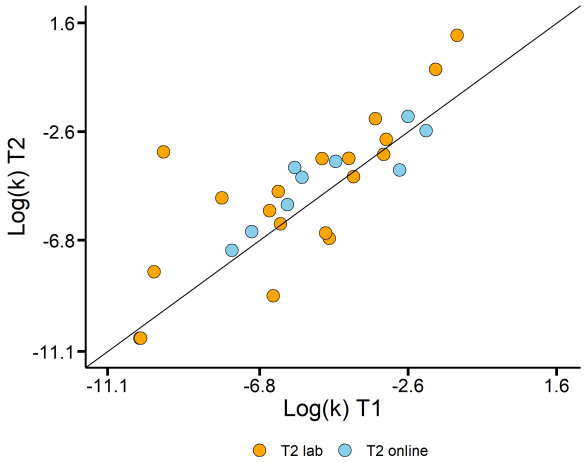
Predicted choices	**Patients (n=9)**	**Controls (n=28)**	**Group Comparison**
T1	0.80 (0.07)	0.88 (0.07)	$t_{(13.839)} = 2.73$; p = 0.02
T2	0.82 (0.09)	0.87 (0.08)	$t_{(12.833)} = 1.65$; p = 0.12
		Controls matched on β (cut-off = 0.5, n=17)	
T1	0.80 (0.07)	0.84(0.06)	$t_{(13.656)} = 1.35$; p =0.20
T2	0.82 (0.09)	0.85(0.06)	$t_{(14.044)} = 0.96$; p =0.35
		Controls matched on β (cut-off = 0.41, n=14)	
T1	0.80 (0.07)	0.83(0.06)	$t_{(13.656)} = 1.18$; p =0.26
T2	0.82 (0.09)	0.83(0.07)	$t_{(14.606)} = 0.51$; p =0.62



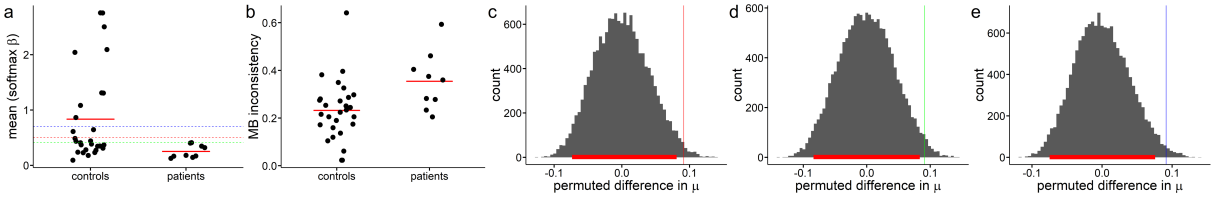
Supplemental Figure S1. Group differences and temporal changes in softmax β. **a/d**, On the group level, OCD patients exhibited lower β values than controls (**a:** T1, controls < patients; directional Bayes Factor (dBF) = 0.003, **d:** T2, controls < patients; directional Bayes Factor (dBF) = 0.0009). β values show inconsistent changes from T1 to T2 (**b:** controls; **c:** patients), or at T2 within each group (**e:** controls, **f:** patients). Thin (thick) horizontal lines denote the 95% (85%) highest density intervals.



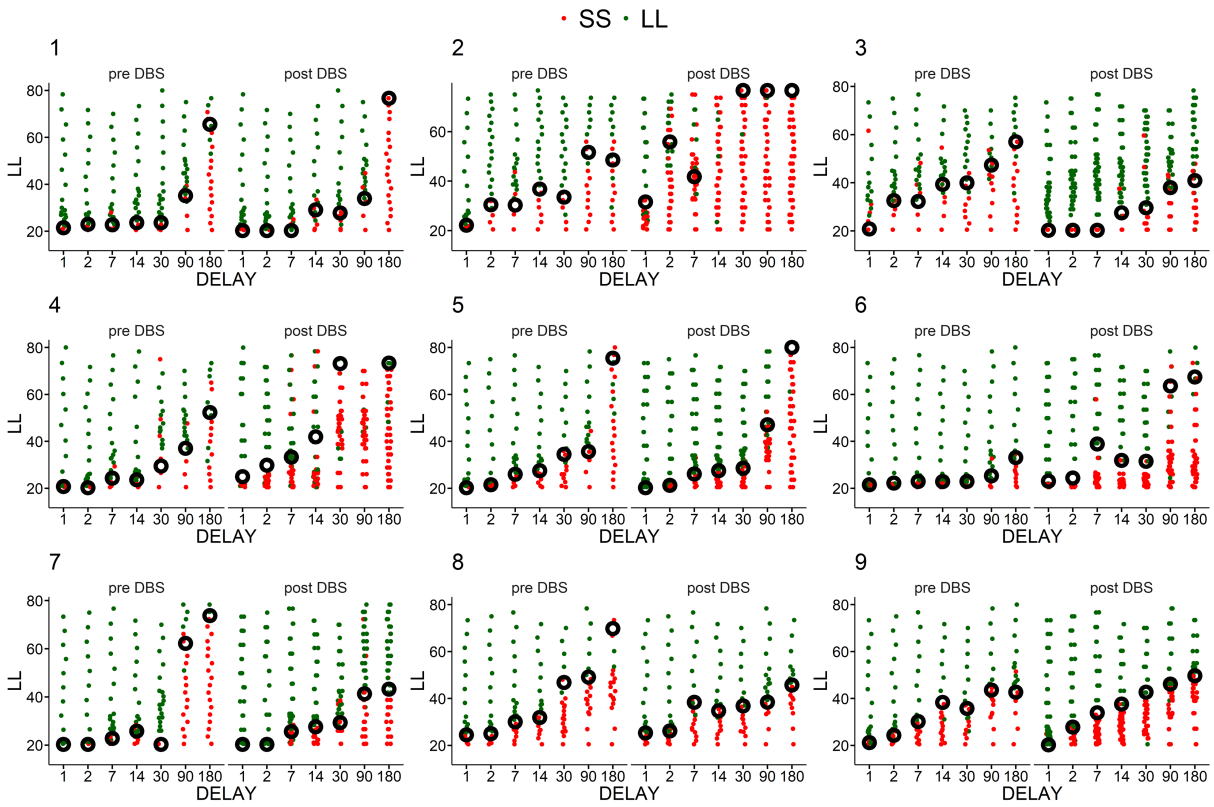
Supplemental Figure S2. Bootstrap analysis across the whole range of log(k) values in controls. y-axis: mean value of each strata´s bootstrap distribution of correlation coefficients with 95 % HDI. In **a**, stratas are ordered according to the mean log(k) value of the strata. In **b**, stratas are ordered according to the mean bootstrap correlation coefficient.



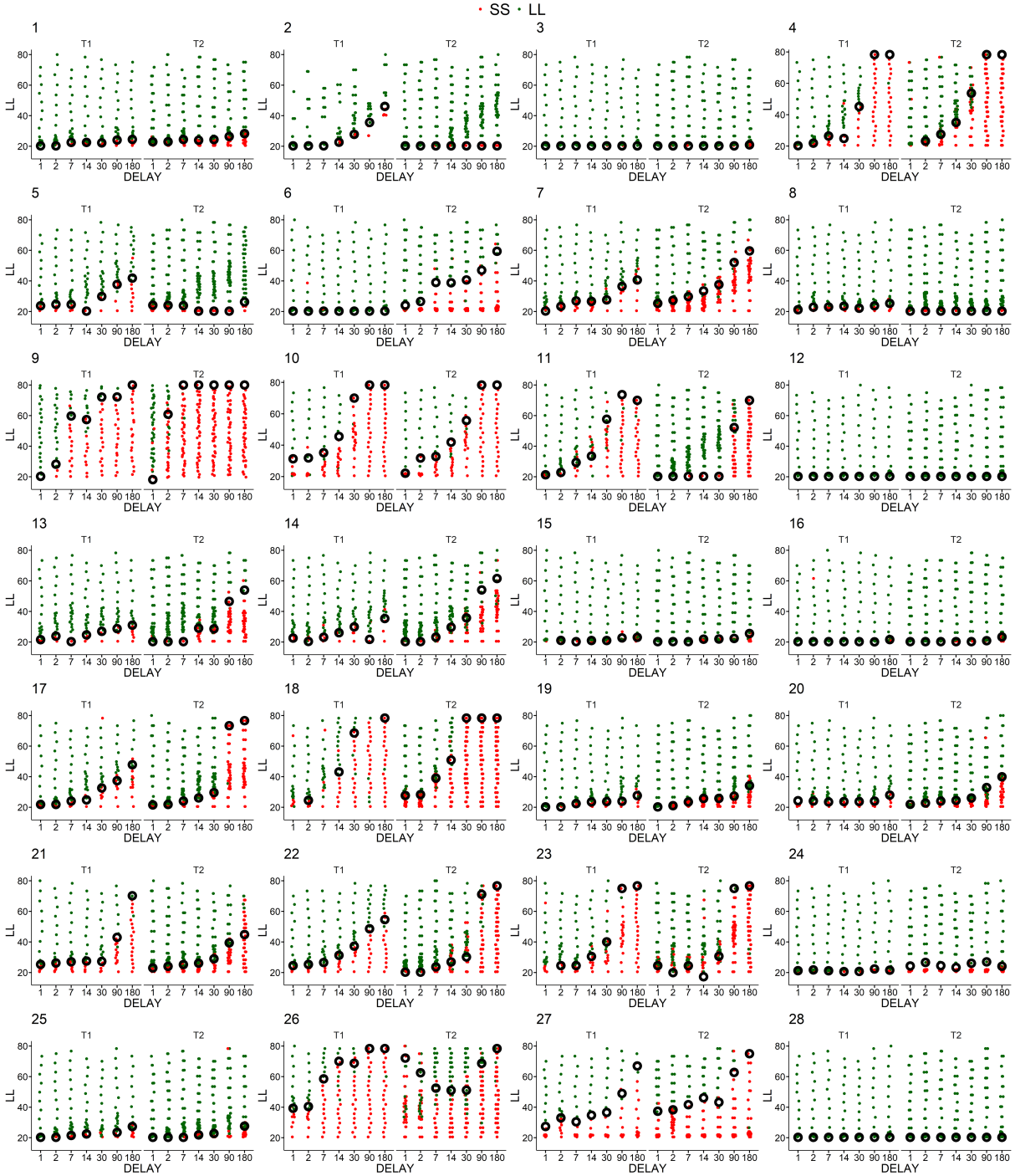
Supplemental Figure S3. Comparison of T2 lab and T2 online sessions. Participants that performed the task in an online-version at T2 (6 months after T1) and participants that completed T2 testing in the lab showed a similar long-term test-retest reliability.



Supplemental Figure S4. a, Individual-subject softmax β parameters (mean of T1 and T2 posterior medians) modelling decision noise. To account for effects of group differences in β, in additional analyses, controls matched on β to the patients were selected across various β cut-offs (β=0.5: red dotted line, n=17; β=0.41: green dotted line, n=14; β=0.7: blue dotted line, n=19). **b**, Individual-subject model-based T1-T2 choice inconsistency in β-matched controls (β cutoff = 0.5*)* and patients. **c-e**, Permutation test results for group differences in model-based choice inconsistency (histograms of group differences with 10k randomly shuffled group labels) across all cut-offs (**c:** β=0.5, **d:** β=0.41, **e:** β=0.7). In line with the analysis across the whole sample, patients showed significantly increased model-based inconsistency for all cut-offs (**c:** p = 0.018, **d:** p = 0.023, **e:** 0.01). Vertical lines in c-e denote the observed group difference, horizontal red lines denote the 95% HDI of the null distribution.



Supplemental Figure S2. Single-subject choice data for all n=9 patients with pre- and post DBS data. Green and red points represent LL and SS choices, respectively, across LL amounts (y-axis) and delays (x-axis). Black circles show estimated indifference-points.



Supplemental Figure S3. Single-subject choice data for all n=28 controls. Green and red points represent LL and SS choices, respectively, across LL amounts (y-axis) and delays (x-axis). Black circles show estimated indifference-points.

General Discussion

The present work assessed intertemporal choice, that is the process of value-based decision-making between two temporally exclusive outcomes. In general humans and other animals discount delayed rewards (Samuelson 1937; Mazur and Coe 1987; Peters and Büchel 2011). Cognitive processes like reward valuation, reward learning and motivation to work for reward largely depend on DAergic structures, i.e. the mesolimbic DA system (Schultz 2013, 2015; Berridge 2016; Berke 2018; Westbrook et al. 2021) . Besides this, the DAergic system is implicated in nearly all psychiatric diseases, and basically all neuroleptic drugs available for treatment target the DA system (Ayano 2016; Meiser et al. 2013; Beaulieu et al. 2015).

The Research Domain Criteria (RDoC) framework of the National Institute of Mental Health called out research to find new and liable tools that measure continuous dimensions of human function. The aim of this initiative is to foster our understanding of the dimensions underlying mental health (Insel et al. 2010; Kozak and Cuthbert 2016; U.S. Department of Health and Human Services, National Institutes of 2016). Research on intertemporal choice, while existing for more than 100 years, has recently gained traction given multiple research groups (Lempert et al. 2019; Bickel et al. 2019) acknowledged its potential as a transdiagnostic marker and potential diagnostic tool that fits within the RDoC framework.

Another promising decision-making construct that was additionally assessed in *Study 2* within this dissertation targets the differentiation of so-called MF and MB contributions to RL. While a MF agent learns stimulus-reward associations, a MB agent additionally takes the structure of the environment into account (Daw 2011). Impairments in MB RL have been associated with habitual tendencies, a lack of goal-directed control and might constitute a general impairment across a range of compulsive disorders (Voon et al. 2015; Gillan et al. 2016; Gillan et al. 2020). Studies in this dissertation contributed to our understanding of such important decision-making constructs in the following ways.

Effects of the D_2-Receptor Antagonist Haloperidol in Healthy Participants

In *Study 1* ("*Dopaminergic Modulation of Intertemporal Choice: a Diffusion Model Analysis using the D_2 Receptor Antagonist Haloperidol*") our results revealed substantially smaller discount-rate parameters (decreased discounting) under haloperidol vs. placebo consistently in two different magnitude conditions. These results challenge one prior study where an effect of haloperidol was absent (Pine et al. 2010), but line up with other studies investigating the effects of moderate increases in DA neurotransmission on discounting (Wit 2002; Hamidovic et al. 2008; Weber et al. 2016). Our data strengthen the view that modulating DA neurotransmission

via lower doses of a D_2-receptor antagonist, here haloperidol, likely enhances DA release by predominantly binding at presynaptic D_2 autoreceptors (Schwarz et al. 2004; Chen et al. 2005; Frank and O'Reilly 2006). This interpretation is supported by an unrelated memory task during fMRI that participants completed prior to the temporal discounting task. Here, the analyses revealed an overall main effect of drug condition on trial onset-related activity in the caudate nucleus, that is caudate activity was increased under haloperidol (Clos et al. 2019). Further, our DDM approach revealed substantially shorter non-decision times (~ 180ms) in the haloperidol group when contrasted to the placebo group. A substantial increase in RT components is more compatible with an increase in DA neurotransmission (Weed and Gold 1998; Guitart-Masip et al. 2011; Beierholm et al. 2013).

Our results are further theoretically plausible with new perspectives and computational models on DAs role in cost-benefit decision-making (Collins and Frank 2014; Westbrook and Braver 2016; Westbrook et al. 2020; Westbrook et al. 2021). Here enhanced DS DA synthesis capacity or elevated DA neurotransmission via methylphenidate and a D_2 receptor antagonist was associated with an increased willingness to invest in effort for receiving reward (Westbrook et al. 2020). A DDM analyses revealed that these subjects increasingly focused on the benefits (reward information) in contrast to the cost (effort information) of their decisions (Westbrook et al. 2021). Note, our results are only compatible with these findings if one defines "delay" as "costs" and "higher reward" as "benefits" of decision options. Limitations and open questions of this perspective are discussed below.

Contextual Effects of Gambling Environments

In *Study 2* ("*Gambling Environment Exposure Increases Temporal Discounting but Improves Model-based Control in Regular Slot-Machine Gamblers*") we examined the contextual modulation of temporal discounting (Lempert et al. 2019; Bickel et al. 2019) and MB control (Gillan et al. 2020; Gillan et al. 2016) in regular slot machine gamblers. Regular slot machine gamblers have previously been characterized by high levels of temporal discounting (Wiehler and Peters 2015) and reduced MB control (Wyckmans et al. 2019). Our results here show that these putatively transdiagnostic traits are modulated in different ways. First, gambling environment exposure increased temporal discounting. A finding in line with lab experiments on gambling cue exposure. In these studies the exposure to gambling cues, i.e. pictures of gambling venues/machines, or sounds increased delay discounting (Miedl et al. 2014; Dale et al. 2019; Genauck et al. 2020). We further replicated a similar finding of Dixon et al. (2006) where regular bet facility visitors showed increased discounting in the context of

betting facilities in contrast to neutral (cue-free) locations (Dixon et al. 2006). Importantly, we further link our findings of increased discounting to maladaptive control beliefs. That is, gamblers who`s discount-rates increased to a greater extend were likewise characterized by higher scores on the Gambling Related Cognition Scale (GRCS) (Raylu and Oei 2004). Taken together this suggests that some underlying mechanism contributes to both, gambling cognition and choice impulsivity in gambling environments. Secondly, we consistently observed increased MB control and reduced MF control of behavior when gamblers performed the RL task in the gambling environment. Importantly, this finding contrasted with our a-priori hypothesis of reduced MB control in the gambling context. Addiction is generally thought to rely on extensive habit formation (Robbins and Everitt 1999; Barry J Everitt and Trevor W Robbins 2005) and cue exposure is believed to trigger those habits (Antons et al. 2020).

