universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Investigating the Effects of D2 Dopamine Receptor Antagonism on Computational Mechanisms Related to Psychotic Experiences”

verfasst von / submitted by

Mag. Nace Mikuš

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2024 / Vienna, 2024

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 796 605 298
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Psychologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Claus Lamm
Betreut von / Supervisor:	Assoz. Prof. Giorgia Silani, PhD
Betreut von / Supervisor:	Univ.-Prof. Christoph Mathys, PhD

Acknowledgements

Many individuals have contributed significantly to the realization of this thesis, and I would like to extend my gratitude to at least some of them:

To Chris (E.), for showing me that science can be exciting and adventurous. Without you I do not think I would have started this journey in the first place.

To Claus, for being an exemplary mentor and role model, both professionally and personally.

To Giorgia, for showing me that science is also about relationships and about sharing the excitement of discovering something new. You seem to navigate this complex space with ease and authenticity, which is incredibly inspiring.

To Chris (M.), for trusting (in) me. Your attitude and support made this journey much less stressful and much more empowering.

To Hana, Lisa, and Jasminka, for that special place we share.

To Ana, for making this world a better place for me and reminding me that life is so much more than models, papers, and grants.

To my parents, for injecting me with self-confidence that allows for high tolerance for uncertainty and the wisdom to prioritize personal values over societal expectations.

To my thesis reviewers: both of you are individuals that have through your work strongly affected my scientific career and I am honored that you have agreed to spend your time on evaluating my thesis.

Contents

Chapter 1: General Introduction

Current challenges in psychopharmacology

Mental health disorders affect millions of people worldwide. Psychiatric drugs, including antidepressants, anxiolytics, mood stabilizers, and antipsychotics, are widely prescribed in every corner of the world and are believed to be essential in managing the symptoms of these disorders. The development and success of various psychiatric medications has not only provided more effective treatment options for patients but has also reinforced the biological perspective in psychiatry (Harrington, 2019). Psychopharmacology, the scientific discipline of researching "mind-altering" substances, has therefore played a crucial role in establishing psychiatry as a key discipline within medicine.

Yet, the field of psychopharmacology currently faces significant challenges. While there has been substantial progress in understanding the biological underpinnings of mental health disorders, this knowledge has had minimal impact on clinical practices or drug development (Kapur, Phillips, & Insel, 2012; Stephan et al., 2016). Treatments continue to be prescribed primarily on a trial-and-error basis, leading to up to 60% of patients not responding to treatment line (Howes, Thase, & Pillinger, 2022; Potkin et al., 2020; A. J. Rush et al., 2004). This inefficiency results in poorer clinical outcomes and decreased quality of life for many patients. Furthermore, despite significant investments, there have been no major breakthroughs in psychiatric drug development in the last five decades, and many global pharmaceutical companies have been abandoning the field of psychiatry altogether (Abbott, 2011).

One of the main reasons for the limited progress in pharmacological treatment in psychiatry is generally attributed to a lack of clinical tests than would reflect latent disease mechanisms and guide treatment selection (Kapur et al., 2012). The current statistical manuals of classification of psychiatric disorders (Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Classification of Diseases (ICD)) are largely based on the Kraepelinian notion that severe mental illnesses fall into discrete categories ("depression" and "schizophrenia") with a clear line between mental illness and normal functioning (Bentall, 2004). Diagnosis is made from observing symptoms and signs with no regard to the underlying biological abnormalities. Although this approach has standardized diagnostic procedures and increased the reliability and consistency of diagnosis across clinicians it is widely considered to fall short of capturing the complex landscape of mental disorders (Hyman, 2012; Kapur et al., 2012). In most cases, a diagnosis based on DSM or ICD provides no insights into the cause or etiology of the mental health problem and therefore has no predictive validity. Despite being constructed to improve reliability, the test-retest scores for some diagnostic categories are surprisingly low (Freedman et al., 2013) and there is large heterogeneity within and considerable comorbidity across diagnostic categories (Kendler, Zachar, & Craver, 2011; Kessler, Chiu, Demler, & Walters, 2005).

The crucial challenge in translating findings from biological psychiatry to clinical practice is that there is a large explanatory gap between the description of clinical symptoms and the molecular targets of psychiatric drugs (Montague, Dolan, Friston, & Dayan, 2012). To move the field forward, initiatives have been launched that aim to shift the focus away from traditional differential diagnostics towards more transdiagnostic, translational and process-based approaches (Borsboom et al., 2021; Kotov et al., 2017; C. Sharp, Monterosso, & Montague, 2012). One of the most commonly used frameworks for translational research in psychiatry is the integrative framework of the Research Domain Criteria (RDoC), by the National Institute of Mental Health in the USA (Cuthbert & Insel, 2013). The RDoC framework describes mechanisms (such as belief updating, reward processing, and arousal regulation), across the entire range of severity (from normal to abnormal), and varying units of analysis (from self-reports to brain circuits). This approach allows for translations from cellular neuroscience and animal studies to humans. However, the focus is predominantly on biological circuits with a poorly specified

bridge from biological systems to behavior and self-reports. The field of Computational Psychiatry provides the methodological tools and theoretical concepts that help in bridging this gap.

Computational approaches in psychiatry

Computational Psychiatry is an interdisciplinary endeavor that employs computational methods, probability theory, and systems theory to address challenges in psychiatry and psychopharmacology. This includes machine learning methods applied to large datasets with the purpose of tackling a clinically relevant problem, such as prediction of treatment outcomes or optimizing treatment selection (for an example see Chekroud et al., 2016; for reviews see Huys, Maia, & Frank, 2016; Rutledge, Chekroud, & Huys, 2019). Despite significant barriers when it comes to real-world deployment (Koutsouleris, Hauser, Skvortsova, & De Choudhury, 2022; Torous, Myrick, & Aguilera, 2023), there is no doubt that this approach will lead to useful applications and improve evidence-based treatment. However, a completely data-driven approach is theory-agnostic and as such provides little insight into the links between explanatory levels. Also, it largely ignores decades of scientific findings about a particular phenomenon, research question, or clinical challenge.

An alternative approach to is to apply theory-driven models that use mathematical descriptions to formalize relationships between variables at different levels of abstraction. In this approach, the causal structure between various data modalities or latent and observed variables is defined with a probabilistic "generative" model – a joint probability distribution over all observations and model parameters. For example, we can provide a generative description of how brain activity generates brain imaging data, or how internal beliefs of an individual affect the choices they make when interacting with others. The power of the generative framework is that we can simulate data and use synthetic data to formalize theories and generate hypotheses based on those theories. These theories can then be tested on experimental and clinical data through data-driven parameter estimation.

The core assumption underlying this approach is that the distinctive function of the brain is computation. Within this paradigm, the brain's essential task is to use information from the past to form models of the world with the purpose of maximizing positive and minimizing negative experiences. When these computations malfunction this leads to suboptimal or even dysfunctional behavioral patterns. Mental illness occurs when there is a large mismatch between the environmental demands and the brain's computational capacity to adapt to these demands. The overarching task of the field is to define specific computational challenges along which individuals with a particular psychiatric diagnosis can be separated from healthy controls on a dimensional basis.