However, if one renders MF control of RL as habitual and MB control as goal directed, a common interpretation (Daw 2011; Gillan et al. 2016), our findings contrast with these assumptions. Thus, our results are important experimental findings for the understanding of addiction. In fact, critics of habit theory emphasized that addiction might be associated with excessive goal-directed behavior, in particular in the presence of addiction-related cues (Hogarth 2020). Our data supports this latter perspective. Overall, these context effects suit the perspective of incentive salience /incentive sensitization theory (Robinson and Berridge 1993; Robinson and Berridge 2008). Further evidence for this interpretation stems from the observation of increased subjective craving in this environment. Increased craving / wanting fits likewise well with incentive sensitization theory (Robinson and Berridge 1993; Robinson and Berridge 2008), which proposes that addiction-related environments exert their influence on behavior in part via a potentiation in dopamine release (Berridge 2016; Anselme and Robinson 2013; Robinson and Berridge 2001), but see section on limitations below. Increases in DA neurotransmission has also been associated with increased MB- (Wunderlich et al. 2012; Sharp et al. 2016) and decreased MF control (Kroemer et al. 2019). The gambling context might thus enhance goal-directed control via an improved construction and/or utilization of the task transition structure potentially modulated by DA. This interpretation furthermore resonates with the computational assumptions of OPAL (Collins and Frank 2014) and experimental findings assuming that DA enhances the willingness to spend cognitive effort (Westbrook and Braver 2016; Westbrook et al. 2020; Westbrook et al. 2021).

Given the latter interpretation and our findings from *Study 1*, it is not straightforward why discounting is modulated differently given a suggested increase in DA in both studies. One possible explanation is, that besides modulating the willingness to promote effort and focus on

the benefits of decision outcomes (Westbrook et al. 2021; Mikhael et al. 2021; Schwartenbeck et al. 2015), DA also signals opportunity costs. These opportunity costs, in theory, do track the average rate of reward in the environment and are suggested to be reflected in VS dopamine, while the willingness to exert and focus on benefits was suggested to be associated with DS DA (Westbrook et al. 2021; Collins and Frank 2014).

Being speculative, these high opportunity costs in the context of gambling environments might modulate performance on intertemporal choice task and RL strategies in the Two Step Task differently. While performance in the Two Step Task can yield bigger reward, performance on the intertemporal choice task does not immediately pay off. Further, all participants performed on the intertemporal choice task first. It is possible that the effects of cue exposure / incentive salience decreased over time, because participants habituated to the environment. Such differential effects are compatible with the perspective of an inverted-U-function of DA where the effects of DA on decision-making differ depending on the degree of DA enhancement (Cools and D'Esposito 2011; Maia and Frank 2017).

Moreover, mechanism like future imagination or self-control (see section on *Theoretical Background*) contribute to intertemporal choice (Peters and Büchel 2010b, 2011). While the willingness to invest in cognitive effort might be increased, spontaneous future imagination might be decreased in the presence of addiction related cues and high opportunity costs. Future prospection has been shown to attenuate temporal discounting in a range of settings (Rösch et al. 2021). While a general impairment of future imagination was not found in gamblers under lab conditions (Wiehler et al. 2015), it is still possible that future imagination is decreased in the context of gambling venues. Gambling venues are designed in a specific way, i.e. they have no windows, subtle lighting and other features that make customers focus on the present and their gambling activities (Sulkunen et al. 2021). Interestingly, it was recently proposed that while higher levels of tonic DA in rodents should generally decrease discounting (promote selection of the larger reward), this relationship is reversed when time perception is distorted (Mikhael and Gershman 2021). Further, participants with vulnerabilities for addiction might in general tend to focus on the present, while under the influence of cues or contexts endowed with high levels of incentive salience (Flagel et al. 2009).

In total our findings show that two computational trans-diagnostic markers with high relevance to gambling disorder in particular and addiction more generally are modulated in opposite ways by exposure to real-life gambling environments. These findings posit a challenge for habit/compulsion theories of addiction. Ecologically valid testing settings such as those investigated here can thus yield novel insights into environmental drivers of maladaptive

behavior underlying mental disorders, i.e. effect sizes during naturalistic cue exposure (e.g. the present study and Dixon et al., 2006) were substantially larger than during lab-based exposure in previous studies (Miedl et al. 2014; Genauck et al. 2020; Dale et al. 2019).

Intertemporal Choice in Patients with Tourette Syndrome

In *Study 3* ("*Temporal Discounting in Adolescents and Adults with Tourette Syndrome*") we examined intertemporal preferences in adolescent and adult patients with TS. TS is a complex neuropsychiatric disorder associated with developmental disturbances in DAergic neurotransmission. These DAergic anomalies are believed to either cause, enable or enhance tics like sudden movements, muscle contractions or phonic and repetitive sounds (Bloch and Leckman 2009; Denys et al. 2013; Maia and Conceição 2018). Moderate pharmacological increases in DA neurotransmission have been associated with decreased discounting (see above). However, studies showed that patients with TS show impairments in RL tasks, likely caused by DA associated reward sensitivity (Palminteri and Pessiglione 2013; Palminteri et al. 2009). To date, effects of the suggested DA hyperinnervative state in TS (Maia and Conceição 2018) on intertemporal preferences are unclear [but see: (Vicario et al. 2020)].

Our results here suggest that the proposed pathophysiology of TS does not give rise to substantial changes in temporal discounting in adult patients with TS. However, in a second study, where adolescent patients with TS performed a experiential discounting task [comparable with (Mischel et al. 1988; Mischel et al. 1989; Johnson 2012)] we found evidence for decreased discounting in TS patients, contrasting with another recent study (Vicario et al. 2020). Functional and structural frontostriatal connectivity undergoes maturation until early adulthood (van den Bos et al. 2015; Jackson et al. 2015; Anandakumar et al. 2018). We speculate that this process might be strengthened in adolescence with TS. Patients with TS are constantly faced by tics and the need to control their motor output. Learning to inhibit tics might foster the ability to inhibit other impulses, thereby strengthening cognitive control more generally (Muraven 2010). However, our adult and adolescent sample performed different tasks. Demands for experiential tasks (adolescents) with delay in the range within 60 seconds and hypothetical tasks (adults) with delays upon months might substantially differ. This is because the delay in the experiential task is directly experienced and the time horizon is substantially different. In consequence, the need for specific subprocesses involved in intertemporal choice, i.e. future imagination, might be less pronounced in the a task where a small delay (in seconds) is experienced immediately.

It is likewise possible that TS pathophysiology in general affects decision subprocesses differently, i.e. effects of a hyperresponsive valuation network [as proposed by: (Palminteri and Pessiglione 2013; Palminteri et al. 2009)] and effects of regions associated with self-control might offset each other's contribution to discounting. However, it was not possible to further disentangle these processes in our study design. Future studies would benefit from adopting a consistent longitudinal approach to further elucidate the developmental trajectory TS and its association with decision-making subprocesses involved in intertemporal choice.

Effects of Chronic Deep Brain Stimulation

In *Study 4* ("*Chronic Deep Brain Stimulation Disrupts the Stability of Intertemporal Preferences*") we used a longitudinal approach to examine effects of acute and chronic DBS in OCD patients. OCD is characterized by impulsions, compulsions and pathophysiology in CSTC loops (Robbins et al. 2019; Kashyap et al. 2012). DBS is one treatment option for otherwise therapy refractory patients and studies report that DBS is effective in restoring pathological activity in CSTC loops (Figee et al. 2013; Wu et al. 2020). Our results show that ALIC /NAcc DBS disrupts the long-term (but not short-term) stability of inter-temporal preferences. While in general intertemporal preferences are relatively stable over time (Kirby 2009) preferences changed significantly more in patients after long-term DBS than in controls. We therefore conclude that long-term ALIC/NAcc DBS (Denys et al. 2010) can disrupt relatively stable choice preferences. Earlier reports noted effects of acute stimulation on risk-taking and impulsivity (Nachev et al. 2015; Luigjes et al. 2011) , however, effects of chronic DBS on the stability of choice preference processes have not been observed to date.

The exact cellular mechanisms underlying changes via DBS are still not completely resolved (Lozano and Lipsman 2013; Figee et al. 2013; Robbins et al. 2019) and potential effects range from DBS acting as an informational lesion, to changes in inter-regional functional connectivity and pathological circuitry (Lozano and Lipsman 2013; Figee et al. 2013). However, our data does not suggest a specific direction or association with changes in OCD symptoms and therefore insights and associations with DBS pathophysiology are somewhat limited. Nevertheless, our results suggest a contribution of the human NAcc region to the maintenance of preference stability over time. In consequence, our findings suggest that long-term DBS can modulate higher cognitive functions that rely on the same CSTC-loops that are affected in OCD (Robbins et al. 2019). Further studies should use this knowledge and thus more carefully monitor subtle long-term effects of DBS on human behavior.

Computational Cognitive Modelling

In *Study 1* and *Study 2* we applied a recent class of value-based sequential-sampling models based on the drift diffusion model (DDM) (Pedersen et al., 2017; Fontanesi et al., 2019; Shahar et al., 2019; Peters and D'Esposito, 2020; see Methods). In both of these studies comprehensive RT-based analysis extended previous studies via a detailed decomposition of the decision process. In our study on DAergic modulation of intertemporal choice in healthy participants this analysis revealed that, beside changes in the discount-rate the haloperidol group was characterized by substantial reductions (~ 180ms) in non-decision time (motor components of the decision process). Thus, our computational assessment of the data allowed for further implications regarding the drug effects of haloperidol as such robust enhancement of lower-level motor and/or perceptual RT components are more compatible with an increase in DA transmission (e.g. Weed and Gold, 1998). An inspection of parameter correlations suggested that both parameters (discount-rate log[k] and non-decision time t_0) might capture aspects of the decision process with an coherent underlying mechanism. Moreover, this finding resonates with previous studies suggesting a DAergic enhancement of RT-based response vigor (Guitart-Masip et al. 2011). Our computational modelling results raise evidence for a pre-synaptic effect of D_2-antagonism given low doses of haloperidol. Importantly, these implications would have not been possible via standard modelling approaches (e.g. softmax action selection).

The same DDM approach enabled us to detect that contextual effects on intertemporal choice in *Study 2* are associated with an attenuation in non-decision time, which mirrors effects of pharmacological enhancement of dopamine transmission in *Study 1*. In contrast to these results, we observed a substantial *increase* in maximum drift rate in the gambling context, reflecting increased value sensitivity of RTs. Pharmacological effects of D_2-antagonism and contextual effects of gambling environments though differ in how they affect latent decision components, i.e. value sensitivity and non-decision time. We further applied our comprehensive RT-based analysis to each stage of the Two Step RL task (Shahar et al. 2019). Our analysis here revealed that latent decision processes (DDM parameters) were largely unaffected by environmental context. Thus, we hypothesize that the gambling venue primarily affected the weighting of MF and MB contributions to evidence accumulation. In consequence we can be more certain that gambling venues most exclusively enhance MB RL strategies, instead of changing other aspects of the decision process. Importantly, model comparison and posterior predictive checks of single subject data showed that in both studies a DDM with non-linear trial-wise drift rate scaling captured the relationship of decision conflict (SS - subjective LL

value differences) and RTs best (Peters and D'Esposito 2020; Wagner et al. 2020; Fontanesi et al. 2019). We further report extensive parameter recovery analyses of our data in *Study 1* and confirm that group-level parameters recovered well (Peters and D'Esposito 2020). In detail, recovery of individual-subject baseline parameters (100 € magnitude condition) was excellent, whereas recovery of parameters modeling condition effects (20 € magnitude condition) was ok. Overall, in both studies the DDM-based modeling approach allowed us to examine the dynamics underlying decision-making much more comprehensively than in previous human pharmacological or cue-reactivity studies (Hamidovic et al. 2008; Wit 2002; Pine et al. 2010; Weber et al. 2016; Petzold et al. 2019; Dixon et al. 2006; Miedl et al. 2014; Genauck et al. 2020; Dale et al. 2019).

Computational modelling in *Study 3* and *Study 4* followed more standard protocols, i.e. softmax-action selection and logistic regression. This decision was due to task constrains, that in consequence rendered RTs not suitable for a similar DDM analysis. However, central findings are: First, data of the experiential discounting task in adolescent patients with TS in *Study 3* was best fit by an exponential discounting function. We propose this to be a consequence of the specific task design. Decisions on the experiential task in adolescents unfolded within seconds upon to a minute, whereas decision outcomes in the adults task unfolded after the task and up to weeks and months in the future.

Secondly, in our analysis in *Study 4* it was necessary to quantify long-term change in intertemporal preferences while controlling for decision noise (softmax[β]). In detail, changes in long-term stability of intertemporal preferences were quantified via changes in trial-wise choice probability for one or the other choice option. However, when both the control and DBS group differ in the decision-noise parameter (softmax[β]), long-term changes in the trial-wise choice probabilities could be caused by changes or a baseline difference in choice stochasticity itself. Therefore one needs to match both groups on decision noise. In this case we decided to only use controls that fall in the same range on stochasticity as patients. This was done because the control group was substantially larger (n = 30 vs. n = 9 patients in the DBS group). Said differently, the substantial aim of this study was to analyze changes in choice preference. Moreover, a more accurate decomposition of the decision process would have complicated the interpretation of the results, simply because it would be unlikely to find consistent effects within such a low number of data points in the DBS group. Unfortunately, it is extremely difficult to find a substantial amount of patients participating on cognitive tasks prior to DBS electrode implantation. Nevertheless, in the ideal case one would prefer to rely on a substantial larger sample size, obtain RTs and further decompose the decision processes via a more

comprehensive analysis. Taken together, studies in this dissertation demonstrate the advantage of sophisticated computational modelling. These methods can draw implications from behavioral data alone and therefore support conclusions (contextual effects or aspects of pharmacological modulation) that would have been much harder to infer without those methods. Further, in some cases it might be useful to analyze data using more standard approaches, i.e. when sample size is limited or data needs to be constrained in specific ways.

Limitations

Aside from the limitations already mentioned above some further limitations and open questions remain. First, we did not directly measure DAergic activity during task performance and therefore our interpretations of DAergic effects are based on pharmacological studies, computational assessment and further implications based on other research and prominent theories in the literature. To illustrate the complexity: While there is evidence that small doses of the D_2-antagonist haloperidol (and likewise amisulpride and sulpiride) increase DA neurotransmission (Frank and O'Reilly 2006; Westbrook et al. 2020; Clos et al. 2019), it is unlikely that these effects are exclusive to specific regions. On the one hand, research suggests a higher ratio of D_2 receptors are found in the DS (Ford 2014; Seamans and Yang 2004). This is also in line with activation patterns following haloperidol administration prior to task performance in our study (Clos et al. 2019) and further theoretically plausible with DAs effect on benefits vs. costs (Westbrook et al. 2020). However, it was also suggested that D_2-antagonists primarily modulate phasic DA (Benoit-Marand et al. 2011), but see (Zhang et al. 2009), and the ratio of phasic vs. tonic DA is believed to be highest in the VS, especially the NAcc area (Tsai et al. 2009). These degrees of freedom can further complicate the exact regional consequences of D_2 antagonism. Moreover, D_2-autoreceptors are commonly found in the amygdala and hippocampus area (Ford 2014; Beaulieu and Gainetdinov 2011), implying the possibility of an effect on processes associated with MTL regions, like future imagination or emotional valence. Thus, neuronal effects of haloperidol on intertemporal choice might be region specific (Wächtler et al. 2020), but it's unknown how these regional effects contribute to intertemporal choice. If we want to understand the exact effects of modulating intertemporal choice via D_2-autoreceptor antagonism this needs to be evaluated carefully.

Intertemporal choice / delay discounting has proven to be a reliable construct with high long-term stability (Kirby 2009; Lempert et al. 2019), however, it is still unknown if this translates to MB control in the Two Step Task assessed in *Study 2*. While split-half reliability proved to be quite good (Shahar et al. 2019), to date likewise one study reported relatively low

overall reliability (Enkavi et al. 2019). Future research needs to clarify if the Two Step Task measures a reliable trait-like construct or is better described as a state-like construct (for trait-state effects in intertemporal choice please see Peters and Büchel [2011]). In *Study 2* we further report a robust increase in impulsive choice and MB RL strategies from the neutral to the gambling environment. While we have some evidence that this finding is associated with gambling cues and exclusive to participants susceptible to these cues (we report an association with GRCS and gambling severity compound scores and observed increased craving in the gambling context; see above) it is not clear if these effects are absent in other subjects, i.e. in a matched control group without prior gambling experience. Said differently, it remains unknown if these context effects are caused by other environmental differences that modulate decision-making independent of gambling cue-reactivity. For example, other context depended changes like novelty effects [novelty and excitement both modulate DA neurotransmission: e.g. (Schultz 2016; Duszkiewicz et al. 2019; Linnet et al. 2011)], a general interest in gambling/ motivation to gamble or arousal (McClelland et al. 1987) might modulate task performance via effects on DA neurotransmission. Moreover, gambling machines are designed to attract attention and might thus constitute salient stimuli to controls (Ungless 2004; Winton-Brown et al. 2014; Duszkiewicz et al. 2019; Linnet et al. 2011; Schultz 2016). Therefore the inclusion of a control group and a careful assessment of other constructs (i.e. novelty, general interest in gambling) that theoretically modulate DA neurotransmission could contribute to our understanding of the underlying mechanisms.

This study is a first step in the direction of ecologically valid testing settings. Context effects such as those investigated here can yield novel insights into environmental drivers of behavior and mental disorders. However, while human behavior is generally shaped by real-world environment it is overly harder to disentangle the unique factors that modulate behavior in such complex environments. For example, the number of customers present varied across participants and context, affecting noise levels like auditory gambling cues. Lighting conditions and environmental cues were not matched and differed between environments. Future work is needed to dissociate important environmental elements that modulate task performance from those irrelevant to decision-making.

With respect to our assessment of intertemporal choice in patients with TS in *Study 3* it is important to note that different discounting tasks and incentives might rely on different mechanisms. While some studies do report correlations of experiential and hypothetical tasks (Steele et al. 2019), others do not (Smits et al. 2013; Patt et al. 2021). This limitation precludes direct comparisons of the discount-rate between age groups. It is therefore not clear if

adolescents with TS show decreased discounting in a hypothetical task with much longer delay periods. Future studies should assess discounting across age groups using identical task designs and if applicable assess further constructs like future imagination or self-control. Directly examining effects of DAergic medication on these processes in TS can further shape our understanding. This is especially important in complex pathophysiologies like TS (Bloch and Leckman 2009) where the contribution of individual subprocesses, i.e. valuation, learning and others (Garcia Lorca 2019), might be differentially modulated by age, course of disease, medication and comorbidities (Hirschtritt et al. 2015). It is therefore essential to adopt a consistent longitudinal approach to further elucidate the developmental trajectory of DAergic disturbances in TS (Singer et al. 2002; Ernst et al. 1999; Buse et al. 2013; Maia and Conceição 2018) and their contributions to decision-making processes like intertemporal choice.

Finally, the results regarding the effects of DBS on preference changes are a bit preliminary. First, the exact cellular mechanism underlying DBS remain unclear (Figee et al. 2013; Lozano and Lipsman 2013). In consequence the mechanism and importance of exact location of DBS electrodes are unknown. To clearly resolve contribution of long-term stimulation of ALIC / NAcc area future studies are necessary. Ideally, those manage to recruit more participants and test for cognitive functioning, i.e. intertemporal preference and if applicable subprocesses (valuation, future imagination, self-control) prior to DBS electrode implantation. With more knowledge about the state prior to DBS, research could expand our knowledge of subtle long-term changes, due to chronic DBS.

Outlook

Understanding the continuous dimensions of human functioning sounds promising, but to approximate this promise, even just for the case of "positive valence systems" many future studies are needed. To resolve the exact role of DA I find it useful to examine, if recently developed theoretical models of DAs role in other decision-making domains or intertemporal choice in rodents hold valuable predictions for human data (Collins and Frank 2014; Mikhael and Gershman 2021; Westbrook et al. 2021).

For example, OPAL (Collins and Frank 2014) and supportive experimental findings suggest that increased baseline DA synthesis capacity or enhancements of DS DA biases decision-making to benefits in contrast to costs in a cognitive effort task (Westbrook et al. 2020; Westbrook et al. 2021). As mentioned above, our results are compatible with this view. However, future research needs to clarify if this interpretation of "benefits" and "costs" in this task generalize to "larger reward" and "delay". For example, imaging methods could further clarify if the effects of D_2 are primarily associated with DS DA. Interestingly, Westbrook et al. (2021) also find that enhancement of DA in the VS decreases the willingness to invest in cognitive effort and suggest that VS DA signals opportunity costs (average reward rate in the environment). To date OPAL´s formalism (Collins and Frank 2014) and the data on cognitive effort (Westbrook et al. 2020) primarily focus on the DS (actor in OPAL) while the role of DA in the critic (VS; state value and opportunity costs) is less formalized. Thus, future research needs to clarify if VS and DS DA modulate intertemporal preferences differently. Further, it is important to resolve how and if different DA drugs primarily modulate DA in specific brain regions. Said differently, are the observed differences of L-DOPA on intertemporal choice (Pine et al. 2010) or impulsive behavior in general (Voon 2017; Canário et al. 2019) and D_2 antagonists (Hamidovic et al. 2008; Weber et al. 2016) associated with regional effects of DA neurotransmission?

A new theory suggests that other animals do behave as if they discount rewards because they underestimate the average reward rate of the LL options (Mikhael and Gershman 2021). DA in this Bayesian-framework is suggested to moderate the effects of context (prior for the reward rate) and temporal and reward estimates (likelihood) via an effect on encoding precision [for DAs effect on encoding precision see also (FitzGerald et al. 2015)]. Here, higher DA levels, under high temporal precision, should optimize the reward rate and thus promote the selection of the LL reward (Mikhael et al. 2021). Our findings in *Study 1* are compatible with this view. However, details differ and it is unclear if one can directly translate this framework to human data. First, humans already choose smaller rewards in the absent of repeated choices, that

renders the computation of a reward rate meaningful. The tasks used in the studies of this dissertation and human tasks in general often use hypothetical rewards or pay out one randomly selected reward. Thus, participants only experience one pre- and post-reward delay. Delays in these tasks are substantially longer than in tasks used in rodents (days to weeks and months vs. seconds to a minute). Nevertheless, the view that also humans compute a reward rate would be accurate if one assumes that humans act as if they will have repeated opportunities to choose in the task (Myerson and Green 1995; Mikhael and Gershman 2021). Another possibility would be that humans assume (belief) that they will encounter a relatively low reward rate in the near future. Being speculative, even if income is relatively high (Green et al. 1996), humans could believe that the average-rate reward will be low and in consequence discount more. However, future research needs to carefully translate this framework to human intertemporal choice tasks. It would be interesting if one could bridge the gap of DAs effects on encoding precision (Mikhael and Gershman 2021), to opportunity costs (theoretically related to the prior/context) and effects of DA on benefits vs. costs of decision outcomes (Westbrook et al. 2020). For example, is there a relationship of encoding precision and future imagination? Does DA affect future imagination, i.e. the noisiness of simulations? If DA reduces simulation noise, future benefits might be more accessible to the individual and in consequence influence decision-making to a greater degree (Westbrook et al. 2020).

Another promising perspective that would complement the work in this dissertation is the focus on environmental context. Human behavior is deeply context-associated (Dixon et al. 2006; Nakahara et al. 2004; Schelp et al. 2017; Waskom et al. 2017), however the specific influence of real-life environments on intertemporal choice beyond gambling, is understudied. While, for example specific demographic factors, emotions or social factors like income (Green et al. 1996) fear (Harris 2012) and trust (Michaelson et al. 2013; Jachimowicz et al. 2017) have shown to modulate discounting, it is unclear how these interact within real-world situations. Moreover, its unknown how and if these individual demographic and emotional factors (see above) contribute to the construction of higher order beliefs (e.g. internal models of external uncertainty) and in consequence affect intertemporal preference. For example, people might believe that imagining the future does not guarantee things to play out in line with this imagination. Said differently, participants might a-priori assume that uncertainty increases with temporal distance. Finally, it would be interesting to test the influence of a-priori models/predictors of external uncertainty on specific decision components in a DDM analysis or the prior mean in a model that quantifies delay discounting primarily with respect to increasing simulation noise [in accordance with: (Gershman and Bhui 2020)].

Conclusion

This dissertation examined contextual, pharmacological and neurological effects on intertemporal decision-making. The discounting of delayed rewards has been suggested to constitute a transdiagnostic trait and therefore the functional and contextual effects that modulate this process are of valuable interest for the understanding of human functioning and mental health. In addition we examined contextual effects on RL strategies associated with goal-directed control. The results of our computational analysis confirm and extend previous research. Using sequential sampling models in combination with established discount-functions and RL rules has proven essential to provide substantial insight to drug and context effects on individual decision-components. Further, simulating from these models reproduced fundamental patterns in the data. Our results strengthen the notion that the discounting of delayed rewards is under DAergic control. Interestingly, small doses of an D_2 antagonist (enhancement of DA neurotransmission) decreased discounting, while an implicated increase of DA in the presence of addiction related cues (incentive salience) had the opposite effect (increased discounting) but improved MB control. A valuable finding for the understanding of addiction in context. Furthermore, we conclude that there are no general changes in intertemporal choice in patients with TS, a population characterized by DAergic hyperinnervation (increased DA neurotransmission). Finally, our data implicates that long-term DBS in regions under DAergic control, i.e. the NAcc area, can modulate higher cognitive function like choice preference. Taken together, intertemporal preferences change as a function of DAergic medication, environmental context and prolonged neurostimulation.

Future research should attempt to further disentangle the subprocesses that contribute to choice impulsivity. For example, how does DA (pharmacological or contextual manipulation) modulate valuation, self-control or future imagination individually? Under which contextual- or regional constraints does DA modulate encoding precision, opportunity costs or cause subjects to focus on the benefits rather than the costs of decision outcomes? In addition, it is proposed to aim at developing a unified model of DAergic action. For this purpose, it might be useful to derive testable hypotheses from theoretical frameworks of DA neurotransmission in other decision-making domains. Resolving these issues will further contribute to the understanding of an important aspect of human functioning, namely trading off the future against the presence, and might elevate the potential for intertemporal decision-making to become an important diagnostic tool and potential treatment target that benefits the individual.

Publication bibliography

Achterberg, Michelle; Peper, Jiska S.; van Duijvenvoorde, Anna C. K.; Mandl, René C. W.; Crone, Eveline A. (2016): Frontostriatal White Matter Integrity Predicts Development of Delay of Gratification: A Longitudinal Study. In *J. Neurosci.* 36 (6), pp. 1954–1961. DOI: 10.1523/JNEUROSCI.3459-15.2016.

Aiello, M.; Terenzi, D.; Furlanis, G.; Catalan, M.; Manganotti, P.; Eleopra, R. et al. (2019): Deep brain stimulation of the subthalamic nucleus and the temporal discounting of primary and secondary rewards. In *J Neurol* 266 (5), pp. 1113–1119. DOI: 10.1007/s00415-019-09240-0.

Ainslie, G. (1975): Specious reward: a behavioral theory of impulsiveness and impulse control. In *Psychological bulletin* 82 (4), pp. 463–496. DOI: 10.1037/h0076860.

Ainslie, G. W. (1974): Impulse control in pigeons. In *Journal of the experimental analysis of behavior* 21 (3), pp. 485–489. DOI: 10.1901/jeab.1974.21-485.

Aizman, O.; Brismar, H.; Uhlén, P.; Zettergren, E.; Levey, A. I.; Forssberg, H. et al. (2000): Anatomical and physiological evidence for D1 and D2 dopamine receptor colocalization in neostriatal neurons. In *Nat Neurosci* 3 (3), pp. 226–230. DOI: 10.1038/72929.

Alexander, Garrett E.; Crutcher, Michael D.; DeLong, Mahlon R. (1990): Chapter 6 Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. In Harry B. M. Uylings (Ed.): The prefrontal cortex. Its structure, function, and pathology ; proceedings of the 16th International Summer School of Brain Research, held at the Royal Tropical Institute and the Royal Netherlands Academy of Sciences Amsterdam, the Netherlands, from 28 August to 1 September 1989, vol. 85. Amsterdam, New York: Elsevier (Progress in Brain Research, v. 85), pp. 119–146.

Alloy, Lauren B.; Olino, Thomas; Freed, Rachel D.; Nusslock, Robin (2016): Role of Reward Sensitivity and Processing in Major Depressive and Bipolar Spectrum Disorders. In *Behavior therapy* 47 (5), pp. 600–621. DOI: 10.1016/j.beth.2016.02.014.

Alonso, Pino; Cuadras, Daniel; Gabriëls, Loes; Denys, Damiaan; Goodman, Wayne; Greenberg, Ben D. et al. (2015): Deep Brain Stimulation for Obsessive-Compulsive Disorder: A Meta-Analysis of Treatment Outcome and Predictors of Response. In *PloS one* 10 (7), e0133591. DOI: 10.1371/journal.pone.0133591.

Alwin, Duane Francis (2006): Margins of error. A study of reliability in survey measurement. Hoboken, NJ: Wiley (Wiley series in survey methodology).

Amlung, M.; Petker, T.; Jackson, J.; Balodis, I.; MacKillop, J. (2016): Steep discounting of delayed monetary and food rewards in obesity: a meta-analysis. In *Psychological medicine* 46 (11), pp. 2423–2434. DOI: 10.1017/S0033291716000866.

Amlung, Michael; Marsden, Emma; Holshausen, Katherine; Morris, Vanessa; Patel, Herry; Vedelago, Lana et al. (2019): Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. In *JAMA psychiatry* 76 (11), pp. 1176–1186. DOI: 10.1001/jamapsychiatry.2019.2102.

Anandakumar, Jeya; Mills, Kathryn L.; Earl, Eric A.; Irwin, Lourdes; Miranda-Dominguez, Oscar; Demeter, Damion V. et al. (2018): Individual differences in functional brain connectivity predict temporal discounting preference in the transition to adolescence. In *Developmental cognitive neuroscience* 34, pp. 101–113. DOI: 10.1016/j.dcn.2018.07.003.

Anokhin, Andrey P.; Golosheykin, Simon; Grant, Julia D.; Heath, Andrew C. (2011): Heritability of delay discounting in adolescence: a longitudinal twin study. In *Behavior genetics* 41 (2), pp. 175–183. DOI: 10.1007/s10519-010-9384-7.

Anokhin, Andrey P.; Grant, Julia D.; Mulligan, Richard C.; Heath, Andrew C. (2015): The genetics of impulsivity: evidence for the heritability of delay discounting. In *Biological psychiatry* 77 (10), pp. 887–894. DOI: 10.1016/j.biopsych.2014.10.022.

Anselme, Patrick; Robinson, Mike J. F. (2013): What motivates gambling behavior? Insight into dopamine's role. In *Front. Behav. Neurosci.* 7, p. 182. DOI: 10.3389/fnbeh.2013.00182.

Antons, Stephanie; Brand, Matthias; Potenza, Marc N. (2020): Neurobiology of cue-reactivity, craving, and inhibitory control in non-substance addictive behaviors. In *Journal of the neurological sciences* 415, p. 116952. DOI: 10.1016/j.jns.2020.116952.

Arrondo, Gonzalo; Aznárez-Sanado, Maite; Fernández-Seara, Maria A.; Goñi, Joaquín; Loayza, Francis R.; Salamon-Klobut, Ewa et al. (2015): Dopaminergic modulation of the trade-off between probability and time in economic decision-making. In *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 25 (6), pp. 817–827. DOI: 10.1016/j.euroneuro.2015.02.011.

Athamneh, Liqa N.; Stein, Jeffrey S.; Bickel, Warren K. (2017): Will delay discounting predict intention to quit smoking? In *Experimental and Clinical Psychopharmacology* 25 (4), pp. 273–280. DOI: 10.1037/pha0000129.

Ayano, Getinet (2016): Dopamine: Receptors, Functions, Synthesis, Pathways, Locations and Mental Disorders: Review of Literatures. In *J Ment Disord Treat* 2 (2). DOI: 10.4172/2471-271x.1000120.

Baldermann, Juan Carlos; Schüller, Thomas; Huys, Daniel; Becker, Ingrid; Timmermann, Lars; Jessen, Frank et al. (2016): Deep Brain Stimulation for Tourette-Syndrome: A Systematic Review and Meta-Analysis. In *Brain Stimulation* 9 (2), pp. 296–304. DOI: 10.1016/j.brs.2015.11.005.

Baldwin, Scott A.; Fellingham, Gilbert W. (2013): Bayesian methods for the analysis of small sample multilevel data with a complex variance structure. In *Psychological methods* 18 (2), pp. 151–164. DOI: 10.1037/a0030642.

Ballard, Ian C.; Kim, Bokyung; Liatsis, Anthony; Aydogan, Gökhan; Cohen, Jonathan D.; McClure, Samuel M. (2017): More Is Meaningful. The Magnitude Effect in Intertemporal Choice Depends on Self-Control. In *Psychological science* 28 (10), pp. 1443–1454. DOI: 10.1177/0956797617711455.

Balleine, Bernard W.; O'Doherty, John P. (2010): Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. In *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 35 (1), pp. 48–69. DOI: 10.1038/npp.2009.131.

Barry J Everitt; Trevor W Robbins (2005): Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. In *Nat Neurosci* 8 (11), pp. 1481–1489. DOI: 10.1038/nn1579.

Bartra, Oscar; McGuire, Joseph T.; Kable, Joseph W. (2013): The valuation system. A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. In *NeuroImage* 76, pp. 412–427. DOI: 10.1016/j.neuroimage.2013.02.063.

Basten, Ulrike; Biele, Guido; Heekeren, Hauke R.; Fiebach, Christian J. (2010): How the brain integrates costs and benefits during decision making. In *PNAS* 107 (50), pp. 21767–21772. DOI: 10.1073/pnas.0908104107.

Bayer, Hannah M.; Glimcher, Paul W. (2005): Midbrain dopamine neurons encode a quantitative reward prediction error signal. In *Neuron* 47 (1), pp. 129–141. DOI: 10.1016/j.neuron.2005.05.020.

Beaulieu, Jean-Martin; Espinoza, Stefano; Gainetdinov, Raul R. (2015): Dopamine receptors - IUPHAR Review 13. In *British journal of pharmacology* 172 (1), pp. 1–23. DOI: 10.1111/bph.12906.

Beaulieu, Jean-Martin; Gainetdinov, Raul R. (2011): The physiology, signaling, and pharmacology of dopamine receptors. In *Pharmacol Rev* 63 (1), pp. 182–217. DOI: 10.1124/pr.110.002642.

Beierholm, Ulrik; Guitart-Masip, Marc; Economides, Marcos; Chowdhury, Rumana; Düzel, Emrah; Dolan, Ray; Dayan, Peter (2013): Dopamine modulates reward-related vigor. In *Neuropsychopharmacology* 38 (8), pp. 1495–1503. DOI: 10.1038/npp.2013.48.

Bello, Estefanía P.; Mateo, Yolanda; Gelman, Diego M.; Noaín, Daniela; Shin, Jung H.; Low, Malcolm J. et al. (2011): Cocaine supersensitivity and enhanced motivation for reward in mice lacking dopamine D2 autoreceptors. In *Nat Neurosci* 14 (8), pp. 1033–1038. DOI: 10.1038/nn.2862.

Benoit, Roland G.; Berkers, Ruud M. W. J.; Paulus, Philipp C. (2018): An adaptive function of mental time travel: Motivating farsighted decisions. In *The Behavioral and brain sciences* 41, e3. DOI: 10.1017/S0140525X1700125X.