Describing how the brain performs computations can be elucidated through several levels, as outlined by Marr (2010/1982). After defining the conceptual problem, we can specify the underlying computational task (what information is needed and what computation needs to be performed) and the algorithm that solves this task, either through a complex deterministic "algorithmic" procedure or with approximations. Finally, we can hypothesize about how these processes are implemented in the brain.

Although these levels are hierarchically arranged and somewhat independent, they are interconnected in practice and mutually constrain each other. For instance, the implementational level, which specifies the limits of information flow between various brain areas, restricts the range of algorithms that are feasible for solving a given computational problem. This computational approach forces researchers to explicitly define variables at different levels and articulate the mapping between them and allows to integrate them within a broader biopsychosocial context.

When this conceptual framework is applied to psychopharmacology, it can delineate how alteration in various neuromodulatory pathways relate to disrupted computational processes that give rise to psychopathology. Through such empirical and computational work, we can develop explanations for therapeutic effects of psychiatric drugs and provide insights into potentially novel targets or patients groups.

In this thesis we adopt this approach to try to elucidate the mechanisms of action of antipsychotic drugs. We define several computational phenotypes relevant for psychopathologies observed in patients with psychotic disorders and investigate how these phenotypes are affected by an acute administration of commonly used antipsychotics in healthy volunteers.

Psychotic symptoms and antipsychotics

Severe and chronic forms of mental illness, such as schizophrenia, schizoaffective disorder, and bipolar disorder, remain one of the most challenging areas of medicine today. Patients on the schizophrenia spectrum are characterized by psychotic symptoms, which manifest as profound changes in thinking and perception that are typically disconnected from reality. The most common psychotic symptoms are hallucinations and delusions, defined as false beliefs that are held with high confidence, despite evidence to the contrary. In the context of schizophrenia, psychotic symptoms are classified as positive symptoms (i.e., they represent an excess or distortion of neurotypical functioning) and are complemented by negative symptoms (such as anhedonia, flattening of affect, and social withdrawal, i.e., reduced "normal" functioning).

Psychotic symptoms are usually treated with a variety of treatment approaches, including psychological and psychosocial support, but there is a large clinical consensus that a pharmacological treatment with antipsychotics is essential (Howes & Kaar, 2018). The first antipsychotic, chlorpromazine, was serendipitously discovered in the 1950s, when physicians were looking for drugs to reduce shock after surgery. It was the first drug ever approved as a specific treatment for a mental disorder and paved the way for the modern psychopharmaceutical industry (Harrington, 2019). Today, more than thirty different antipsychotic drugs are available in the clinic (Kapur et al., 2004). Antipsychotics are administered worldwide for the treatment of delusions and other psychotic symptoms and are considered to be effective, with a meta-analytic effect size of 0.45 for the treatment of positive symptoms compared to placebo (Leucht et al., 2017). However, despite their effectiveness, we still have a very poor understanding of why and how precisely they work.

Although antipsychotics differ in the receptors to which they bind, all currently approved antipsychotics block D2 dopamine receptors (Howes, Egerton, et al., 2009). The D2 receptor subtype is one of five dopamine receptor subtypes that are grouped into two families: D1-like receptors (including D1, D4 and D5 subtypes) and D2-like receptors (including D2 and D3 subtypes). Positron emission tomography (PET) studies have shown that for a drug to have a therapeutic response a sufficient occupancy of D2/3 receptors is required (Kaar, Natesan, McCutcheon, & Howes, 2020). However, adequate D2/3 blockade is a necessary but not sufficient condition for a positive clinical response (Howes, Egerton, et al., 2009). In fact, up to one-third of patients with schizophrenia do not respond to the first two antipsychotic treatment regimens despite adequate receptor occupancy (Potkin et al., 2020). Individuals who do not respond to treatment have worse outcomes in terms of symptoms, functioning, and quality of life compared to those who do respond (Elkis & Buckley, 2016). Furthermore, patients who do respond to treatment often have persistent symptoms despite being considered in remission (Schennach et al., 2015). Understanding the mechanisms by which D2 receptor antagonists affect behavioral and psychological processes relevant to patients with psychotic symptoms would undoubtedly help to identify treatment responders, enable patient stratification, and pave the way for more personalized treatment approaches.

Dopamine and the aberrant salience hypothesis

Given the relative success of dopamine antagonists in treating psychotic symptoms in schizophrenia and evidence from post-mortem studies in patients showing increased D2 dopamine receptor levels (Zakzanis & Hansen, 1998), it was initially hypothesized that dopamine receptors are the primary biological cause of psychotic symptoms (Snyder, 1976). This hypothesis also explained the findings that psychostimulants, drugs that increase synaptic dopamine levels, induce psychotic symptoms (Lieberman, Kane, & Alvir, 1987). However, the advent of molecular imaging with PET and

similar chemical imaging methods in the 1980s revealed a different picture. PET studies enable insight into various aspects of dopaminergic function in the living brain. A meta-analysis of PET imaging studies with dopaminergic ligands showed that although D2/3 receptor density is slightly increased in some patients with schizophrenia compared to controls, this is not the case for drug-naïve patients (Howes et al., 2012). Instead, these studies reveal that patients with schizophrenia consistently show an increase in presynaptic dopamine function with a large meta-analytic effect size. Presynaptic dopamine function is measured with PET imaging by radiolabeling the L-dihydroxyphenylalanine, a dopamine precursor. High levels of the precursor imply an increased capacity for dopamine synthesis and release. Using this technique, higher dopamine synthesis in the striatum was found to be elevated in individuals at risk for developing schizophrenia (Howes, Montgomery, et al., 2009), positively correlated with psychotic symptoms regardless whether the primary diagnosis was schizophrenia or bipolar disorder (Jauhar et al., 2017), and significantly predicted the response to antipsychotic treatment in first-episode patients with schizophrenia (Jauhar et al., 2019). Taken together, these findings make a compelling case for a link between elevated presynaptic dopamine and psychotic symptoms but provide no insight into how one leads to the other.

One well-accepted model linking biological abnormalities to symptoms is called the aberrant attribution of salience hypothesis (Heinz, 2002; Kapur, 2003). This idea rests on seminal electrophysiological studies in animals showing that dopamine plays a crucial role in reward processing, motivation, and salience. Schultz and colleagues (Schultz, Dayan, & Montague, 1997) demonstrated that the rapid ("phasic") activation of dopaminergic mesolimbic neurons is remarkably tuned to the mismatch between expected and actual rewards. This has been interpreted as a signal of incentive salience that underpins attention orientation and action selection (Berridge, 2007). A dopaminergic response to a stimulus endows it with value or importance. On this basis, it has been suggested that delusions and other psychotic symptoms are caused by dysregulated dopaminergic signaling, resulting in too much meaning and motivational salience being attached to otherwise irrelevant stimuli. This may result in a world where everything seems to be imbued with meaning, which provokes bizarre explanations. Furthermore, the same dysregulated and noisy dopaminergic activity also implies a lack of sustained signaling for relevant reward-related stimuli, providing an explanation for negative symptoms of schizophrenia psychopathology, such as anhedonia, or avolition (Heinz & Schlagenhauf, 2010).