Benoit-Marand, Marianne; Ballion, Bérangère; Borrelli, Emiliana; Boraud, Thomas; Gonon, Francois (2011): Inhibition of dopamine uptake by D2 antagonists: an in vivo study. In *Journal of neurochemistry* 116 (3), pp. 449–458. DOI: 10.1111/j.1471-4159.2010.07125.x.

Berke, Joshua D. (2018): What does dopamine mean? In *Nat Neurosci* 21 (6), pp. 787–793. DOI: 10.1038/s41593-018-0152-y.

Bermudez, Maria A.; Göbel, Carl; Schultz, Wolfram (2012): Sensitivity to temporal reward structure in amygdala neurons. In *Current biology : CB* 22 (19), pp. 1839–1844. DOI: 10.1016/j.cub.2012.07.062.

Bernheimer, H.; Birkmayer, W.; Hornykiewicz, O.; Jellinger, K.; Seitelberger, F. (1973): Brain dopamine and the syndromes of Parkinson and Huntington Clinical, morphological and neurochemical correlations. In *Journal of the neurological sciences* 20 (4), pp. 415–455. DOI: 10.1016/0022-510X(73)90175-5.

Berridge, K. (2000): Measuring hedonic impact in animals and infants: microstructure of affective taste reactivity patterns. In *Neuroscience & Biobehavioral Reviews* 24 (2), pp. 173–198. DOI: 10.1016/S0149-7634(99)00072-X.

Berridge, K. C. (2016): Incentive Motivation and Incentive Salience☆. In John Stein (Ed.): Reference module in neuroscience and biobehavioral psychology. Place of publication not identified: Elsevier. Available online at https://www.sciencedirect.com/science/article/pii/B9780128093245003424.

Berridge, Kent C. (2012): From prediction error to incentive salience: mesolimbic computation of reward motivation. In *The European journal of neuroscience* 35 (7), pp. 1124–1143. DOI: 10.1111/j.1460-9568.2012.07990.x.

Berridge, Kent C.; Robinson, Terry E. (1998): What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? In *Brain Research Reviews* 28 (3), pp. 309–369. DOI: 10.1016/S0165-0173(98)00019-8.

Berridge, Kent C.; Robinson, Terry E. (2016): Liking, wanting, and the incentive-sensitization theory of addiction. In *The American psychologist* 71 (8), pp. 670–679. DOI: 10.1037/amp0000059.

Berridge, Kent C.; Venier, Isabel L.; Robinson, Terry E. (1989): Taste reactivity analysis of 6-hydroxydopamine-induced aphagia: Implications for arousal and anhedonia hypotheses of dopamine function. In *Behavioral Neuroscience* 103 (1), pp. 36–45. DOI: 10.1037/0735-7044.103.1.36.

Bichot, Narcisse P.; Rao, S.Chenchal; Schall, Jeffrey D. (2001): Continuous processing in macaque frontal cortex during visual search. In *Neuropsychologia* 39 (9), pp. 972–982. DOI: 10.1016/S0028-3932(01)00022-7.

Bickel, Warren K.; Athamneh, Liqa N.; Basso, Julia C.; Mellis, Alexandra M.; DeHart, William B.; Craft, William H.; Pope, Derek (2019): Excessive discounting of delayed reinforcers as a trans-disease process: Update on the state of the science. In *Current opinion in psychology* 30, pp. 59–64. DOI: 10.1016/j.copsyc.2019.01.005.

Bickel, Warren K.; Jarmolowicz, David P.; Mueller, E. Terry; Koffarnus, Mikhail N.; Gatchalian, Kirstin M. (2012): Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: emerging evidence. In *Pharmacology & therapeutics* 134 (3), pp. 287–297. DOI: 10.1016/j.pharmthera.2012.02.004.

Bickel, Warren K.; Koffarnus, Mikhail N.; Moody, Lara; Wilson, A. George (2014): The behavioral- and neuro-economic process of temporal discounting: A candidate behavioral marker of addiction. In *Neuropharmacology* 76 Pt B, pp. 518–527. DOI: 10.1016/j.neuropharm.2013.06.013.

Björklund, Anders; Dunnett, Stephen B. (2007): Dopamine neuron systems in the brain: an update. In *Trends in neurosciences* 30 (5), pp. 194–202. DOI: 10.1016/j.tins.2007.03.006.

Blaszczynski, Alex; Nower, Lia (2002): A pathways model of problem and pathological gambling. In *Addiction (Abingdon, England)* 97 (5), pp. 487–499. DOI: 10.1046/j.1360-0443.2002.00015.x.

Bloch, Michael H.; Leckman, James F. (2009): Clinical course of Tourette syndrome. In *Journal of Psychosomatic Research* 67 (6), pp. 497–501. DOI: 10.1016/j.jpsychores.2009.09.002.

Böhm-Bawerk (1891): Capital and Interest.

Braem, Senne; King, Joseph A.; Korb, Franziska M.; Krebs, Ruth M.; Notebaert, Wim; Egner, Tobias (2017): The Role of Anterior Cingulate Cortex in the Affective Evaluation of Conflict. In *Journal of Cognitive Neuroscience* 29 (1), pp. 137–149. DOI: 10.1162/jocn_a_01023.

Bromberg, Uli; Lobatcheva, Maria; Peters, Jan (2017): Episodic future thinking reduces temporal discounting in healthy adolescents. In *PloS one* 12 (11), e0188079. DOI: 10.1371/journal.pone.0188079.

Brown, Robert (1828): XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and … Available online at https://www.tandfonline.com/doi/pdf/10.1080/14786442808674769?casa_token=agrmy1gnilaaaaaa:zq5uh8s2tqcnhg3izncevng5iw8gsf0w2icxyux6on57o00ljw0wia2eejv_ecxptpkvsfowvftpng.

Buckholtz, Joshua W.; Treadway, Michael T.; Cowan, Ronald L.; Woodward, Neil D.; Li, Rui; Ansari, M. Sib et al. (2010): Dopaminergic network differences in human impulsivity. In *Science* 329 (5991), p. 532. DOI: 10.1126/science.1185778.

Bulley, Adam; Henry, Julie; Suddendorf, Thomas (2016): Prospection and the Present Moment: The Role of Episodic Foresight in Intertemporal Choices between Immediate and Delayed Rewards. In *Review of General Psychology* 20 (1), pp. 29–47. DOI: 10.1037/gpr0000061.

Buse, Judith; Schoenefeld, Katja; Münchau, Alexander; Roessner, Veit (2013): Neuromodulation in Tourette syndrome: dopamine and beyond. In *Neuroscience and biobehavioral reviews* 37 (6), pp. 1069–1084. DOI: 10.1016/j.neubiorev.2012.10.004.

Caballero, Adriana; Granberg, Rachel; Tseng, Kuei Y. (2016): Mechanisms contributing to prefrontal cortex maturation during adolescence. In *Neuroscience and biobehavioral reviews* 70, pp. 4–12. DOI: 10.1016/j.neubiorev.2016.05.013.

Cagniard, Barbara; Beeler, Jeff A.; Britt, Jonathan P.; McGehee, Daniel S.; Marinelli, Michela; Zhuang, Xiaoxi (2006): Dopamine scales performance in the absence of new learning. In *Neuron* 51 (5), pp. 541–547. DOI: 10.1016/j.neuron.2006.07.026.

Canário, Nádia; Sousa, Mário; Moreira, Fradique; Duarte, Isabel Catarina; Oliveira, Francisco; Januário, Cristina; Castelo-Branco, Miguel (2019): Impulsivity across reactive, proactive and cognitive domains in Parkinson's disease on dopaminergic medication: Evidence for multiple domain impairment. In *PloS one* 14 (2), e0210880. DOI: 10.1371/journal.pone.0210880.

Cannon, Claire Matson; Palmiter, Richard D. (2003): Reward without Dopamine. In *J. Neurosci.* 23 (34), pp. 10827–10831. DOI: 10.1523/JNEUROSCI.23-34-10827.2003.

Carlisi, C. O.; Norman, L.; Murphy, C. M.; Christakou, A.; Chantiluke, K.; Giampietro, V. et al. (2017): Comparison of neural substrates of temporal discounting between youth with autism spectrum disorder and with obsessive-compulsive disorder. In *Psychological medicine* 47 (14), pp. 2513–2527. DOI: 10.1017/s0033291717001088.

Casey, B. J.; Oliveri, Mary Ellen; Insel, Thomas (2014): A neurodevelopmental perspective on the research domain criteria (RDoC) framework. In *Biological psychiatry* 76 (5), pp. 350–353. DOI: 10.1016/j.biopsych.2014.01.006.

Chamberlain, Samuel R.; Menzies, Lara; Hampshire, Adam; Suckling, John; Fineberg, Naomi A.; Del Campo, Natalia et al. (2008): Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. In *Science* 321 (5887), pp. 421–422. DOI: 10.1126/science.1154433.

Chen, Yin-Ching I.; Choi, Ji-Kyung; Andersen, Susan L.; Rosen, Bruce R.; Jenkins, Bruce G. (2005): Mapping dopamine D2/D3 receptor function using pharmacological magnetic resonance imaging. In *Psychopharmacology* 180 (4), pp. 705–715. DOI: 10.1007/s00213-004-2034-0.

Chesson, Harrell W.; Leichliter, Jami S.; Zimet, Gregory D.; Rosenthal, Susan L.; Bernstein, David I.; Fife, Kenneth H. (2006): Discount rates and risky sexual behaviors among teenagers and young adults. In *J Risk Uncertainty* 32 (3), pp. 217–230. DOI: 10.1007/s11166-006-9520-1.

Ciliax, Brian J.; Drash, Genny W.; Staley, Julie K.; Haber, Sharon; Mobley, Catherine J.; Miller, Gary W. et al. (1999): Immunocytochemical localization of the dopamine transporter in human brain. In *The Journal of comparative neurology* 409 (1), pp. 38–56. DOI: 10.1002/(SICI)1096-9861(19990621)409:1<38::AID-CNE4>3.0.CO;2-1.

Cleary, Daniel R.; Ozpinar, Alp; Raslan, Ahmed M.; Ko, Andrew L. (2015): Deep brain stimulation for psychiatric disorders: where we are now. In *FOC* 38 (6), E2. DOI: 10.3171/2015.3.FOCUS1546.

Clithero, John A.; Rangel, Antonio (2014): Informatic parcellation of the network involved in the computation of subjective value. In *Social cognitive and affective neuroscience* 9 (9), pp. 1289–1302. DOI: 10.1093/scan/nst106.

Clos, Mareike; Bunzeck, Nico; Sommer, Tobias (2019): Dopamine Enhances Item Novelty Detection via Hippocampal and Associative Recall via Left Lateral Prefrontal Cortex Mechanisms. In *Journal of Neuroscience* 39 (40), pp. 7920–7933. DOI: 10.1523/JNEUROSCI.0495-19.2019.

Coffey, Barbara J.; Biederman, Joseph; Geller, Daniel; Frazier, Jean; Spencer, Tom; Doyle, Robert et al. (2004): Reexamining Tic persistence and Tic-associated impairment in Tourette's Disorder: findings

from a naturalistic follow-up study. In *Journal of Nervous & Mental Disease* 192 (11), pp. 776–780. DOI: 10.1097/01.nmd.0000144696.14555.c4.

Collins, Anne G. E.; Frank, Michael J. (2014): Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. In *Psychological Review* 121 (3), pp. 337–366. DOI: 10.1037/a0037015.

Colombo, Matteo (2014): Deep and beautiful. The reward prediction error hypothesis of dopamine. In *Studies in history and philosophy of biological and biomedical sciences* 45, pp. 57–67. DOI: 10.1016/j.shpsc.2013.10.006.

Cools, Roshan (2008): Role of dopamine in the motivational and cognitive control of behavior. In *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 14 (4), pp. 381–395. DOI: 10.1177/1073858408317009.

Cools, Roshan; D'Esposito, Mark (2011): Inverted-U-shaped dopamine actions on human working memory and cognitive control. In *Biological psychiatry* 69 (12), e113-25. DOI: 10.1016/j.biopsych.2011.03.028.

Costello, C. G. (1972): Depression: Loss of reinforcers of loss of reinforcer effectiveness? In *Behavior therapy* 3 (2), pp. 240–247. DOI: 10.1016/S0005-7894(72)80084-4.

Coyle, J. T.; Snyder, S. H. (1969): Antiparkinsonian drugs: inhibition of dopamine uptake in the corpus striatum as a possible mechanism of action. In *Science (New York, N.Y.)* 166 (3907), pp. 899–901. DOI: 10.1126/science.166.3907.899.

Crone, Eveline A.; Steinbeis, Nikolaus (2017): Neural Perspectives on Cognitive Control Development during Childhood and Adolescence. In *Trends in cognitive sciences* 21 (3), pp. 205–215. DOI: 10.1016/j.tics.2017.01.003.

Dabney, Will; Kurth-Nelson, Zeb; Uchida, Naoshige; Starkweather, Clara Kwon; Hassabis, Demis; Munos, Rémi; Botvinick, Matthew (2020): A distributional code for value in dopamine-based reinforcement learning. In *Nature* 577 (7792), pp. 671–675. DOI: 10.1038/s41586-019-1924-6.

Dale, Gerard; Rock, Adam J.; Clark, Gavin I. (2019): Cue-Reactive Imagery Mediates the Relationships of Reward Responsiveness with Both Cue-Reactive Urge to Gamble and Positive Affect in Poker-Machine Gamblers. In *Journal of gambling studies*. DOI: 10.1007/s10899-019-09864-x.

Daw (2011): Model-Based Influences on Humans' Choices and Striatal Prediction Errors. In *Neuron* 69 (6), pp. 1204–1215. DOI: 10.1016/j.neuron.2011.02.027.

Day, Jeremy J.; Jones, Joshua L.; Wightman, R. Mark; Carelli, Regina M. (2010): Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. In *Biological psychiatry* 68 (3), pp. 306–309. DOI: 10.1016/j.biopsych.2010.03.026.

Dayan, Peter; Niv, Yael (2008): Reinforcement learning: the good, the bad and the ugly. In *Current opinion in neurobiology* 18 (2), pp. 185–196. DOI: 10.1016/j.conb.2008.08.003.

Denys, Damiaan; Mantione, Mariska; Figee, Martijn; van den Munckhof, Pepijn; Koerselman, Frank; Westenberg, Herman et al. (2010): Deep brain stimulation of the nucleus accumbens for treatment-refractory obsessive-compulsive disorder. In *Archives of general psychiatry* 67 (10), pp. 1061–1068. DOI: 10.1001/archgenpsychiatry.2010.122.

Denys, Damiaan; Vries, Froukje de; Cath, Danielle; Figee, Martijn; Vulink, Nienke; Veltman, Dick J. et al. (2013): Dopaminergic activity in Tourette syndrome and obsessive-compulsive disorder. In *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 23 (11), pp. 1423–1431. DOI: 10.1016/j.euroneuro.2013.05.012.

Dixon, Mark R.; Jacobs, Eric A.; Sanders, Scott (2006): CONTEXTUAL CONTROL OF DELAY DISCOUNTING BY PATHOLOGICAL GAMBLERS. In *Journal of applied behavior analysis* 39 (4), pp. 413–422. DOI: 10.1901/jaba.2006.173-05.

Dixon, Mark R.; Marley, Janice; Jacobs, Eric A. (2003): Delay discounting by pathological gamblers. In *Journal of applied behavior analysis* 36 (4), pp. 449–458. DOI: 10.1901/jaba.2003.36-449.

Doll, Bradley B.; Simon, Dylan A.; Daw, Nathaniel D. (2012): The ubiquity of model-based reinforcement learning. In *Current opinion in neurobiology* 22 (6), pp. 1075–1081. DOI: 10.1016/j.conb.2012.08.003.

Doucet, G.; Descarries, L.; Garcia, S. (1986): Quantification of the dopamine innervation in adult rat neostriatum. In *Neuroscience* 19 (2), pp. 427–445. DOI: 10.1016/0306-4522(86)90272-1.

Draganski, Bogdan; Kherif, Ferath; Klöppel, Stefan; Cook, Philip A.; Alexander, Daniel C.; Parker, Geoff J. M. et al. (2008): Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. In *J. Neurosci.* 28 (28), pp. 7143–7152. DOI: 10.1523/JNEUROSCI.1486-08.2008.

Duszkiewicz, Adrian J.; McNamara, Colin G.; Takeuchi, Tomonori; Genzel, Lisa (2019): Novelty and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. In *Trends in neurosciences* 42 (2), pp. 102–114. DOI: 10.1016/j.tins.2018.10.002.

Dutcher, Janine M.; Creswell, J. David (2018): The role of brain reward pathways in stress resilience and health. In *Neuroscience and biobehavioral reviews* 95, pp. 559–567. DOI: 10.1016/j.neubiorev.2018.10.014.

Dwyer, Jennifer B. (2018): A Developmental Perspective of Dopaminergic Dysfunction in Tourette Syndrome. In *Biological psychiatry* 84 (5), e33-e35. DOI: 10.1016/j.biopsych.2018.07.008.

Ebitz, R. Becket; Smith, Elliot H.; Horga, Guillermo; Schevon, Catherine A.; Yates, Mark J.; McKhann, Guy M. et al. (2020): Human dorsal anterior cingulate neurons signal conflict by amplifying task-relevant information. In *bioRxiv*, 2020.03.14.991745. DOI: 10.1101/2020.03.14.991745.

Einstein, Albert (1905): On the movement of small particles suspended in stationary liquids required by the molecularkinetic theory of heat. Annalen der Physik. Available online at https://www2.stat.duke.edu/courses/spring12/sta357/refs/eins1905en.pdf.

Enkavi, A. Zeynep; Eisenberg, Ian W.; Bissett, Patrick G.; Mazza, Gina L.; MacKinnon, David P.; Marsch, Lisa A.; Poldrack, Russell A. (2019): Large-scale analysis of test-retest reliabilities of self-regulation measures. In *PNAS* 116 (12), pp. 5472–5477. DOI: 10.1073/pnas.1818430116.

Ernst, M.; Zametkin, A. J.; Jons, P. H.; Matochik, J. A.; Pascualvaca, D.; Cohen, R. M. (1999): High presynaptic dopaminergic activity in children with Tourette's disorder. In *Journal of the American Academy of Child & Adolescent Psychiatry* 38 (1), pp. 86–94. DOI: 10.1097/00004583-199901000-00024.

Farrell, Simon; Lewandowsky, Stephan (2018): Computational modeling of cognition and behavior. Cambridge: Cambridge University Press.

Feller, William (2003): An introduction to probability theory and its applications. 3. ed., rev. print., [Nachdr.]. S.l.: Wiley (Wiley series in probability and mathematical statistics).

Ferreri, Laura; Mas-Herrero, Ernest; Zatorre, Robert J.; Ripollés, Pablo; Gomez-Andres, Alba; Alicart, Helena et al. (2019): Dopamine modulates the reward experiences elicited by music. In *Proceedings of the National Academy of Sciences* 116 (9), pp. 3793–3798. DOI: 10.1073/pnas.1811878116.

Figee, Martijn; Luigjes, Judy; Smolders, Ruud; Valencia-Alfonso, Carlos-Eduardo; van Wingen, Guido; Kwaasteniet, Bart de et al. (2013): Deep brain stimulation restores frontostriatal network activity in obsessive-compulsive disorder. In *Nature neuroscience* 16 (4), pp. 386–387. DOI: 10.1038/nn.3344.

Figner, Bernd; Knoch, Daria; Johnson, Eric J.; Krosch, Amy R.; Lisanby, Sarah H.; Fehr, Ernst; Weber, Elke U. (2010): Lateral prefrontal cortex and self-control in intertemporal choice. In *Nature neuroscience* 13 (5), pp. 538–539. DOI: 10.1038/nn.2516.

FitzGerald, Thomas H. B.; Dolan, Raymond J.; Friston, Karl (2015): Dopamine, reward learning, and active inference. In *Front. Comput. Neurosci.* 9, p. 136. DOI: 10.3389/fncom.2015.00136.

FitzGerald, Thomas H. B.; Seymour, Ben; Dolan, Raymond J. (2009): The role of human orbitofrontal cortex in value comparison for incommensurable objects. In *J. Neurosci.* 29 (26), pp. 8388–8395. DOI: 10.1523/JNEUROSCI.0717-09.2009.

Flagel, Shelly B.; Akil, Huda; Robinson, Terry E. (2009): Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. In *Neuropharmacology* 56 Suppl 1, pp. 139–148. DOI: 10.1016/j.neuropharm.2008.06.027.

Foerde, Karin; Figner, Bernd; Doll, Bradley B.; Woyke, Isabel C.; Braun, Erin Kendall; Weber, Elke U.; Shohamy, Daphna (2016): Dopamine Modulation of Intertemporal Decision-making: Evidence from Parkinson Disease. In *Journal of Cognitive Neuroscience* 28 (5), pp. 657–667. DOI: 10.1162/jocn_a_00929.

Fontanesi, Laura; Gluth, Sebastian; Spektor, Mikhail S.; Rieskamp, Jörg (2019): A reinforcement learning diffusion decision model for value-based decisions. In *Psychonomic Bulletin & Review* 26 (4), pp. 1099–1121. DOI: 10.3758/s13423-018-1554-2.

Fontenelle, Leonardo F.; Mendlowicz, Mauro V.; Versiani, Marcio (2005): Impulse control disorders in patients with obsessive-compulsive disorder. In *Psychiatry and Clinical Neurosciences* 59 (1), pp. 30–37. DOI: 10.1111/j.1440-1819.2005.01328.x.

Ford, C. P. (2014): The role of D2-autoreceptors in regulating dopamine neuron activity and transmission. In *Neuroscience* 282, pp. 13–22. DOI: 10.1016/j.neuroscience.2014.01.025.

Forstmann, B. U.; Ratcliff, R.; Wagenmakers, E-J (2016): Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. In *Annual review of psychology* 67, pp. 641–666. DOI: 10.1146/annurev-psych-122414-033645.

Frank, Michael J. (2005): Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. In *Journal of Cognitive Neuroscience* 17 (1), pp. 51–72. DOI: 10.1162/0898929052880093.

Frank, Michael J.; O'Reilly, Randall C. (2006): A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. In *Behavioral Neuroscience* 120 (3), pp. 497–517. DOI: 10.1037/0735-7044.120.3.497.

Freund, Nadja; Jordan, Chloe J.; Lukkes, Jodi L.; Norman, Kevin J.; Andersen, Susan L. (2019): Juvenile exposure to methylphenidate and guanfacine in rats: effects on early delay discounting and later cocaine-taking behavior. In *Psychopharmacology* 236 (2), pp. 685–698. DOI: 10.1007/s00213-018-5096-0.

Friedel, Robert O. (2004): Dopamine Dysfunction in Borderline Personality Disorder: A Hypothesis. In *Neuropsychopharmacol* 29 (6), pp. 1029–1039. DOI: 10.1038/sj.npp.1300424.

Friston, Karl (2010): The free-energy principle: a unified brain theory? In *Nature Reviews Neuroscience* 11 (2), pp. 127–138. DOI: 10.1038/nrn2787.

Friston, Karl (2012): The history of the future of the Bayesian brain. In *NeuroImage* 62 (2), pp. 1230–1233. DOI: 10.1016/j.neuroimage.2011.10.004.

Garcia Lorca, Federico (2019): Dysregulation of the Reward and Learning Systems in Tourette Syndrome. In *JAMA Neurology. American Medical Association* 76, p. 1124.

Genauck, Alexander; Andrejevic, Milan; Brehm, Katharina; Matthis, Caroline; Heinz, Andreas; Weinreich, André et al. (2020): Cue-induced effects on decision-making distinguish subjects with gambling disorder from healthy controls. In *Addiction biology* 25 (6), e12841. DOI: 10.1111/adb.12841.

Gershman, Samuel J.; Bhui, Rahul (2020): Rationally inattentive intertemporal choice. In *Nature communications* 11 (1), p. 3365. DOI: 10.1038/s41467-020-16852-y.

Gilbert, Daniel T.; Wilson, Timothy D. (2007): Prospection: experiencing the future. In *Science* 317 (5843), pp. 1351–1354. DOI: 10.1126/science.1144161.

Gillan, Claire M.; Kalanthroff, Eyal; Evans, Michael; Weingarden, Hilary M.; Jacoby, Ryan J.; Gershkovich, Marina et al. (2020): Comparison of the Association Between Goal-Directed Planning and Self-reported Compulsivity vs Obsessive-Compulsive Disorder Diagnosis. In *JAMA psychiatry* 77 (1), pp. 77–85. DOI: 10.1001/jamapsychiatry.2019.2998.

Gillan, Claire M.; Kosinski, Michal; Whelan, Robert; Phelps, Elizabeth A.; Daw, Nathaniel D. (2016): Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. In *eLife* 5. DOI: 10.7554/eLife.11305.

Giros, Bruno; Caron, Marc G. (1993): Molecular characterization of the dopamine transporter. In *Trends in Pharmacological Sciences* 14 (2), pp. 43–49. DOI: 10.1016/0165-6147(93)90029-J.

Gläscher, Jan; Daw, Nathaniel; Dayan, Peter; O'Doherty, John P. (2010): States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. In *Neuron* 66 (4), pp. 585–595. DOI: 10.1016/j.neuron.2010.04.016.

Grace, A. A.; Bunney, B. S. (1984): The control of firing pattern in nigral dopamine neurons: single spike firing. In *J. Neurosci.* 4 (11), pp. 2866–2876. DOI: 10.1523/JNEUROSCI.04-11-02866.1984.

Green, Leonard; Myerson, Joel (2004): A discounting framework for choice with delayed and probabilistic rewards. In *Psychological bulletin* 130 (5), pp. 769–792. DOI: 10.1037/0033-2909.130.5.769.

Green, Leonard; Myerson, Joel; Lichtman, David; Rosen, Suzanne; Fry, Astrid (1996): Temporal discounting in choice between delayed rewards: The role of age and income. In *Psychology and Aging* 11 (1), pp. 79–84. DOI: 10.1037/0882-7974.11.1.79.

Grimm, Pamela (2010): Social Desirability Bias: American Cancer Society (Wiley International Encyclopedia of Marketing).

Groth, Camilla; Mol Debes, Nanette; Rask, Charlotte Ulrikka; Lange, Theis; Skov, Liselotte (2017): Course of Tourette Syndrome and Comorbidities in a Large Prospective Clinical Study. In *Journal of the American Academy of Child and Adolescent Psychiatry* 56 (4), pp. 304–312. DOI: 10.1016/j.jaac.2017.01.010.

Grüne-Yanoff, Till (2015): Models of Temporal Discounting 1937-2000: An Interdisciplinary Exchange between Economics and Psychology. In *Science in context* 28 (4), pp. 675–713. DOI: 10.1017/S0269889715000307.

Guitart-Masip, Marc; Beierholm, Ulrik R.; Dolan, Raymond; Duzel, Emrah; Dayan, Peter (2011): Vigor in the face of fluctuating rates of reward: an experimental examination. In *Journal of Cognitive Neuroscience* 23 (12), pp. 3933–3938. DOI: 10.1162/jocn_a_00090.

Haber, Suzanne N.; Knutson, Brian (2010): The reward circuit: linking primate anatomy and human imaging. In *Neuropsychopharmacology* 35 (1), pp. 4–26. DOI: 10.1038/npp.2009.129.

Hamidovic, Ajna; Kang, Un Jung; Wit, Harriet de (2008): Effects of low to moderate acute doses of pramipexole on impulsivity and cognition in healthy volunteers. In *Journal of clinical psychopharmacology* 28 (1), pp. 45–51. DOI: 10.1097/jcp.0b013e3181602fab.

Hampton, William H.; Asadi, Nima; Olson, Ingrid R. (2018): Good Things for Those Who Wait: Predictive Modeling Highlights Importance of Delay Discounting for Income Attainment. In *Front. Psychol.* 9, p. 1545. DOI: 10.3389/fpsyg.2018.01545.

Hare, Todd A.; Camerer, Colin F.; Rangel, Antonio (2009): Self-control in decision-making involves modulation of the vmPFC valuation system. In *Science (New York, N.Y.)* 324 (5927), pp. 646–648. DOI: 10.1126/science.1168450.

Hare, Todd A.; Hakimi, Shabnam; Rangel, Antonio (2014): Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. In *Frontiers in Neuroscience* 8, p. 50. DOI: 10.3389/fnins.2014.00050.

Hariri, Ahmad R.; Brown, Sarah M.; Williamson, Douglas E.; Flory, Janine D.; Wit, Harriet de; Manuck, Stephen B. (2006): Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26 (51), pp. 13213–13217. DOI: 10.1523/JNEUROSCI.3446-06.2006.

Harris, Alison; Hare, Todd; Rangel, Antonio (2013): Temporally dissociable mechanisms of self-control: early attentional filtering versus late value modulation. In *J. Neurosci.* 33 (48), pp. 18917–18931. DOI: 10.1523/JNEUROSCI.5816-12.2013.

Harris, Christine R. (2012): Feelings of dread and intertemporal choice. In *J. Behav. Decis. Making* 25 (1), pp. 13–28. DOI: 10.1002/bdm.709.

Hassabis, Demis; Kumaran, Dharshan; Vann, Seralynne D.; Maguire, Eleanor A. (2007): Patients with hippocampal amnesia cannot imagine new experiences. In *Proc Natl Acad Sci USA* 104 (5), pp. 1726–1731. DOI: 10.1073/pnas.0610561104.

Heimer, L.; Switzer, R. D.; van Hoesen, G. W. (1982): Ventral striatum and ventral pallidum. In *Trends in neurosciences* 5, pp. 83–87. DOI: 10.1016/0166-2236(82)90037-6.

Hershfield, Hal E.; Goldstein, Daniel G.; Sharpe, William F.; Fox, Jesse; Yeykelis, Leo; Carstensen, Laura L.; Bailenson, Jeremy N. (2011): INCREASING SAVING BEHAVIOR THROUGH AGE-PROGRESSED RENDERINGS OF THE FUTURE SELF. In *JMR, Journal of marketing research* 48, S23-S37. DOI: 10.1509/jmkr.48.SPL.S23.

Hikida, Takatoshi; Kimura, Kensuke; Wada, Norio; Funabiki, Kazuo; Nakanishi, Shigetada (2010): Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. In *Neuron* 66 (6), pp. 896–907. DOI: 10.1016/j.neuron.2010.05.011.

Hirschtritt, Matthew E.; Lee, Paul C.; Pauls, David L.; Dion, Yves; Grados, Marco A.; Illmann, Cornelia et al. (2015): Lifetime prevalence, age of risk, and genetic relationships of comorbid psychiatric disorders in Tourette syndrome. In *JAMA psychiatry* 72 (4), pp. 325–333. DOI: 10.1001/jamapsychiatry.2014.2650.

Huang, Yunyun; Hu, Ping; Li, Xueting (2017): Undervaluing delayed rewards explains adolescents' impulsivity in inter-temporal choice: an ERP study. In *Sci Rep* 7 (1), p. 42631. DOI: 10.1038/srep42631.

Huys, Daniel; Kohl, Sina; Baldermann, Juan Carlos; Timmermann, Lars; Sturm, Volker; Visser-Vandewalle, Veerle; Kuhn, Jens (2019): Open-label trial of anterior limb of internal capsule-nucleus accumbens deep brain stimulation for obsessive-compulsive disorder: insights gained. In *Journal of neurology, neurosurgery, and psychiatry* 90 (7), pp. 805–812. DOI: 10.1136/jnnp-2018-318996.

Insel, Thomas; Cuthbert, Bruce; Garvey, Marjorie; Heinssen, Robert; Pine, Daniel S.; Quinn, Kevin et al. (2010): Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. In *American Journal of Psychiatry* 167 (7), pp. 748–751. DOI: 10.1176/appi.ajp.2010.09091379.

Jachimowicz, Jon M.; Chafik, Salah; Munrat, Sabeth; Prabhu, Jaideep C.; Weber, Elke U. (2017): Community trust reduces myopic decisions of low-income individuals. In *Proceedings of the National Academy of Sciences* 114 (21), pp. 5401–5406. DOI: 10.1073/pnas.1617395114.

Jackson, Georgina M.; Draper, Amelia; Dyke, Katherine; Pépés, Sophia E.; Jackson, Stephen R. (2015): Inhibition, Disinhibition, and the Control of Action in Tourette Syndrome. In *Trends in cognitive sciences* 19 (11), pp. 655–665. DOI: 10.1016/j.tics.2015.08.006.

Jackson, Jacob N. S.; MacKillop, James (2016): Attention-Deficit/Hyperactivity Disorder and Monetary Delay Discounting: A Meta-Analysis of Case-Control Studies. In *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 1 (4), pp. 316–325. DOI: 10.1016/J.BPSC.2016.01.007.

Jevons, William Stanley (2013): The Theory of Political Economy. Basingstoke: Palgrave Macmillan. Available online at http://gbv.eblib.com/patron/FullRecord.aspx?p=1609114.

Jimura, Koji; Chushak, Maria S.; Westbrook, Andrew; Braver, Todd S. (2018): Intertemporal Decision-Making Involves Prefrontal Control Mechanisms Associated with Working Memory. In *Cerebral cortex (New York, N.Y. : 1991)* 28 (4), pp. 1105–1116. DOI: 10.1093/cercor/bhx015.

Johnson, Matthew W. (2012): An efficient operant choice procedure for assessing delay discounting in humans: initial validation in cocaine-dependent and control individuals. In *Experimental and Clinical Psychopharmacology* 20 (3), pp. 191–204. DOI: 10.1037/a0027088.

Joshua, Mati; Adler, Avital; Mitelman, Rea; Vaadia, Eilon; Bergman, Hagai (2008): Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. In *J. Neurosci.* 28 (45), pp. 11673–11684. DOI: 10.1523/JNEUROSCI.3839-08.2008.

Joutsa, Juho; Johansson, Jarkko; Niemelä, Solja; Ollikainen, Antti; Hirvonen, Mika M.; Piepponen, Petteri et al. (2012): Mesolimbic dopamine release is linked to symptom severity in pathological gambling. In *NeuroImage* 60 (4), pp. 1992–1999. DOI: 10.1016/j.neuroimage.2012.02.006.

Kable, Joseph W.; Glimcher, Paul W. (2007): The neural correlates of subjective value during intertemporal choice. In *Nature neuroscience* 10 (12), pp. 1625–1633. DOI: 10.1038/nn2007.

Kaplan, Robert M.; Saccuzzo, Dennis P. (2018): Psychological testing. Principles, applications, & issues. Ninth edition. Boston, MA.: Cengage Learning.

Kashyap, Himani; Fontenelle, Leonardo F.; Miguel, Euripedes C.; Ferrão, Ygor A.; Torres, Albina R.; Shavitt, Roseli G. et al. (2012): 'Impulsive compulsivity' in obsessive-compulsive disorder: a phenotypic marker of patients with poor clinical outcome. In *Journal of Psychiatric Research* 46 (9), pp. 1146–1152. DOI: 10.1016/j.jpsychires.2012.04.022.

Kéri, Szabolcs; Szlobodnyik, Csaba; Benedek, György; Janka, Zoltán; Gádoros, Júlia (2002): Probabilistic classification learning in Tourette syndrome. In *Neuropsychologia* 40 (8), pp. 1356–1362. DOI: 10.1016/S0028-3932(01)00210-X.

Kirby, K. N.; Maraković, N. N. (1996): Delay-discounting probabilistic rewards: Rates decrease as amounts increase. In *Psychonomic Bulletin & Review* 3 (1), pp. 100–104. DOI: 10.3758/BF03210748.