Empirical support for the aberrant salience hypothesis comes from several studies in patients with schizophrenia that use functional magnetic resonance imaging (fMRI) in conjunction with various learning paradigms, where participants need to learn about associations between actions, stimuli, and outcomes. It has repeatedly been demonstrated that compared to healthy controls, patients and their relatives respond more to stimuli that have low predictive quality, while, at the same time, the responses to relevant stimuli are reduced (Corlett & Fletcher, 2012; Jensen et al., 2008; L. et al., 2010; Murray, Corlett, et al., 2008; Roiser, Howes, Chaddock, Joyce, & McGuire, 2013; Roiser, Stephan, Den Ouden, Friston, & Joyce, 2010). In reversal learning tasks, marked by latent changes in contingencies or hidden states, patients misinterpret random (less likely) outcomes as indicative of a reversal, thus resulting in higher switching behavior (Deserno et al., 2019; Fromm et al., 2023; Schlagenhauf et al., 2014). At the neural level, these findings are reflected in increased striatal responses to neutral cues, which poorly predict rewarding outcomes, and reduced activation in the same region in response to reward-predicting stimuli (Deserno, Boehme, Heinz, & Schlagenhauf, 2013; Deserno, Heinz, & Schlagenhauf, 2015; Murray, Cheng, et al., 2008; Murray, Corlett, et al., 2008; Schlagenhauf et al., 2014). This suggests that excessive dopamine availability may be responsible for both positive and negative aspects of the schizophrenia psychopathology. In support of this, higher presynaptic dopamine synthesis measured with PET imaging in healthy volunteers appears to be both negatively correlated with an appetitive prediction error signal in the striatum and positively correlated with behavioral measures of aberrant salience (Boehme et al., 2015; Schlagenhauf et al., 2013).

To understand how dopamine dysregulation could lead to such a complex array of behavioral and neural alterations observed in patients with schizophrenia, and how these are affected by D2 antagonism, it is necessary to embed these findings within a more comprehensive neurocomputational framework. Here, we adopt the framework of Bayesian inference.

Bayesian inference and the Bayesian Brain

Bayesian inference formalizes the process of updating internal beliefs using normative principles of probability theory. At its core is Bayes' theorem, that specifies the only way in which, given a probabilistic model, probability distributions can be updated in the face of new information. Within this framework, when new input is available a posterior belief is generated by combining the prior belief (the initial expectation) with the likelihood (the probability of that input given the model and the prior belief).

Bayesian accounts of brain function postulate that the brain's fundamental task is to infer the causes of its sensory observations by building predictive models using the principles described in a recently developed framework termed hierarchical Bayesian inference (Friston, Stephan, Montague, & Dolan, 2014). This idea of the brain as a predictive machine can be traced back to work of Helmholtz (1867) and is theoretically supported by systems theory. In particular, the good regulator theorem states that any (self-)regulating system must build a model of that system (Conant & Ashby, 1970). The brain can be seen as a system than attempts to regulate its environment, and therefore needs to build a good model of it (Mathys, 2016).

The environment is infinitely complex, and the observations are often imperfect, or noisy. To account for this the brain must build hierarchical probabilistic models that can generate complex sensory observations. According to Bayesian inference, the crucial challenge that the brain must solve is to infer the most likely cause of its sensory observations, and this can be achieved by inverting its generative model of the environment (Dayan, Hinton, Neal, & Zemel, 1995; Rao & Ballard, 1999).

This inferential process can be formalized by defining beliefs as probability distributions with sufficient statistics, such as the expectation (mean) and the uncertainty associated with this expectation (variance or its inverse, precision). The degree to which new information affects prior beliefs depends on the size of the mismatch between prediction and input ("prediction error"), weighted by various sources of uncertainty. If a belief is held with high precision, new input can largely be ignored and will have minimal effect on belief updating. In contrast, if a belief has low precision, novel unexpected information will profoundly alter the prior belief. This concept of uncertainty or precision-weighting of prediction errors can be applied to hierarchical message passing schemes such as Predictive coding.

The Predictive coding framework is a biologically plausible proposal for how hierarchical Bayesian inference might be implemented in the brain (Friston, 2010). According to this framework, representations in higher levels of cortical hierarchies generate "top-down" predictions of "bottom-up" inputs from the lower levels, all the way down to sensory input. At each level of the hierarchy top-down predictions are compared with bottom-up signals, and prediction errors are sent up the hierarchy to update higher-level representations. The prediction errors at each level are weighted by the precision of the beliefs at the higher level and the reliability (likelihood) of the signal from the lower level (Friston, 2009; Rao & Ballard, 1999). In the context of Predictive coding and hierarchical Bayesian inference more generally, aberrant attribution of salience arises from an increased precision-weighting of prediction errors, caused either by lower precision (higher uncertainty) of beliefs at higher levels of the hierarchy, or by overestimating the precision or reliability of sensory evidence (Adams, Stephan, Brown, Frith, & Friston, 2013; Sterzer et al., 2018).

This notion of a bias towards sensory evidence and away from prior expectations is supported by the observation that patients with psychotic symptoms are less likely to experience visual illusions that are based on prior beliefs about the world (e.g. the hollow mask illusion, where convexity is expected because of the prior belief that light comes from above). Hallucinations could also be

understood as resulting from failures to downregulate self-generated internal sensory signals, such as inner speech (Frith, Rees, & Friston, 1998), or motor actions. More generally, it aligns with the aberrant salience hypothesis outlined above, where patients experience both mental events (thoughts and sensations) and random events in their environment as important occurrences that demand attention (Kapur, 2003). Through persistent bottom-up signaling and failures to form adequate predictions of sensory sensations, individuals with psychotic experiences may resort to theories and explanations that are not falsifiable and therefore resistant to rational argument, while having problems with reality (or source) monitoring (Bentall, Baker, & Havers, 1991; Bentall, Kinderman, & Kaney, 1994).

How is this hierarchical message passing implemented in the brain? Several neuromodulators have been proposed to drive or modulate the information flow between the hierarchies. Bottom-up prediction errors are signaled predominantly via the glutamatergic alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors, whereas the N-methyl-D-aspartate (NMDA) receptors play an important role in synaptic plasticity in the prefrontal cortex and the hippocampus and signal the precision (or confidence) of top-down predictions (Adams et al., 2013). The role of dopamine in this message passing is to modulate the precision of bottom-up signals relative to top-down predictions (Friston et al., 2012). On this basis, psychotic symptoms could result from a variety of perturbations within the hierarchical circuitry, including prefrontal NMDA receptor hypofunction and striatal dopamine dysregulation.