Kirby, Kris N. (2009): One-year temporal stability of delay-discount rates. In *Psychonomic Bulletin & Review* 16 (3), pp. 457–462. DOI: 10.3758/PBR.16.3.457.

Kirby, Kris N.; Winston, Gordon C.; Santiesteban, Mariana (2005): Impatience and grades: Delay-discount rates correlate negatively with college GPA. In *Learning and Individual Differences* 15 (3), pp. 213–222. DOI: 10.1016/j.lindif.2005.01.003.

Knutson, Brian; Cooper, Jeffrey C. (2005): Functional magnetic resonance imaging of reward prediction. In *Current opinion in neurology* 18 (4), pp. 411–417. DOI: 10.1097/01.wco.0000173463.24758.f6.

Kobayashi, Kenji; Hsu, Ming (2019): Common neural code for reward and information value. In *Proceedings of the National Academy of Sciences of the United States of America* 116 (26), pp. 13061–13066. DOI: 10.1073/pnas.1820145116.

Kobayashi, Shunsuke; Schultz, Wolfram (2014): Reward contexts extend dopamine signals to unrewarded stimuli. In *Current biology : CB* 24 (1), pp. 56–62. DOI: 10.1016/j.cub.2013.10.061.

Kohl, Sina; Schönherr, Deva M.; Luigjes, Judy; Denys, Damiaan; Mueller, Ulf J.; Lenartz, Doris et al. (2014): Deep brain stimulation for treatment-refractory obsessive compulsive disorder: a systematic review. In *BMC psychiatry* 14, p. 214. DOI: 10.1186/s12888-014-0214-y.

Kozak, Michael J.; Cuthbert, Bruce N. (2016): The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. In *Psychophysiology* 53 (3), pp. 286–297. DOI: 10.1111/psyp.12518.

Kravitz, Alexxai V.; Tye, Lynne D.; Kreitzer, Anatol C. (2012): Distinct roles for direct and indirect pathway striatal neurons in reinforcement. In *Nat Neurosci* 15 (6), pp. 816–818. DOI: 10.1038/nn.3100.

Kroemer, Nils B.; Lee, Ying; Pooseh, Shakoor; Eppinger, Ben; Goschke, Thomas; Smolka, Michael N. (2019): L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action. In *NeuroImage* 186, pp. 113–125. DOI: 10.1016/j.neuroimage.2018.10.075.

Kruschke, John K. (2011): Bayesian Assessment of Null Values Via Parameter Estimation and Model Comparison. In *Perspectives on psychological science : a journal of the Association for Psychological Science* 6 (3), pp. 299–312. DOI: 10.1177/1745691611406925.

Kruschke, John K. (2013): Bayesian estimation supersedes the t test. In *Journal of experimental psychology. General* 142 (2), pp. 573–603. DOI: 10.1037/a0029146.

Kruschke, John K. (Ed.) (2015): Doing Bayesian data analysis. A tutorial with R, JAGS, and Stan. 2. ed. Amsterdam: AP Academic Press/Elsevier.

Kühn, Simone; Gallinat, Jürgen (2012): The neural correlates of subjective pleasantness. In *NeuroImage* 61 (1), pp. 289–294. DOI: 10.1016/j.neuroimage.2012.02.065.

Lak, Armin; Stauffer, William R.; Schultz, Wolfram (2016): Dopamine neurons learn relative chosen value from probabilistic rewards. In *eLife* 5. DOI: 10.7554/eLife.18044.

Larson, Paul S. (2014): Deep brain stimulation for movement disorders. In *Neurotherapeutics* 11 (3), pp. 465–474. DOI: 10.1007/s13311-014-0274-1.

Lawrance, Emily C. (1991): Poverty and the Rate of Time Preference: Evidence from Panel Data. In *Journal of Political Economy* 99 (1), pp. 54–77. DOI: 10.1086/261740.

Lebreton, Maël; Bertoux, Maxime; Boutet, Claire; Lehericy, Stéphane; Dubois, Bruno; Fossati, Philippe; Pessiglione, Mathias (2013): A critical role for the hippocampus in the valuation of imagined outcomes. In *PLOS Biology* 11 (10), e1001684. DOI: 10.1371/journal.pbio.1001684.

Lempert, Karolina M.; Phelps, Elizabeth A. (2016): The Malleability of Intertemporal Choice. In *Trends in cognitive sciences* 20 (1), pp. 64–74. DOI: 10.1016/j.tics.2015.09.005.

Lempert, Karolina M.; Steinglass, Joanna E.; Pinto, Anthony; Kable, Joseph W.; Simpson, Helen Blair (2019): Can delay discounting deliver on the promise of RDoC? In *Psychological medicine* 49 (2), pp. 190–199. DOI: 10.1017/S0033291718001770.

Leyton, Marco; Casey, Kevin F.; Delaney, J. Scott; Kolivakis, Theodore; Benkelfat, Chawki (2005): Cocaine craving, euphoria, and self-administration: a preliminary study of the effect of catecholamine precursor depletion. In *Behavioral Neuroscience* 119 (6), pp. 1619–1627. DOI: 10.1037/0735-7044.119.6.1619.

Linnet, Jakob; Møller, Arne; Peterson, Ericka; Gjedde, Albert; Doudet, Doris (2011): Dopamine release in ventral striatum during Iowa Gambling Task performance is associated with increased excitement levels in pathological gambling. In *Addiction* 106 (2), pp. 383–390. DOI: 10.1111/j.1360-0443.2010.03126.x.

Loewenstein, George; Elster, Jon (1992): Choice over time. New York: Russell Sage Foundation.

Lozano, Andres M.; Lipsman, Nir (2013): Probing and regulating dysfunctional circuits using deep brain stimulation. In *Neuron* 77 (3), pp. 406–424. DOI: 10.1016/j.neuron.2013.01.020.

Luigjes, Judy; Mantione, Mariska; van den Brink, Wim; Schuurman, P. Richard; van den Munckhof, Pepijn; Denys, Damiaan (2011): Deep brain stimulation increases impulsivity in two patients with obsessive-compulsive disorder. In *International clinical psychopharmacology* 26 (6), pp. 338–340. DOI: 10.1097/YIC.0b013e32834af505.

MacKillop, James; Amlung, Michael T.; Few, Lauren R.; Ray, Lara A.; Sweet, Lawrence H.; Munafò, Marcus R. (2011): Delayed reward discounting and addictive behavior: a meta-analysis. In *Psychopharmacology* 216 (3), pp. 305–321. DOI: 10.1007/s00213-011-2229-0.

Maia, Tiago V.; Conceição, Vasco A. (2018): Dopaminergic Disturbances in Tourette Syndrome: An Integrative Account. In *Biological psychiatry* 84 (5), pp. 332–344. DOI: 10.1016/j.biopsych.2018.02.1172.

Maia, Tiago V.; Frank, Michael J. (2017): An Integrative Perspective on the Role of Dopamine in Schizophrenia. In *Biological psychiatry* 81 (1), pp. 52–66. DOI: 10.1016/j.biopsych.2016.05.021.

Mariano, T. Y.; Bannerman, D. M.; McHugh, S. B.; Preston, T. J.; Rudebeck, P. H.; Rudebeck, S. R. et al. (2009): Impulsive choice in hippocampal but not orbitofrontal cortex-lesioned rats on a nonspatial decision-making maze task. In *The European journal of neuroscience* 30 (3), pp. 472–484. DOI: 10.1111/j.1460-9568.2009.06837.x.

Masuda, Akira; Sano, Chie; Zhang, Qi; Goto, Hiromichi; McHugh, Thomas J.; Fujisawa, Shigeyoshi; Itohara, Shigeyoshi (2020): The hippocampus encodes delay and value information during delay-discounting decision making. In *eLife* 9. DOI: 10.7554/eLife.52466.

Mazo, Robert M. (2009): Brownian motion. Fluctuations, dynamics, and applications. 1. publ. in paperback. Oxford: Oxford Univ. Press (International series of monographs on physics, 112).

Mazur, J. E.; Coe, D. (1987): Tests of transitivity in choices between fixed and variable reinforcer delays. In *Journal of the experimental analysis of behavior* 47 (3), pp. 287–297. DOI: 10.1901/jeab.1987.47-287.

McClelland, David C.; Patel, Vandana; Stier, Deborah; Brown, Don (1987): The relationship of affiliative arousal to dopamine release. In *Motiv Emot* 11 (1), pp. 51–66. DOI: 10.1007/BF00992213.

McClure, Samuel M.; Laibson, David I.; Loewenstein, George; Cohen, Jonathan D. (2004): Separate neural systems value immediate and delayed monetary rewards. In *Science (New York, N.Y.)* 306 (5695), pp. 503–507. DOI: 10.1126/science.1100907.

McGuigan, Sara; Zhou, Shou-Han; Brosnan, Méadhbh B.; Thyagarajan, Dominic; Bellgrove, Mark A.; Chong, Trevor T-J (2019): Dopamine restores cognitive motivation in Parkinson's disease. In *Brain* 142 (3), pp. 719–732. DOI: 10.1093/brain/awy341.

McKerchar, Todd L.; Green, Leonard; Myerson, Joel; Pickford, T. Stephen; Hill, Jade C.; Stout, Steven C. (2009): A comparison of four models of delay discounting in humans. In *Behavioural Processes* 81 (2), pp. 256–259. DOI: 10.1016/j.beproc.2008.12.017.

Meier, Stephan; Sprenger, Charles D. (2012): Time discounting predicts creditworthiness. In *Psychol Sci* 23 (1), pp. 56–58. DOI: 10.1177/0956797611425931.

Meiser, Johannes; Weindl, Daniel; Hiller, Karsten (2013): Complexity of dopamine metabolism. In *Cell communication and signaling : CCS* 11 (1), p. 34. DOI: 10.1186/1478-811X-11-34.

Michaelson, Laura; La Vega, Alejandro de; Chatham, Christopher H.; Munakata, Yuko (2013): Delaying gratification depends on social trust. In *Front. Psychol.* 4, p. 355. DOI: 10.3389/fpsyg.2013.00355.

Miedl, Stephan F.; Büchel, Christian; Peters, Jan (2014): Cue-induced craving increases impulsivity via changes in striatal value signals in problem gamblers. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 34 (13), pp. 4750–4755. DOI: 10.1523/JNEUROSCI.5020-13.2014.

Miedl, Stephan F.; Peters, Jan; Büchel, Christian (2012): Altered neural reward representations in pathological gamblers revealed by delay and probability discounting. In *Archives of general psychiatry* 69 (2), pp. 177–186. DOI: 10.1001/archgenpsychiatry.2011.1552.

Mikhael, John G.; Gershman, Samuel J. (2021): Impulsivity and risk-seeking as Bayesian inference under dopaminergic control. In *Neuropsychopharmacol*. DOI: 10.1038/s41386-021-01125-z.

Mikhael, John G.; Lai, Lucy; Gershman, Samuel J. (2021): Rational inattention and tonic dopamine. In *PLOS Computational Biology* 17 (3), e1008659. DOI: 10.1371/journal.pcbi.1008659.

Milad, Mohammed R.; Rauch, Scott L. (2012): Obsessive-compulsive disorder: beyond segregated cortico-striatal pathways. In *Trends in cognitive sciences* 16 (1), pp. 43–51. DOI: 10.1016/j.tics.2011.11.003.

Miletić, Steven; Boag, Russell J.; Forstmann, Birte U. (2020): Mutual benefits: Combining reinforcement learning with sequential sampling models. In *Neuropsychologia* 136, p. 107261. DOI: 10.1016/j.neuropsychologia.2019.107261.

Miller, E. K. (2000): The prefrontal cortex and cognitive control. In *Nat Rev Neurosci* 1 (1), pp. 59–65. DOI: 10.1038/35036228.

Mirenowicz, J.; Schultz, W. (1994): Importance of unpredictability for reward responses in primate dopamine neurons. In *Journal of neurophysiology* 72 (2), pp. 1024–1027. DOI: 10.1152/jn.1994.72.2.1024.

Mischel, W.; Shoda, Y.; Peake, P. K. (1988): The nature of adolescent competencies predicted by preschool delay of gratification. In *Journal of Personality and Social Psychology* 54 (4), pp. 687–696. DOI: 10.1037//0022-3514.54.4.687.

Mischel, W.; Shoda, Y.; Rodriguez, M. I. (1989): Delay of gratification in children. In *Science (New York, N.Y.)* 244 (4907), pp. 933–938. DOI: 10.1126/science.2658056.

Mischel, Walter; Ebbesen, Ebbe B. (1970): Attention in delay of gratification. In *Journal of Personality and Social Psychology* 16 (2), pp. 329–337. DOI: 10.1037/h0029815.

Modell, J. G.; Mountz, J. M.; Curtis, G. C.; Greden, J. F. (1989): Neurophysiologic dysfunction in basal ganglia/limbic striatal and thalamocortical circuits as a pathogenetic mechanism of obsessive-compulsive disorder. In *The Journal of Neuropsychiatry and Clinical Neurosciences* 1 (1), pp. 27–36. DOI: 10.1176/jnp.1.1.27.

Mohebi, Ali; Pettibone, Jeffrey R.; Hamid, Arif A.; Wong, Jenny-Marie T.; Vinson, Leah T.; Patriarchi, Tommaso et al. (2019): Dissociable dopamine dynamics for learning and motivation. In *Nature* 570 (7759), pp. 65–70. DOI: 10.1038/s41586-019-1235-y.

Montague, P. R.; Dayan, P.; Sejnowski, T. J. (1996): A framework for mesencephalic dopamine systems based on predictive Hebbian learning. In *J. Neurosci.* 16 (5), pp. 1936–1947. DOI: 10.1523/JNEUROSCI.16-05-01936.1996.

Muggleton, Naomi; Parpart, Paula; Newall, Philip; Leake, David; Gathergood, John; Stewart, Neil (2021): The association between gambling and financial, social and health outcomes in big financial data. In *Nature human behaviour*. DOI: 10.1038/s41562-020-01045-w.

Muraven, Mark (2010): Building Self-Control Strength: Practicing Self-Control Leads to Improved Self-Control Performance. In *Journal of Experimental Social Psychology* 46 (2), pp. 465–468. DOI: 10.1016/j.jesp.2009.12.011.

Myerson, J.; Green, L. (1995): Discounting of delayed rewards: Models of individual choice. In *Journal of the experimental analysis of behavior* 64 (3), pp. 263–276. DOI: 10.1901/jeab.1995.64-263.

Myerson, J.; Green, L.; Warusawitharana, M. (2001): Area under the curve as a measure of discounting. In *Journal of the experimental analysis of behavior* 76 (2), pp. 235–243. DOI: 10.1901/jeab.2001.76-235.

Nachev, Parashkev; Lopez-Sosa, Fernando; Gonzalez-Rosa, Javier Jesus; Galarza, Ana; Avecillas, Josue; Pineda-Pardo, Jose Angel et al. (2015): Dynamic risk control by human nucleus accumbens. In *Brain : a journal of neurology* 138 (Pt 12), pp. 3496–3502. DOI: 10.1093/brain/awv285.

Nakahara, Hiroyuki; Itoh, Hideaki; Kawagoe, Reiko; Takikawa, Yoriko; Hikosaka, Okihide (2004): Dopamine Neurons Can Represent Context-Dependent Prediction Error. In *Neuron* 41 (2), pp. 269–280. DOI: 10.1016/S0896-6273(03)00869-9.

Nelson, Edward (1967): Dynamical Theories of Brownian Motion. Princeton, NJ: Princeton University Press (Mathematical Notes, 106).

Nieoullon, A.; Cheramy, A.; Glowinski, J. (1978): Release of dopamine evoked by electrical stimulation of the motor and visual areas of the cerebral cortex in both caudate nuclei and in the substantia nigra in the cat. In *Brain Research* 145 (1), pp. 69–83. DOI: 10.1016/0006-8993(78)90797-7.

Nunes, E. J.; Randall, P. A.; Santerre, J. L.; Given, A. B.; Sager, T. N.; Correa, M.; Salamone, J. D. (2010): Differential effects of selective adenosine antagonists on the effort-related impairments induced by dopamine D1 and D2 antagonism. In *Neuroscience* 170 (1), pp. 268–280. DOI: 10.1016/j.neuroscience.2010.05.068.

Oberg, Scott A. K.; Christie, Gregory J.; Tata, Matthew S. (2011): Problem gamblers exhibit reward hypersensitivity in medial frontal cortex during gambling. In *Neuropsychologia* 49 (13), pp. 3768–3775. DOI: 10.1016/j.neuropsychologia.2011.09.037.

Odum, Amy L. (2011): Delay discounting: trait variable? In *Behavioural Processes* 87 (1), pp. 1–9. DOI: 10.1016/j.beproc.2011.02.007.

Palminteri, Stefano; Lebreton, Maël; Worbe, Yulia; Grabli, David; Hartmann, Andreas; Pessiglione, Mathias (2009): Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. In *Proceedings of the National Academy of Sciences* 106 (45), pp. 19179–19184. DOI: 10.1073/pnas.0904035106.

Palminteri, Stefano; Pessiglione, Mathias (2013): Reinforcement learning and Tourette syndrome: Elsevier (International Review of Neurobiology, 112). In *International review of neurobiology*, pp. 131–153.

Parent, André; Hazrati, Lili-Naz (1995): Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. In *Brain Research Reviews* 20 (1), pp. 91–127. DOI: 10.1016/0165-0173(94)00007-C.

Patt, Virginie M.; Hunsberger, Renee; Jones, Dominoe A.; Keane, Margaret M.; Verfaellie, Mieke (2021): Temporal discounting when outcomes are experienced in the moment: Validation of a novel paradigm and comparison with a classic hypothetical intertemporal choice task. In *PloS one* 16 (5), e0251480. DOI: 10.1371/journal.pone.0251480.

Peciña, Susana; Berridge, Kent C. (2005): Hedonic hot spot in nucleus accumbens shell: where do mu-opioids cause increased hedonic impact of sweetness? In *J. Neurosci.* 25 (50), pp. 11777–11786. DOI: 10.1523/JNEUROSCI.2329-05.2005.