Before outlining how increased presynaptic dopamine levels in the striatum might affect the hierarchical message passing, we will briefly summarize the literature on the prefrontal dysfunction and NMDA hypofunction in patients with schizophrenia and how they relate to symptoms and belief updating. The idea that prefrontal glutamate receptors are involved in schizophrenia aetiology is well established (Howes, McCutcheon, & Stone, 2015). Evidence for specific involvement of glutamatergic NMDA receptors comes from studies showing that NMDA antagonists produce psychological effects that closely resemble both positive and negative symptoms of schizophrenia (Krystal et al., 1994). Acute administration of ketamine, a noncompetitive NMDA antagonist, is known to induce delusion-like experiences in healthy volunteers (Lahti, Koffel, LaPorte, & Tamminga, 1995) and to impair prediction error signalling in learning tasks in the prefrontal cortex in a manner that is comparable to patients with psychotic symptoms (Corlett et al., 2006, 2007). This prefrontal prediction error response has been shown to be related to the severity of delusions in first-episode patients (Corlett et al., 2006), as well as to the distress that healthy individuals feel about their delusional ideas (Corlett & Fletcher, 2012).

There is a well-established connection between prefrontal and hippocampal dysfunction in schizophrenia and cognitive deficits, such as working memory. Within predictive coding, disruption of prefrontal and hippocampal circuits through NMDA (or D1) receptor hypofunction reduces the capacity to form representations and more abstract or conceptual beliefs. This includes building structured internal representations of associative connections in the world. Beyond leading to more general cognitive impairments, this has several other implications for psychopathology. Weak prefrontal representations imply that predictions are less precise, making the world seem more unpredictable or surprising. In this context, a delusional belief may serve as a compensatory mechanism that explains a seemingly chaotic world (Corlett, Taylor, Wang, Fletcher, & Krystal, 2010). Furthermore, in absence of adequate cognitive models, behavior and thought processes may become more habitual. Habitual behavior implies repeating actions that have been rewarded in the past, and avoiding actions that have led to aversive outcomes (Thorndike, 1911). It is considered to be reflexive, fast and inflexible (e.g. taking the same route to work every day without deliberately thinking about where to go). Habitual behavior is contrasted with goal-directed behavior, in which choices are made with a particular goal in mind, often relying on an internal model of the world (Tolman, 1948). In acquisition of novel habits, the control of behavior gradually shifts from relying on internal models through frontal, temporal and ventral striatal brain regions towards more reflexive habitual behavior supported by dorsal striatal circuits (Daw, Niv, & Dayan, 2005). Although habitual behavior can save valuable cognitive resources and time (Kool, Cushman, & Gershman, 2016; Otto, Gershman, & Daw,

2013; Patzelt, Kool, Millner, & Gershman, 2018), it can be maladaptive and lead to entrenched behavioral and thought patterns with undesired consequences. In line with this, a bias towards habitual behavior is associated with a range of psychopathologies related to compulsivity and addiction (Gillan, Kosinski, Whelan, Phelps, & Daw, 2016; Voon et al., 2015; Voon, Reiter, Sebold, & Groman, 2016).

In the case of schizophrenia, a bias towards habitual decision-making has been implicated in both the formation and persistence of delusional beliefs. Irrational belief structures are formed through an inflexible selection of alternative explanations leading to a bias against disconfirming evidence (Woodward, Moritz, Menon, & Klinge, 2008). Once a delusional belief is formed, its use to explain an endogenous and exaggerated prediction error reinforces it through a process similar to memory consolidation (Corlett, Krystal, Taylor, & Fletcher, 2009). Similarly, negative symptoms in schizophrenia and symptoms of depression have been explained by inflexible prior beliefs or cognitive schemas (Huys & Dayan, 2009) in which desirable events are unattainable, or negative events are unavoidable (as in learned helplessness (Amat et al., 2005)). These negative schemas may be coupled with an inability to flexibly search for alternative and imagine more pleasurable options (Hartmann et al., 2015). One study looked at the trade-off between habitual and goal-directed behavior in stable patients with schizophrenia and found increased habitual behavior compared to controls, even after controlling for general intelligence and working memory performance (Culbreth, Westbrook, Daw, Botvinick, & Barch, 2016), however no associations with either negative or positive symptoms were found.

In summary, hierarchical Bayesian inference is a key concept in understanding how the brain integrates incoming signals to update internal beliefs and representations and in this way forms predictions about the future. The brain can implement this Bayesian belief updating through hierarchical message passing where predictions and prediction errors travel in opposite directions. Delusions and other psychotic symptoms are proposed to emerge as an imbalance in this message passing resulting from striatal dopamine dysregulation and abnormal functioning of the hippocampal and prefrontal NMDA receptors. This imbalance could result in an increased precision of bottom-up signalling and a reduced ability of patients to form and maintain higher-level belief representations (internal models or cognitive maps), leading to less accurate predictive signals, reduced goal-directed behaviour, and a greater reliance on established habitual thought processes. In the next section, I outline how these computational substrates relate to increased dopamine synthesis and what role D2 dopamine receptors may play in this.

Dopamine receptors and hierarchical message passing

In general, dopamine's role in hierarchical message passing is to modulate the precision of sensory signals relative to top-down predictions. Evidence for this comes from electrophysiological studies in animals showing that the activity of the mesolimbic dopaminergic neurons is tightly coupled both to the temporal precision of reward prediction and reward uncertainty (Fiorillo, Newsome, & Schultz, 2008; Fiorillo, Tobler, & Schultz, 2003; Tobler, Fiorillo, & Schultz, 2005). However, to understand how increased dopamine affects information propagation we need to outline the function of dopamine receptors within the corticostriatal circuit.

It is well-established that striatal medium spiny neurons expressing dopamine D1 and dopamine D2 receptors exert opposing control over thalamic output to the cortex (Frank, 2005). The two receptor subtypes have been associated with two regulatory pathways of striatal control over cortical output via the thalamus. Striatal neurons on the direct pathway, predominantly expressing D1 dopamine receptors, stimulate thalamic cortical projections, by disinhibiting the basal ganglia control of the thalamus through substantia nigra and globus pallidus. The indirect pathway, involving neurons expressing D2 receptors, inhibits thalamic output to the cortex via an indirect route to the substantia nigra over the globus pallidus and the subthalamic nucleus. Dopamine binding to D1 receptors increases activity in the direct pathway, while activity on D2 receptors suppresses the firing of neurons in the indirect pathway (Burke, Rotstein, & Alvarez, 2017).

Furthermore, data from electrophysiological experiments suggest that phasic bursts of activity of dopamine neurons primarily increase the occupancy of the D1 receptor subtype, whereas slow changes in "tonic", baseline firing are primarily reflected in the occupancy of D2-type receptors, which appear to have higher affinity for dopamine (Dreyer, Herrik, Berg, & Hounsgaard, 2010; Goto & Grace, 2005). Based on this, it is thought that sufficient phasic change in dopamine levels is required to increase the excitability of striatal neurons expressing D1 receptors and propagate reinforcement and action initiation (Frank, Seeberger, & O'Reilly, 2004).