Peciña, Susana; Smith, Kyle S.; Berridge, Kent C. (2006): Hedonic hot spots in the brain. In *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 12 (6), pp. 500–511. DOI: 10.1177/1073858406293154.

Pedersen, Mads Lund; Frank, Michael J.; Biele, Guido (2017): The drift diffusion model as the choice rule in reinforcement learning. In *Psychonomic Bulletin & Review* 24 (4), pp. 1234–1251. DOI: 10.3758/s13423-016-1199-y.

Peisker, Canan B.; Schüller, Thomas; Peters, Jan; Wagner, Ben J.; Schilbach, Leonhard; Müller, Ulf J. et al. (2018): Nucleus Accumbens Deep Brain Stimulation in Patients with Substance Use Disorders and Delay Discounting. In *Brain Sciences* 8 (2). DOI: 10.3390/brainsci8020021.

Peters, J.; Büchel, C. (2010a): Neural representations of subjective reward value. In *Behavioural brain research* 213 (2), pp. 135–141. DOI: 10.1016/j.bbr.2010.04.031.

Peters, Jan; Bromberg, Uli; Schneider, Sophia; Brassen, Stefanie; Menz, Mareike; Banaschewski, Tobias et al. (2011): Lower ventral striatal activation during reward anticipation in adolescent smokers. In *American Journal of Psychiatry* 168 (5), pp. 540–549. DOI: 10.1176/appi.ajp.2010.10071024.

Peters, Jan; Büchel, Christian (2009): Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. In *The Journal of neuroscience : the official*

journal of the Society for Neuroscience 29 (50), pp. 15727–15734. DOI: 10.1523/JNEUROSCI.3489-09.2009.

Peters, Jan; Büchel, Christian (2010b): Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. In *Neuron* 66 (1), pp. 138–148. DOI: 10.1016/j.neuron.2010.03.026.

Peters, Jan; Büchel, Christian (2011): The neural mechanisms of inter-temporal decision-making. Understanding variability. In *Trends in cognitive sciences* 15 (5), pp. 227–239. DOI: 10.1016/j.tics.2011.03.002.

Peters, Jan; D'Esposito, Mark (2016): Effects of Medial Orbitofrontal Cortex Lesions on Self-Control in Intertemporal Choice. In *Current biology : CB* 26 (19), pp. 2625–2628. DOI: 10.1016/j.cub.2016.07.035.

Peters, Jan; D'Esposito, Mark (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. In *PLOS Computational Biology* 16 (4), e1007615. DOI: 10.1371/journal.pcbi.1007615.

Petzold, Johannes; Kienast, Annika; Lee, Ying; Pooseh, Shakoor; London, Edythe D.; Goschke, Thomas; Smolka, Michael N. (2019): Baseline impulsivity may moderate L-DOPA effects on value-based decision-making. In *Scientific reports* 9 (1), p. 5652. DOI: 10.1038/s41598-019-42124-x.

Pine, Alex; Shiner, Tamara; Seymour, Ben; Dolan, Raymond J. (2010): Dopamine, time, and impulsivity in humans. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30 (26), pp. 8888–8896. DOI: 10.1523/JNEUROSCI.6028-09.2010.

Pinto, Anthony; Steinglass, Joanna E.; Greene, Ashley L.; Weber, Elke U.; Simpson, H. Blair (2014): Capacity to delay reward differentiates obsessive-compulsive disorder and obsessive-compulsive personality disorder. In *Biological psychiatry* 75 (8), pp. 653–659. DOI: 10.1016/j.biopsych.2013.09.007.

Plummer, Martyn (Ed.) (2003): JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003).

Potenza, Marc N. (2013): How central is dopamine to pathological gambling or gambling disorder? In *Front. Behav. Neurosci.* 7, p. 206. DOI: 10.3389/fnbeh.2013.00206.

Potenza, Marc N. (2018): Searching for Replicable Dopamine-Related Findings in Gambling Disorder. In *Biological psychiatry* 83 (12), pp. 984–986. DOI: 10.1016/j.biopsych.2018.04.011.

Race, Elizabeth; Keane, Margaret M.; Verfaellie, Mieke (2011): Medial temporal lobe damage causes deficits in episodic memory and episodic future thinking not attributable to deficits in narrative construction. In *J. Neurosci.* 31 (28), pp. 10262–10269. DOI: 10.1523/JNEUROSCI.1145-11.2011.

Ratcliff, Roger (1978): A theory of memory retrieval. In *Psychological Review* 85 (2), pp. 59–108. DOI: 10.1037/0033-295x.85.2.59.

Ratcliff, Roger; McKoon, Gail (2008): The diffusion decision model: theory and data for two-choice decision tasks. In *Neural Computation* 20 (4), pp. 873–922. DOI: 10.1162/neco.2008.12-06-420.

Raylu, Namrata; Oei, Tian P. S. (2004): The Gambling Related Cognitions Scale (GRCS): Development, confirmatory factor validation and psychometric properties. In *Addiction (Abingdon, England)* 99 (6), pp. 757–769.

Reimers, Stian; Maylor, Elizabeth A.; Stewart, Neil; Chater, Nick (2009): Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. In *Personality and Individual Differences* 47 (8), pp. 973–978. DOI: 10.1016/j.paid.2009.07.026.

Reynolds, Brady (2006): A review of delay-discounting research with humans: relations to drug use and gambling. In *Behavioural pharmacology* 17 (8), pp. 651–667. DOI: 10.1097/FBP.0b013e3280115f99.

Ripke, Stephan; Hübner, Thomas; Mennigen, Eva; Müller, Kathrin U.; Rodehacke, Sarah; Schmidt, Dirk et al. (2012): Reward processing and intertemporal decision making in adults and adolescents: the role of impulsivity and decision consistency. In *Brain Research* 1478, pp. 36–47. DOI: 10.1016/j.brainres.2012.08.034.

Robbins, T. W.; Everitt, B. J. (1999): Drug addiction: bad habits add up. In *Nature* 398 (6728), pp. 567–570. DOI: 10.1038/19208.

Robbins, Trevor W.; Gillan, Claire M.; Smith, Dana G.; Wit, Sanne de; Ersche, Karen D. (2012): Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. In *Trends in cognitive sciences* 16 (1), pp. 81–91. DOI: 10.1016/j.tics.2011.11.009.

Robbins, Trevor W.; Vaghi, Matilde M.; Banca, Paula (2019): Obsessive-Compulsive Disorder: Puzzles and Prospects. In *Neuron* 102 (1), pp. 27–47. DOI: 10.1016/j.neuron.2019.01.046.

Robertson, Mary May (2012): The Gilles de la Tourette syndrome: the current status. In *Archives of disease in childhood. Education and practice edition* 97 (5), pp. 166–175. DOI: 10.1136/archdischild-2011-300585.

Robinson, Siobhan; Sandstrom, Suzanne M.; Denenberg, Victor H.; Palmiter, Richard D. (2005): Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. In *Behavioral Neuroscience* 119 (1), pp. 5–15. DOI: 10.1037/0735-7044.119.1.5.

Robinson, Siobhan; Smith, David M.; Mizumori, Sheri J. Y.; Palmiter, Richard D. (2004): Firing properties of dopamine neurons in freely moving dopamine-deficient mice: effects of dopamine receptor activation and anesthesia. In *Proc Natl Acad Sci USA* 101 (36), pp. 13329–13334. DOI: 10.1073/pnas.0405084101.

Robinson, T.; Berridge, Kent C. (1993): The neural basis of drug craving: An incentive-sensitization theory of addiction. In *Brain Research Reviews* 18 (3), pp. 247–291. DOI: 10.1016/0165-0173(93)90013-P.

Robinson, T. E.; Berridge, K. C. (2001): Incentive-sensitization and addiction. In *Addiction (Abingdon, England)* 96 (1), pp. 103–114. DOI: 10.1046/j.1360-0443.2001.9611038.x.

Robinson, Terry E.; Berridge, Kent C. (2008): Review. The incentive sensitization theory of addiction: some current issues. In *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1507), pp. 3137–3146. DOI: 10.1098/rstb.2008.0093.

Rodriguez-Moreno, Diana V.; Cycowicz, Yael M.; Figner, Bernd; Wang, Zhishun; He, Xiaofu; Geronazzo-Alman, Lupo et al. (2021): Delay discounting and neurocognitive correlates among inner city adolescents with and without family history of substance use disorder. In *Developmental cognitive neuroscience* 48, p. 100942. DOI: 10.1016/j.dcn.2021.100942.

Rogers, Robert D. (2011): The roles of dopamine and serotonin in decision making: evidence from pharmacological experiments in humans. In *Neuropsychopharmacol* 36 (1), pp. 114–132. DOI: 10.1038/npp.2010.165.

Rösch, Sarah Alica; Stramaccia, Davide Francesco; Benoit, Roland Georg (2021): Promoting farsighted decisions via episodic future thinking: A meta-analysis.

Rossow, Ingeborg (2008): Alcohol consumption and discounting. In *Addiction Research & Theory* 16 (6), pp. 572–584. DOI: 10.1080/16066350801896248.

Saddoris, Michael P.; Cacciapaglia, Fabio; Wightman, R. Mark; Carelli, Regina M. (2015): Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 35 (33), pp. 11572–11582. DOI: 10.1523/JNEUROSCI.2344-15.2015.

Samuelson, Paul A. (1937): A Note on Measurement of Utility. In *The Review of Economic Studies* 4 (2), p. 155. DOI: 10.2307/2967612.

Satoh, Takemasa; Nakai, Sadamu; Sato, Tatsuo; Kimura, Minoru (2003): Correlated Coding of Motivation and Outcome of Decision by Dopamine Neurons. In *The Journal of Neuroscience* 23 (30), pp. 9913–9923. DOI: 10.1523/JNEUROSCI.23-30-09913.2003.

Schacter, Daniel L.; Addis, Donna Rose; Hassabis, Demis; Martin, Victoria C.; Spreng, R. Nathan; Szpunar, Karl K. (2012): The future of memory: remembering, imagining, and the brain. In *Neuron* 76 (4), pp. 677–694. DOI: 10.1016/j.neuron.2012.11.001.

Schelp, Scott A.; Pultorak, Katherine J.; Rakowski, Dylan R.; Gomez, Devan M.; Krzystyniak, Gregory; Das, Raibatak; Oleson, Erik B. (2017): A transient dopamine signal encodes subjective value and causally influences demand in an economic context. In *PNAS* 114 (52), E11303-E11312. DOI: 10.1073/pnas.1706969114.

Schneider, S.; Peters, J.; Peth, J. M.; Büchel, C. (2014): Parental inconsistency, impulsive choice and neural value representations in healthy adolescents. In *Transl Psychiatry* 4 (4), e382. DOI: 10.1038/tp.2014.20.

Schultz, W.; Apicella, P.; Ljungberg, T. (1993): Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. In *J. Neurosci.* 13 (3), pp. 900–913. DOI: 10.1523/JNEUROSCI.13-03-00900.1993.

Schultz, W.; Romo, R. (1990): Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. In *Journal of neurophysiology* 63 (3), pp. 607–624. DOI: 10.1152/jn.1990.63.3.607.

Schultz, Wolfram (2007): Multiple dopamine functions at different time courses. In *Annual review of neuroscience* 30, pp. 259–288. DOI: 10.1146/annurev.neuro.28.061604.135722.

Schultz, Wolfram (2013): Updating dopamine reward signals. In *Current opinion in neurobiology* 23 (2), pp. 229–238. DOI: 10.1016/j.conb.2012.11.012.

Schultz, Wolfram (2015): Neuronal Reward and Decision Signals: From Theories to Data. In *Physiological reviews* 95 (3), pp. 853–951. DOI: 10.1152/physrev.00023.2014.

Schultz, Wolfram (2016): Dopamine reward prediction-error signalling: a two-component response. In *Nature reviews. Neuroscience* 17 (3), pp. 183–195. DOI: 10.1038/nrn.2015.26.

Schultz, Wolfram; Stauffer, Wiliam R.; Lak, Armin (2017): The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility. In *Current opinion in neurobiology* 43, pp. 139–148. DOI: 10.1016/j.conb.2017.03.013.

Schwartenbeck, Philipp; FitzGerald, Thomas H. B.; Mathys, Christoph; Dolan, Ray; Friston, Karl (2015): The Dopaminergic Midbrain Encodes the Expected Certainty about Desired Outcomes. In *Cerebral cortex (New York, N.Y. : 1991)* 25 (10), pp. 3434–3445. DOI: 10.1093/cercor/bhu159.

Schwarz, Adam; Gozzi, Alessandro; Reese, Torsten; Bertani, Simone; Crestan, Valerio; Hagan, Jim et al. (2004): Selective dopamine D(3) receptor antagonist SB-277011-A potentiates phMRI response to acute amphetamine challenge in the rat brain. In *Synapse* 54 (1), pp. 1–10. DOI: 10.1002/SYN.20055.

Seaman, Kendra L.; Brooks, Nickolas; Karrer, Teresa M.; Castrellon, Jaime J.; Perkins, Scott F.; Dang, Linh C. et al. (2018): Subjective value representations during effort, probability and time discounting across adulthood. In *Social cognitive and affective neuroscience* 13 (5), pp. 449–459. DOI: 10.1093/scan/nsy021.

Seamans, Jeremy K.; Yang, Charles R. (2004): The principal features and mechanisms of dopamine modulation in the prefrontal cortex. In *Progress in Neurobiology* 74 (1), pp. 1–58. DOI: 10.1016/J.PNEUROBIO.2004.05.006.

Seinstra, Maayke; Wojtecki, Lars; Storzer, Lena; Schnitzler, Alfons; Kalenscher, Tobias (2016): No Effect of Subthalamic Deep Brain Stimulation on Intertemporal Decision-Making in Parkinson Patients. In *eNeuro* 3 (2), ENEURO.0019-16.2016. DOI: 10.1523/ENEURO.0019-16.2016.

Selen, Luc P. J.; Shadlen, Michael N.; Wolpert, Daniel M. (2012): Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. In *J. Neurosci.* 32 (7), pp. 2276–2286. DOI: 10.1523/JNEUROSCI.5273-11.2012.

Sescousse, Guillaume; Barbalat, Guillaume; Domenech, Philippe; Dreher, Jean-Claude (2013): Imbalance in the sensitivity to different types of rewards in pathological gambling. In *Brain* 136 (Pt 8), pp. 2527–2538. DOI: 10.1093/brain/awt126.

Shadlen, Michael N.; Kiani, Roozbeh (2013): Decision making as a window on cognition. In *Neuron* 80 (3), pp. 791–806. DOI: 10.1016/j.neuron.2013.10.047.

Shahar, Nitzan; Hauser, Tobias U.; Moutoussis, Michael; Moran, Rani; Keramati, Mehdi; Dolan, Raymond J. (2019): Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. In *PLoS computational biology* 15 (2), e1006803. DOI: 10.1371/journal.pcbi.1006803.

Sharp, Madeleine E.; Foerde, Karin; Daw, Nathaniel D.; Shohamy, Daphna (2016): Dopamine selectively remediates 'model-based' reward learning: a computational approach. In *Brain* 139 (Pt 2), pp. 355–364. DOI: 10.1093/brain/awv347.

Sheffer, Christine E.; Christensen, Darren R.; Landes, Reid; Carter, Larry P.; Jackson, Lisa; Bickel, Warren K. (2014): Delay discounting rates: a strong prognostic indicator of smoking relapse. In *Addictive behaviors* 39 (11), pp. 1682–1689. DOI: 10.1016/j.addbeh.2014.04.019.

Shen, Weixing; Flajolet, Marc; Greengard, Paul; Surmeier, D. James (2008): Dichotomous dopaminergic control of striatal synaptic plasticity. In *Science* 321 (5890), pp. 848–851. DOI: 10.1126/science.1160575.

Shiels, Keri; Hawk, Larry W.; Reynolds, Brady; Mazzullo, Rebecca J.; Rhodes, Jessica D.; Pelham, William E. et al. (2009): Effects of methylphenidate on discounting of delayed rewards in attention deficit/hyperactivity disorder. In *Experimental and Clinical Psychopharmacology* 17 (5), pp. 291–301. DOI: 10.1037/a0017259.

Sienkiewicz-Jarosz, H.; Scinska, A.; Kuran, W.; Ryglewicz, D.; Rogowski, A.; Wrobel, E. et al. (2005): Taste responses in patients with Parkinson's disease. In *Journal of Neurology, Neurosurgery & Psychiatry* 76 (1), pp. 40–46. DOI: 10.1136/jnnp.2003.033373.

Sienkiewicz-Jarosz, Halina; Scinska, Anna; Swiecicki, Lukasz; Lipczynska-Lojkowska, Wanda; Kuran, Wlodzimierz; Ryglewicz, Danuta et al. (2013): Sweet liking in patients with Parkinson's disease. In *Journal of the neurological sciences* 329 (1-2), pp. 17–22. DOI: 10.1016/j.jns.2013.03.005.

Singer, Harvey S. (2013): The Neurochemistry of Tourette Syndrome. In : Tourette Syndrome: Oxford University Press, pp. 276–300.

Singer, Harvey S.; Szymanski, Sally; Giuliano, Joseph; Yokoi, Fuji; Dogan, A. Semih; Brasic, James R. et al. (2002): Elevated intrasynaptic dopamine release in Tourette's syndrome measured by PET. In *The American journal of psychiatry* 159 (8), pp. 1329–1336. DOI: 10.1176/appi.ajp.159.8.1329.

Smith, Christopher T.; San Juan, M. Danica; Dang, Linh C.; Katz, Daniel T.; Perkins, Scott F.; Burgess, Leah L. et al. (2018): Ventral striatal dopamine transporter availability is associated with lower trait motor impulsivity in healthy adults. In *Translational psychiatry* 8 (1), p. 269. DOI: 10.1038/s41398-018-0328-y.