More recently, it has been shown that the mapping between D1 and D2 receptor subtypes and the direct and indirect pathways, while largely true for the dorsal striatum, is not so clear-cut in the ventral striatum. There, neurons in the direct pathway express both D1 and D2 receptors, and either excitation or inhibition of both receptor subtypes increases or decreases motivational (Soares-Cunha, Coimbra, David-Pereira, et al., 2016; Soares-Cunha, Coimbra, Sousa, & Rodrigues, 2016). Furthermore, there is evidence for lateral inhibition, by which the indirect pathway inhibits the direct pathway (Dobbs et al., 2016).

Given this model of corticostriatal circuitry, there are several possible mechanisms by which D2 hyperstimulation resulting from increased presynaptic dopamine availability could affect processes related to precision-weighted belief updating. First, excessive dopamine binding to D2 receptors on the inhibitory pathway could inhibit alternative actions (via the subthalamic nucleus pathway), leading to reduced behavioral flexibility and a reliance on habitual actions. In Bayesian terms, this would cause belief rigidity and increase the precision of prior beliefs. This may explain the impervious nature of delusional beliefs, which would then be rendered looser under tonic D2 antagonism.

The observation that D2 overactivation increases the reliance on habitual behaviors is also consistent with work showing that striatal D2 receptors modulate cortical input relevant for goal-directed behavior (Goto & Grace, 2005). In particular, prefrontal signals are suppressed by D2 agonism and evoked by D2 antagonism. Increased striatal dopamine availability in chronic patients and individuals at risk of developing schizophrenia has been associated with abnormal prefrontal activation and poorer cognitive functioning (Fusar-Poli et al., 2011; Meyer-Lindenberg et al., 2002). Therefore, it is possible that striatal D2 overstimulation impairs prefrontal representations and the formation and use of higher-order models for prediction.

On the other hand, increased tonic dopamine signaling, primarily affecting the D2 receptors, suppresses the lateral inhibition of the direct pathway and might reduce the suppression of otherwise irrelevant signals, leading to unwanted thoughts and behaviors (McCutcheon, Abi-Dargham, & Howes, 2019). Through this mechanism, bottom-up prediction errors, that would otherwise be suppressed, are propagated forward, leading to an overreliance on sensory input.

Although this summary of the corticostriatal circuitry is by no means exhaustive, it serves to demonstrate that there are several potential mechanistic pathways by which excessive dopamine binding to D2 receptors could affect hierarchical message passing. In the following section, I outline several important gaps in the literature and how these are addressed in the subsequent chapters of the thesis.

Outline of the thesis

The main goal of this thesis is to understand the mechanistic effects of antipsychotic medication on the psychological processes and computational mechanisms underlying psychotic symptoms. The hope is that such work will not only contribute to our understanding of psychotic symptoms and the medications we use to treat them, but also enable the development of novel clinical assessments that can be used for mechanism-based treatment monitoring, or even to guide treatment selection. However, several research gaps remain, which I aimed to address together with my co-authors in Chapters 2 to 4.

In Chapter 2, we define several computational mechanisms of belief updating that are altered in individuals with psychotic-like experiences across the clinical severity spectrum. As we review above, a particularly promising framework within which we can conceptualize psychotic symptoms, such as delusions and hallucinations, is through hierarchical Bayesian inference. Within this framework, positive symptoms emerge from two interrelated computational substrates: increased sensory precision and decreased precision of top-down predictions. In learning experiments, this imbalance is reflected in an increased behavioral and neural response to random chance occurrences. In addition, patients with psychotic symptoms also show a reduced response to relevant cues, reflecting belief rigidity and a hyporesponsive motivational system. Previous work has already implicated all three mechanisms in psychosis psychopathology; however, they have never been jointly addressed under one computational framework. In Chapter 2, we first introduce a new model that describes interindividual differences in how participants update beliefs about the unobservable (hidden) states when observations are noisy. The model is a particular instantiation of the hierarchical Gaussian filter (HGF), a family of models that provides a principled description of how agents integrate various sources of uncertainty to update beliefs about unobservable states (Mathys, Daunizeau, Friston, & Stephan, 2011; Mathys et al., 2014). The HGF allows us to describe how participants form beliefs about the underlying cause, about the reliability of their observations, as well as more higher-order beliefs, such as beliefs about the likelihood of latent changes in a reversal learning task. Belief updates across the hierarchy are weighted by the ratio of the precision of top-down predictions and bottom-up signals, making the HGF a good model for testing the predictions about psychotic symptoms that arise from the predictive coding framework.

A major challenge for computational psychiatry has been to define computational mechanisms that are consistently impaired across clinical and non-clinical populations and stable across test sessions. One important source of inconsistency in findings across studies has been the reliance on existing DSM-based disease categories and group comparisons of individuals with and without a formal diagnosis of a mental health disorder (or a cut-off point on a clinical scale). This is in contrasts with the notion that mental health problems are dimensional processes (Gillan & Whelan, 2017; Kotov et al., 2017; Patzelt, Hartley, & Gershman, 2018). Given the high degree of heterogeneity within patients with schizophrenia, it is not surprising that it has been difficult to draw meaningful, replicable conclusions. In Chapter 2, we address this challenge by including individuals with both subclinical delusional ideation and patients with psychotic symptoms.

For any assessment tool that aims to measure trait-like processes relevant to clinical practice, it is a prerequisite that the test has good psychometric properties, such as high test-retest reliability. In experimental research in psychiatry and psychopharmacology it is often implicitly assumed that behaviors measured by experimental tasks and the parameters describing these behaviors reflect measures that are stable within participants. However, the test-retest metrics are either not known or not at an acceptable level (Bland et al., 2016; Enkavi et al., 2019). In the paper presented in Chapter 2, we aim to address this issue by exploring the task conditions that allow for good psychometric scores through data simulation, before conducting a test-retest study with four test sessions. We then investigate how psychotic experiences across the severity spectrum relate to three computational substrates: belief rigidity; precision of higher-order beliefs, and precision of sensory observations.

As outlined above, D2 dopamine receptors are implicated in various aspects of hierarchical inference and D2 receptor antagonism is central to the clinical response of antipsychotics. It might therefore be expected that D2 antagonists would modulate the computational processes described above, thereby facilitating treatment response. Although some studies have attempted to assess the relationship between antipsychotic effects on learning-based paradigms and the clinical effects of antipsychotics (reviewed in Winton-Brown, Fusar-Poli, Ungless, & Howes, 2014), a major limitation of these studies is that several other treatment-related processes might confound the effect of D2 receptor occupancy. In the studies presented in Chapters 3 and 4 we instead conducted two randomized, double-blind, placebo-controlled psychopharmacological studies in healthy volunteers to investigating how

acute D2 antagonism affects belief rigidity (Chapter 3) and the ability to use higher-order representations (cognitive maps) for planning and decision-making (Chapter 4).

Given that dopamine binding to D2 receptors within the corticostriatal circuit reduces the activation of postsynaptic neurons and thalamic output to the cortex, D2 antagonism should increase signal propagation, leading to higher prediction error signaling and more flexible belief updating. If this is the case, the effect of D2 antagonists should also be disproportionately greater in individuals with greater presynaptic dopamine availability. In Chapter 3, we investigated how the antipsychotic drug sulpiride, a highly selective D2-type receptor antagonist, affects learning about the prosocial attitudes of others in healthy volunteers that were preselected to have genetically conferring higher or lower levels of dopamine availability. By explicitly modelling the latent belief trajectories through an HGF model, we investigated how the drug affected belief rigidity when learning about others, and how it interacted with the genetic polymorphism related to presynaptic dopamine availability.

D2 receptors have also been shown to be involved in routing information from striatal to prefrontal circuits and D2 overstimulation might impair the manipulation of information needed for cognitive performance or decision-making. Within hierarchical inference, this might imply an impairment in the use of prior knowledge for prediction. On the other hand, both D1 and D2 overstimulation could also lead to increased habitual decision-making. In Chapter 4, we investigated the effects of amisulpride, another highly selective D2-type receptor antagonist, on the trade-off between goal-directed and habitual decision-making. The distinction between the two types of decision-making has been mapped onto model-based and model-free reinforcement learning components. The model-free system relies on prediction error driven associative learning of action-outcome contingencies, whereas the model-based system makes prospective choices by relying on internal models of the environment facilitated by prefrontal cortical areas (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) including the dorsolateral prefrontal cortex (Smittenaar, FitzGerald, Romei, Wright, & Dolan, 2013). Because of the various pathways through which a D2 antagonist could impact the trade-off between the two systems, we compared its effects to those of an opioid antagonist. The rationale for this was that the opioid system is notorious for hijacking the habitual reward system, while having little direct effect on cognitive functions.

Finally, in Chapter 5 we first summarize the main findings from Chapters 2 to 4, before discussing the recurring themes and the implications of the findings for clinical practice.

Chapter 2: Computational phenotyping of aberrant belief updating in individuals with schizotypal traits and schizophrenia

Non-peer reviewed manuscript as submitted to a peer-review journal

Nace Mikus [1, 2], Claus Lamm[1, *] & Christoph Mathys[2, 3, 4, *]

[1] Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Austria;

[2] Interacting Minds Centre, Aarhus University, Denmark;

[3] Translational Neuromodeling Unit, University of Zurich and ETH Zurich, Zurich, Switzerland;

[4] Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy

*Shared senior authors

Correspondence to: Nace Mikus: nace.mikus@univie.ac.at

Abstract

Psychotic and psychotic-like experiences are thought to emerge from various patterns of disrupted belief updating. These include belief rigidity, overestimating the reliability of sensory information, and misjudging task volatility. Yet, these substrates have never been jointly addressed under one computational framework and it is not clear to what degree they reflect trait-like computational patterns. Here, we introduce a computational model that describes interindividual differences in how individuals update their beliefs about a volatile environment through noisy observations and use this model in a series of studies. In a test-retest study with healthy participants (N = 45, 4 sessions), we find that interclass correlations were moderate to high for session-level model parameters and excellent for averaged belief trajectories and learning rates estimated through hierarchical Bayesian inference. Across three studies (total N=590) we then demonstrate that two distinct computational patterns describe two different transdiagnostic categories. Higher uncertainty about the task volatility is related to schizotypal traits (N = 45, d = 0.687, P = 0.02, N = 437, d = 0.14, P = 0.032) and to positive symptoms in a sample of patients with schizophrenia (N = 108, d = 0.187, P = 0.039), when learning to gain rewards. In contrast, depressive-anxious traits were associated with more rigid beliefs about the underlying mean (N = 437, d = -0.125, P = 0.006) and outcome variance (N = 437, d = -0.111, P = 0.013), as were negative symptoms in patients with schizophrenia (d = -0.223, P = 0.026, d = -0.298, P = 0.003), when learning to avoid losses. These findings suggest that individuals high on schizotypal traits across the psychosis continuum are less likely to learn or utilize higher-order statistical regularities of the environment and showcase the potential of clinically relevant computational phenotypes for differentiating symptom groups in a transdiagnostic manner.

Introduction

A computational phenotype is a computational parameter that provides a trait-like, mechanistic description of a psychological, behavioral, or neural process. As such, computational phenotypes can enable transdiagnostic assessments of psychiatric symptoms that span many levels of analysis and have the potential to guide psychiatric treatment (Paulus, Huys, & Maia, 2016; Stephan & Mathys, 2014). One area where computational approaches have proven particularly promising is in describing neurocomputational mechanisms of how humans and non-human animals infer latent unobservable states from sensory observations. In particular, a rich body of theoretical and empirical work supports the notion that schizotypal traits across the severity spectrum can be cast as aberrant inference or learning (Adams et al., 2013; Corlett, Frith, & Fletcher, 2009; Fletcher & Frith, 2009; Stephan, Baldeweg, & Friston, 2006).

Under normative accounts of statistical inference, the degree to which new observations reflect meaningful occurrences in our environment should depend on several distinct sources of uncertainty, such as how reliable we feel our observation was (outcome uncertainty), or how likely we believe such an occurrence to be (environmental uncertainty). This weighting of relative uncertainties (or their inverses, precisions) when integrating sensory information with prior expectations has been referred to as precision-weighted belief updating (Friston, 2009; Mathys et al., 2011). Overestimating either the precision of sensory sensations or the probability of unlikely events can both lead to attributing too much importance to random chance occurrences. Within this framework, the computational substrate underlying psychotic symptoms has been suggested to arise from this aberrant attribution of salience to internal sensations and random external events, leading to a psychological state where the world seems imbued with significant experiences that provoke bizarre explanations (Fletcher & Frith, 2009; Jaspers, 1997; Kapur, 2003).

The challenge of separating outcomes arising from noise in observation from outcomes that reflect latent environmental changes has been operationalized in probabilistic reversal learning tasks. Empirical research using various variations of learning tasks has shown that patients with schizophrenia tend to update their beliefs more rapidly (Garety, Hemsley, & Wessely, 1991; Morawetz, Bode, Baudewig, Jacobs, & Heekeren, 2016), overestimate the likelihood of a contextual change (Kaplan et al., 2016), and are more likely to switch between choices in two-choice tasks (Fromm et al., 2023; Schlagenhauf et al., 2014). Computational work implies that these behavioral patterns result from increased beliefs (and uncertainty) about the task volatility (Deserno et al., 2019; Henco et al., 2020; Reed et al., 2020); however, several important issues remain to be resolved.