Smith, Kyle S.; Berridge, Kent C.; Aldridge, J. Wayne (2011): Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. In *PNAS* 108 (27), E255-64. DOI: 10.1073/pnas.1101920108.

Smits, Rochelle R.; Stein, Jeffrey S.; Johnson, Patrick S.; Odum, Amy L.; Madden, Gregory J. (2013): Test-retest reliability and construct validity of the Experiential Discounting Task. In *Experimental and Clinical Psychopharmacology* 21 (2), pp. 155–163. DOI: 10.1037/a0031725.

Sohn, Sung Yun; Kang, Jee in; Namkoong, Kee; Kim, Se Joo (2014): Multidimensional measures of impulsivity in obsessive-compulsive disorder: cannot wait and stop. In *PloS one* 9 (11), e111739. DOI: 10.1371/journal.pone.0111739.

Stan Development Team (2020): Stan. Version 2.26. Available online at https://mc-stan.org, checked on 2021.

Stanger, Catherine; Ryan, Stacy R.; Fu, Hongyun; Landes, Reid D.; Jones, Bryan A.; Bickel, Warren K.; Budney, Alan J. (2012): Delay discounting predicts adolescent substance abuse treatment outcome. In *Experimental and Clinical Psychopharmacology* 20 (3), pp. 205–212. DOI: 10.1037/a0026543.

Steele, Catherine C.; Gwinner, MacKenzie; Smith, Travis; Young, Michael E.; Kirkpatrick, Kimberly (2019): Experience Matters: The Effects of Hypothetical versus Experiential Delays and Magnitudes on Impulsive Choice in Delay Discounting Tasks. In *Brain Sciences* 9 (12), p. 379. DOI: 10.3390/brainsci9120379.

Steinbeis, Nikolaus; Haushofer, Johannes; Fehr, Ernst; Singer, Tania (2016): Development of Behavioral Control and Associated vmPFC-DLPFC Connectivity Explains Children's Increased Resistance to Temptation in Intertemporal Choice. In *Cerebral cortex (New York, N.Y. : 1991)* 26 (1), pp. 32–42. DOI: 10.1093/cercor/bhu167.

Steinglass, Joanna E.; Lempert, Karolina M.; Choo, Tse-Hwei; Kimeldorf, Marcia B.; Wall, Melanie; Walsh, B. Timothy et al. (2017): Temporal discounting across three psychiatric disorders: Anorexia nervosa, obsessive compulsive disorder, and social anxiety disorder. In *Depression and Anxiety* 34 (5), pp. 463–470. DOI: 10.1002/da.22586.

Steketee, Jeffery D.; Kalivas, Peter W. (2011): Drug wanting: behavioral sensitization and relapse to drug-seeking behavior. In *Pharmacol Rev* 63 (2), pp. 348–365. DOI: 10.1124/pr.109.001933.

Sulkunen, Pekka; Babor, Thomas F.; Cisneros Örnberg, Jenny; Egerer, Michael; Hellman, Matilda; Livingstone, Charles et al. (2021): Setting Limits: Gambling, Science and Public Policy-summary of results. In *Addiction* 116 (1), pp. 32–40. DOI: 10.1111/add.15241.

Sutton, R. S.; Barto, A. G. (1990): Time-derivative models of Pavlovian reinforcement. Available online at https://psycnet.apa.org/record/1991-97439-012.

Sutton, R. S.; Barto, A. G. (1998): Reinforcement Learning: An Introduction. In *IEEE Trans. Neural Netw.* 9 (5), p. 1054. DOI: 10.1109/TNN.1998.712192.

Sutton, Richard S. (1988): Learning to predict by the methods of temporal differences. In *Mach Learn* 3 (1), pp. 9–44. DOI: 10.1007/bf00115009.

Taber, Katherine H.; Black, Deborah N.; Porrino, Linda J.; Hurley, Robin A. (2012): Neuroanatomy of dopamine: reward and addiction. In *The Journal of Neuropsychiatry and Clinical Neurosciences* 24 (1), pp. 1–4. DOI: 10.1176/appi.neuropsych.24.1.1.

Tabibnia, Golnaz; Monterosso, John R.; Baicy, Kate; Aron, Adam R.; Poldrack, Russell A.; Chakrapani, Shruthi et al. (2011): Different forms of self-control share a neurocognitive substrate. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31 (13), pp. 4805–4810. DOI: 10.1523/JNEUROSCI.2859-10.2011.

Tai, Lung-Hao; Lee, A. Moses; Benavidez, Nora; Bonci, Antonello; Wilbrecht, Linda (2012): Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. In *Nat Neurosci* 15 (9), pp. 1281–1289. DOI: 10.1038/nn.3188.

Tanabe, Jody; Thompson, Laetitia; Claus, Eric; Dalwani, Manish; Hutchison, Kent; Banich, Marie T. (2007): Prefrontal cortex activity is reduced in gambling and nongambling substance users during decision-making. In *Hum. Brain Mapp.* 28 (12), pp. 1276–1286. DOI: 10.1002/hbm.20344.

Tsai, Hsing-Chen; Zhang, Feng; Adamantidis, Antoine; Stuber, Garret D.; Bonci, Antonello; Lecea, Luis de; Deisseroth, Karl (2009): Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. In *Science* 324 (5930), pp. 1080–1084. DOI: 10.1126/science.1168878.

Tye, Kay M.; Mirzabekov, Julie J.; Warden, Melissa R.; Ferenczi, Emily A.; Tsai, Hsing-Chen; Finkelstein, Joel et al. (2013): Dopamine neurons modulate neural encoding and expression of depression-related behaviour. In *Nature* 493 (7433), pp. 537–541. DOI: 10.1038/nature11740.

U.S. Department of Health and Human Services, National Institutes of (2016): Behavioral Assessment Methods for RDoC Constructs. Available online at https://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/behavioral-assessment-methods-for-rdoc-constructs.

Ungless, Mark A. (2004): Dopamine: the salient issue. In *Trends in neurosciences* 27 (12), pp. 702–706. DOI: 10.1016/j.tins.2004.10.001.

van den Bos, Wouter; Rodriguez, Christian A.; Schweitzer, Julie B.; McClure, Samuel M. (2014): Connectivity strength of dissociable striatal tracts predict individual differences in temporal discounting. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 34 (31), pp. 10298–10310. DOI: 10.1523/JNEUROSCI.4105-13.2014.

van den Bos, Wouter; Rodriguez, Christian A.; Schweitzer, Julie B.; McClure, Samuel M. (2015): Adolescent impatience decreases with increased frontostriatal connectivity. In *PNAS* 112 (29), E3765-74. DOI: 10.1073/pnas.1423095112.

van Holst, Ruth J.; Sescousse, Guillaume; Janssen, Lieneke K.; Janssen, Marcel; Berry, Anne S.; Jagust, William J.; Cools, Roshan (2018): Increased Striatal Dopamine Synthesis Capacity in Gambling Addiction. In *Biological psychiatry* 83 (12), pp. 1036–1043. DOI: 10.1016/j.biopsych.2017.06.010.

van Holstein, Mieke; MacLeod, Paula E.; Floresco, Stan B. (2020): Basolateral amygdala - nucleus accumbens circuitry regulates optimal cue-guided risk/reward decision making. In *Progress in neuro-psychopharmacology & biological psychiatry* 98, p. 109830. DOI: 10.1016/j.pnpbp.2019.109830.

van Schouwenburg, Martine; Aarts, Esther; Cools, Roshan (2010): Dopaminergic modulation of cognitive control: distinct roles for the prefrontal cortex and the basal ganglia. In *CPD* 16 (18), pp. 2026–2032. DOI: 10.2174/138161210791293097.

Vicario, Carmelo M.; Gulisano, Mariangela; Maugeri, Nicoletta; Rizzo, Renata (2020): Delay Reward Discounting in Adolescents With Tourette's Syndrome. In *Movement Disorders* 35 (7), pp. 1279–1280. DOI: 10.1002/mds.28096.

Volkow, Nora D.; Baler, Ruben D. (2015): NOW vs LATER brain circuits: implications for obesity and addiction. In *Trends in neurosciences* 38 (6), pp. 345–352. DOI: 10.1016/j.tins.2015.04.002.

Volkow, Nora D.; Boyle, Maureen (2018): Neuroscience of Addiction: Relevance to Prevention and Treatment. In *American Journal of Psychiatry* 175 (8), pp. 729–740. DOI: 10.1176/appi.ajp.2018.17101174.

Volkow, Nora D.; Wang, Gene-Jack; Baler, Ruben D. (2011): Reward, dopamine and the control of food intake: implications for obesity. In *Trends in cognitive sciences* 15 (1), pp. 37–46. DOI: 10.1016/j.tics.2010.11.001.

Voon, V. (2017): Decision-Making and Impulse Control Disorders in Parkinson's Disease. In *Decision Neuroscience*, pp. 305–314. DOI: 10.1016/B978-0-12-805308-9.00024-5.

Voon, V.; Derbyshire, K.; Rück, C.; Irvine, M. A.; Worbe, Y.; Enander, J. et al. (2015): Disorders of compulsivity: a common bias towards learning habits. In *Molecular Psychiatry* 20 (3), pp. 345–352. DOI: 10.1038/mp.2014.44.

Voon, Valerie; Reiter, Andrea; Sebold, Miriam; Groman, Stephanie (2017): Model-Based Control in Dimensional Psychiatry. In *Biological psychiatry* 82 (6), pp. 391–400. DOI: 10.1016/j.biopsych.2017.04.006.

Wächtler, Christian Ole; Chakroun, Karima; Clos, Mareike; Bayer, Janine; Hennies, Nora; Beaulieu, Jean Martin; Sommer, Tobias (2020): Region-specific effects of acute haloperidol in the human midbrain, striatum and cortex. In *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 35, pp. 126–135. DOI: 10.1016/j.euroneuro.2020.04.008.

Wagenmakers, Eric-Jan; Marsman, Maarten; Jamil, Tahira; Ly, Alexander; Verhagen, Josine; Love, Jonathon et al. (2018): Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. In *Psychonomic Bulletin & Review* 25 (1), pp. 35–57. DOI: 10.3758/s13423-017-1343-3.

Wagner, Ben; Clos, Mareike; Sommer, Tobias; Peters, Jan (2020): Dopaminergic Modulation of Human Intertemporal Choice: A Diffusion Model Analysis Using the D2-Receptor Antagonist Haloperidol. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* 40 (41), pp. 7936–7948. DOI: 10.1523/JNEUROSCI.0592-20.2020.

Waskom, Michael L.; Frank, Michael C.; Wagner, Anthony D. (2017): Adaptive Engagement of Cognitive Control in Context-Dependent Decision Making. In *Cerebral cortex (New York, N.Y. : 1991)* 27 (2), pp. 1270–1284. DOI: 10.1093/cercor/bhv333.

Wassum, Kate M.; Ostlund, Sean B.; Balleine, Bernard W.; Maidment, Nigel T. (2011): Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. In *Learning & Memory* 18 (7), pp. 475–483. DOI: 10.1101/lm.2229311.

Water, Erik de; Cillessen, Antonius H. N.; Scheres, Anouk (2014): Distinct age-related differences in temporal discounting and risk taking in adolescents and young adults. In *Child development* 85 (5), pp. 1881–1897. DOI: 10.1111/cdev.12245.

Weber, S. C.; Beck-Schimmer, B.; Kajdi, M-E; Müller, D.; Tobler, P. N.; Quednow, B. B. (2016): Dopamine D2/3- and μ-opioid receptor antagonists reduce cue-induced responding and reward impulsivity in humans. In *Translational psychiatry* 6 (7), e850. DOI: 10.1038/tp.2016.113.

Weed, M. R.; Gold, L. H. (1998): The effects of dopaminergic agents on reaction time in rhesus monkeys. In *Psychopharmacology* 137 (1), pp. 33–42. DOI: 10.1007/s002130050590.

Westbrook, A.; van den Bosch, R.; Määttä, J. I.; Hofmans, L.; Papadopetraki, D.; Cools, R.; Frank, M. J. (2020): Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. In *Science* 367 (6484), pp. 1362–1366. DOI: 10.1126/science.aaz5891.

Westbrook, Andrew; Braver, Todd S. (2016): Dopamine Does Double Duty in Motivating Cognitive Effort. In *Neuron* 89 (4), pp. 695–710. DOI: 10.1016/j.neuron.2015.12.029.

Westbrook, Andrew; Frank, Michael J.; Cools, Roshan (2021): A mosaic of cost-benefit control over cortico-striatal circuitry. In *Trends in cognitive sciences*. DOI: 10.1016/j.tics.2021.04.007.

Whiteside, Stephen P.; Port, John D.; Abramowitz, Jonathan S. (2004): A meta-analysis of functional neuroimaging in obsessive-compulsive disorder. In *Psychiatry research* 132 (1), pp. 69–79. DOI: 10.1016/j.pscychresns.2004.07.001.

Wiehler, Antonius; Bromberg, Uli; Peters, Jan (2015): The Role of Prospection in Steep Temporal Reward Discounting in Gambling Addiction. In *Frontiers in Psychiatry* 6, p. 112. DOI: 10.3389/fpsyt.2015.00112.

Wiehler, Antonius; Peters, Jan (2015): Reward-based decision making in pathological gambling: the roles of risk and delay. In *Neuroscience research* 90, pp. 3–14. DOI: 10.1016/j.neures.2014.09.008.

Winton-Brown, Toby T.; Fusar-Poli, Paolo; Ungless, Mark A.; Howes, Oliver D. (2014): Dopaminergic basis of salience dysregulation in psychosis. In *Trends in neurosciences* 37 (2), pp. 85–94. DOI: 10.1016/j.tins.2013.11.003.

Wise, Roy A. (1980): The dopamine synapse and the notion of 'pleasure centers' in the brain. In *Trends in neurosciences* 3 (4), pp. 91–95. DOI: 10.1016/0166-2236(80)90035-1.

Wise, Roy A. (2004): Dopamine, learning and motivation. In *Nat Rev Neurosci* 5 (6), pp. 483–494. DOI: 10.1038/nrn1406.

Wit, H. de (2002): Acute Administration of d-Amphetamine Decreases Impulsivity in Healthy Volunteers. In *Neuropsychopharmacology* 27 (5), pp. 813–825. DOI: 10.1016/S0893-133X(02)00343-3.

Wu, Hemmings; Hariz, Marwan; Visser-Vandewalle, Veerle; Zrinzo, Ludvic; Coenen, Volker A.; Sheth, Sameer A. et al. (2020): Deep brain stimulation for refractory obsessive-compulsive disorder (OCD): emerging or established therapy? In *Molecular Psychiatry*. DOI: 10.1038/s41380-020-00933-x.

Wunderlich, Klaus; Smittenaar, Peter; Dolan, Raymond J. (2012): Dopamine enhances model-based over model-free choice behavior. In *Neuron* 75 (3), pp. 418–424. DOI: 10.1016/j.neuron.2012.03.042.

Wyckmans, Florent; Otto, A. Ross; Sebold, Miriam; Daw, Nathaniel; Bechara, Antoine; Saeremans, Mélanie et al. (2019): Reduced model-based decision-making in gambling disorder. In *Scientific reports* 9 (1), p. 19625. DOI: 10.1038/s41598-019-56161-z.

Zhang, Lifen; Doyon, William M.; Clark, Jeremy J.; Phillips, Paul E. M.; Dani, John A. (2009): Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. In *Molecular pharmacology* 76 (2), pp. 396–404. DOI: 10.1124/mol.109.056317.

Contribution Statement

Study 1: Ben Jonathan Wagner, Mareike Clos, Tobias Sommer and Jan Peters: Dopaminergic modulation of human intertemporal choice – a diffusion model analysis using the D2 antagonist haloperidol (2020). The Journal of Neuroscience.

B.W. and J.P. contributed unpublished reagents/analytic tools; B.W. and J.P. analyzed data; B.W. wrote the first draft of the paper; B.W. and J.P. wrote the paper; M.C., T.S., and J.P. designed research; M.C. performed research; M.C. and T.S. edited the paper.

Study 2: Ben Jonathan Wagner, David Mathar and Jan Peters: Gambling environment exposure increases temporal discounting but improves model-based control in regular slot-machine gamblers (2021). Preprint BioRiv

JP conceived the idea and acquired the funding. JP and BJW designed the study. BJW acquired the data. BJW analyzed the data and performed the modeling. DM contributed analytical tools/software. BJW wrote the paper. JP and DM provided revisions. JP supervised the project.

Study 3: Canan Beate Schüller* , Ben Jonathan Wagner* , Thomas Schüller, Juan Carlos Baldermann, Daniel Huys, Julia Kerner auch Koerner, Eva Niessen, Alexander Münchau, Valerie Brandt, Jan Peters and Jens Kuhn (2021): Temporal discounting in adolescents and adults with Tourette syndrome. PLOS One. *These authors contributed equally

C.B.S., T.S., A.M., V.B, J.K., and J.P. conceived the idea. C.B.S., T.S., E.N., A.M., V.B., J.K.K., and J.P. conceived and designed the experiments; C.B.S., E.N., and V.B. performed the experiments; C.B.S., B.W., and J.P. analyzed the data; J.P. and B.W contributed analysis tools; B.W. performed the modelling. C.B.S. and B.W. wrote the paper. All authors reviewed and approved the final manuscript.

Study 4: Ben J. Wagner*, Canan B. Schüller*, Thomas Schüller, Juan C. Baldermann, Sina Kohl, Veerle Visser-Vandewalle, Daniel Huys, Jens Kuhn, and Jan Peters (2020): Chronic deep brain stimulation disrupts the stability of inter-temporal preferences. Preprint BioRiv. *These authors contributed equally

J.K., J.P. C.S. and B.W. designed the experiment. C.S. and B.W. acquired the data. C.S. and B.W. analyzed the data. B.W. performed the modeling and statistical analyses. J.P. and J.K. supervised the project. B.W. and J.P. wrote the paper, and all authors provided revisions.