First, all computational work in this domain only investigated either beliefs about volatility or beliefs about observation noise and has not accounted for both simultaneously and within one common framework. This is an important limitation because the two mechanisms are computationally interdependent, and abnormalities in one might be attributed to abnormalities in the other (Piray & Daw, 2021). They might also represent two different factors of the illness and recruit different brain circuits (Iglesias, Tomiello, Schneebeli, & Stephan, 2017; Sterzer et al., 2018). Second, findings of increased updating in patients with schizophrenia are hard to reconcile with a range of studies showing patients demonstrate reduced belief updating and impaired flexibility (Baker, Konova, Daw, & Horga, 2019; Doll et al., 2014; Nassar, Waltz, Albrecht, Gold, & Frank, 2021; Waltz & Gold, 2007). Third, a major setback in defining outputs of computational models as clinically relevant phenotypes has been that their test-retest reliability is either unknown (Browning et al., 2020), or modest at best (Enkavi et al., 2019). Despite recent progress demonstrating fair reliability of behavioral patterns in reinforcement learning tasks (Loosen, Seow, & Hauser, 2022; Waltmann, Schlagenhauf, & Deserno, 2022), test-retest scores of parameters representing beliefs about more abstract task features, such as task volatility or outcome variance, are simply not known. Finally, it is not clear to what degree the proposed belief updating patterns are specific to positive symptoms, rather than the syndrome of schizophrenia more generally, and whether they represent transdiagnostic patterns on a whole spectrum of clinical severity.

The present study aimed to address these issues. We first introduced a model that describes how individuals dynamically update their beliefs about environmental volatility as well as observational noise when both can dynamically change throughout the task. We then used the generative computational framework to determine the minimal task conditions under which satisfactory reliability of computational parameters is feasible. We then empirically investigated whether the parameters of the computational model are reliable and clinically meaningful measures. To this end, we first conducted a test-retest study in healthy volunteers and looked at which of the three computational patterns predicts subclinical delusional ideation (Experiment 1). We then replicated and extended the findings of Experiment 1 in a larger data set of healthy participants (Experiment 2), and a data set of patients with schizophrenia (Experiment 3).

Methods and Materials

Computational Modelling

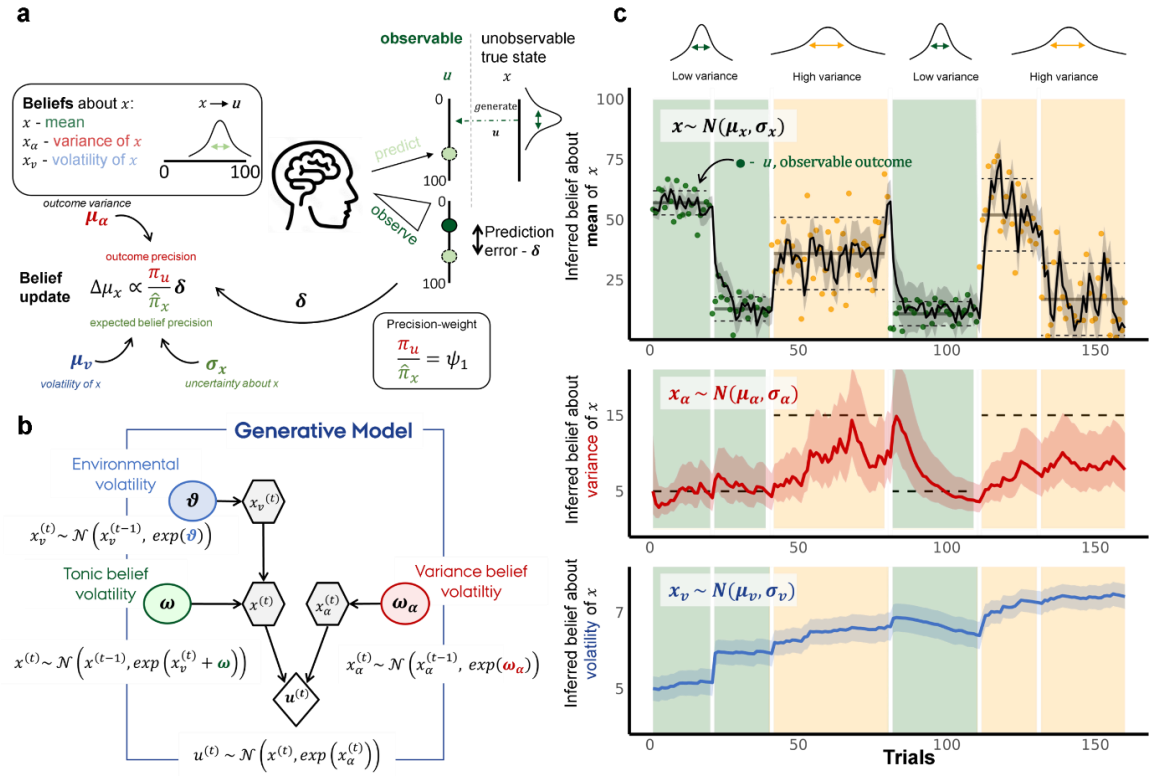


Figure 1. Computational model. **a**, Observations are generated from an unobservable dynamically changing Gaussian variable with varying variance. Observers make predictions based on their beliefs and use the prediction errors to update their beliefs about the underlying states. Belief updates are regulated by the precision weight ψ_1: ratio of the outcome precision (π_u) and the expected mean belief precisions (π_x), which is in turn governed by both prior belief precision or uncertainty σ_x and belief about volatility μ_v. **b**, Generative model of the observer is defined through two cascades of random Gaussian walks that determine the evolutions of beliefs about the mean (and its volatility) and the outcome variance. **c**, Model inversion through the hierarchical Gaussian filter provides both the mean and the precision of beliefs about the mean, variance, and volatility of the latent variable.

The fundamental problem that we need to solve is to provide a principled description of how an agent updates their beliefs about a latent cause in an environment where the (sensory) observations are noisy and the underlying cause changes, whereby this agent's behavior is determined with a set of specific parameters that could account for interindividual differences. In one simplified example of such an environment, the probabilistic model used to generate outcomes (observations) is a Gaussian variable, where both summary statistics are stable for several trials but can change anytime (Figure 1a). A task with such an outcome generating model has been introduced as the Predictive inference task

(Nassar, Wilson, Heasly, & Gold, 2010), whereby one includes reversals in the degree of observational noise within one task session.

The model we present here is based on the generalization of hierarchical generative models invertible by the hierarchical Gaussian filter (HGF, Mathys, Daunizeau, Friston, & Stephan, 2011; Mathys et al., 2014). We define the generative model (Figure 1b) as two parallel cascades of Gaussian random walks, where the first governs the evolution of the underlying mean (x), and its (log) volatility (x_v), and the second governs the evolution of the outcome variance (x_α), also on a log scale. The generative model of what causes sensory observations (u) is then defined as the mean state x and the variance state x_v (Figure 1b). Importantly, the evolutions of the three described states are governed by three agent-specific parameters: variance volatility parameter ω_α (the rate of change of the variance state), environmental volatility parameter ϑ (the rate of change of the volatility of the mean), and tonic volatility parameter ω (the rate of change of the mean x, after accounting for the environmental volatility).

The generative model is inverted using hierarchical Gaussian Filtering (HGF, Mathys, Daunizeau, Friston, & Stephan, 2011; Mathys et al., 2014) to derive at trial-level update equations that resemble other prediction error driven updating, such as reinforcement learning (Sutton, 1992) or the Kalman filter (Gershman, 2015; Kakade & Dayan, 2002). The model inversion is described in detail in work introducing the HGF (Mathys et al., 2011, 2014) and it's more generalizable form (L. A. Weber et al., 2023). Through model inversion we derive inferred posterior distributions of participants' belief trajectories as Gaussians with the mean μ and variance σ^2 or its inverse, the precision π (see Supplementary Table 1 for update equations). The likelihood function (or the action model) maps beliefs about the mean to behavior (y) with a Gaussian function:

$$y^{(t)} = N\left(\mu^{(t)}, \exp(\eta)\right),$$

where η represents decision noise, or measurement error (on a log scale).

Parameter estimation

Initial parameter estimation and simulations were done for each individual and session separately using maximum a posteriori estimates and the HGF toolbox in Matlab (Mathworks) and is available at www.translationalneuromodeling.org/tapas as open-source code (GPLv3).

In the test-retest study, we first analyzed behavior without pooling. We compared three different models using the Random-Effects Bayesian Model Selections (Stephan, Penny, Daunizeau, Moran, & Friston, 2009). The priors used for the models are described in Supplementary Table 2.

Because hierarchical inference can improve estimation of parameters across individuals and at the group level (Waltmann et al., 2022), we also analyzed our data within a hierarchical Bayesian framework. Models were implemented in custom code in Stan (Carpenter et al., 2017) using R as the programming language and RStudio as the integrated development environment for R (for the full model definition see Supplementary Methods). Models were estimated using Markov Chain Monte Carlo sampling, with four independent chains and 3000 iterations (1000 warm-up). Convergence of sampling chains was estimated through the Gelman-Rubin $\hat{R}$ statistic (Gelman & Rubin, 1992), whereby we considered $\hat{R}$ values smaller than or equal to 1.01 as acceptable.

The intraclass correlation (*ICC*) for the parameters was defined as the ratio of between-participant variance and total variance and estimated either with custom code or using the *icc* function from the R package *psych* (using bootstrapping with 1000 samples to obtain distributions).

Simulating test-retest scores

To calculate test-retest we drew 50 sets of model parameters from the prior distributions (Supplementary Table 2) and set these as unobservable "true" parameters. We generated 2 pairs of task

trajectories outcomes (observable inputs) for five different trial numbers (16, 48, 120, 240, and 480). For each of the 50 computational parameter sets we simulate data with three noise levels on two different task versions and estimate the parameters. The re-estimated parameters are then used to calculate the simulated ICC using the *psych* package in R.

Experiment 1

Participants and procedures

Fifty-five (n = 55) healthy adults took part in four in-person testing sessions separated by at least 1 week, out of which n = 45 finished all four sessions and were included in the final analysis. On each session participants performed a behavioral battery that included the Predictive inference task, as well as an exploration-exploitation task and a losing of control task (Mikus, Mancinelli, Lamm, & Mathys, n.d.) that will be published elsewhere. On top of this, in the third session, participants filled out the Peters et al. Delusion Inventory (PDI, Peters, Joseph, Day, & Garety, 2004) and in the fourth session they played the Two-step task (Daw et al., 2005; Kool et al., 2016). The participants earned a base fee of € 40 and could earn up to € 80 through bonus payments from the tasks. All participants gave written informed consent, and the experiment was approved by the ethics committee of the University of Vienna (# 1918/2015) and conducted following the latest revision of the Declaration of Helsinki (World Medical Association, 2013).

The predictive inference task

For this experiment we adapted the Predictive inference task from Nassar et al. (2010). In this task, participants are required to predict the outcome of each trial. The outcome is a random number between 0 and 100, drawn from a normal Gaussian distribution with a certain mean and standard deviation, with a rebounding boundary. In contrast to the task by Nassar and colleagues, both the mean and the standard deviation change several times throughout the task. The probability of the mean changing is 0.1 except for the first 3 trials of each block (as in Nassar et al. 2010, 2012) and the probability of the standard deviation to change when the mean changed is 0.4, and could be either high (15) or low (5). Participants could express their confidence with a longer button press and performance was incentivized based on prediction errors. Participants played 240 trials in four sessions with four different trajectories that did not change across participants. For a detailed description of the task see Supplementary Methods.

The two-step task

The two-step task is often used as a behavioral measure for the capacity for model-based decision making (Daw et al., 2005; Kool et al., 2016), whereby keeping a "model" of the mapping between step one and step two can improve points earned in the task. For a detailed description of the task refer to the Supplementary Methods. For 5 participants, the data of the task was not saved, therefore, a total of 40 participants were included in the analysis of this data.

Behavioral analysis

For frequentist mixed linear models, the lme4 package in R was used, and for Bayesian mixed linear models, the brms package was used. For frequentist models we report Confidence Intervals (CI)s and p values (P). for Bayesian models we report Credibility intervals (CrI) and the probability of the posterior interval that lies above (or below) zero (P). For details on behavioral analysis see Supplementary Methods.

Experiment 2

In the second experiment, we reanalyzed data from a previously published online study (for details see Seow & Gillan, 2020). Here, a sample of 437 (249 female) MTurk participants, aged 20–65 (mean = 36.3. SD = 10.2) played a version of the Predictive inference task with constant noise and varying volatility (see Supplementary Methods for details) and filled out a wide range of questionnaires

assessing general psychopathology including the Short Scales for Measuring Schizotypy that consists of four subscales (Mason, Linney, & Claridge, 2005). As described in Seow and Gillan (2020), all questionnaires together consisted of 209 items, and weights from a factor analysis done in a previous study (Gillan et al., 2016) were used to define three factor scores named as "compulsive behavior and intrusive thought" (CIT), "anxious-depression" (AD), and "social withdrawal" (SW).

Experiment 3

In the third experiment, we reanalyzed data from a previously published study with patients with schizophrenia (Nassar et al., 2021). Data was used from 108 participants with a diagnosis of schizophrenia or schizoaffective disorder collected at the Maryland Psychiatric Research Center, University of Maryland School of Medicine. Participants were stable outpatients, most treated with antipsychotic medications. Inclusion criteria in patients were a presence of a schizophrenia spectrum disorder in patients. Exclusion criteria were a presence of a current Axis I disorder, a neurological disorder, or a cognitively impairing medical disorder. The severity of positive and negative symptoms were assessed with the Brief Psychiatric Rating Scale (BPRS, (Overall & Gorham, 1962)), and Scale for the Assessment of Negative Symptoms (SANS, (Andreasen, 1989)), respectively. In the predictive inference task performance was incentivized in two different ways across two different task sessions, each lasting for 100 trials. In the appetitive condition, participants could earn points and in the aversive condition, participants would have points taken away from their initial endowment in a performance contingent way (see Supplementary Methods for details).

Results

Computational Model

The three model parameters ω, ω_α and ϑ determine interindividual (and inter-session) differences in how participants update beliefs about the mean (μ), the outcome variance (μ_α) or the task volatility (μ_v), by specifying the precision weights on the respective belief trajectories (see update equations Supplementary Table 1). Low values of ω, ω_α and ϑ imply rigid beliefs about the mean, the outcome variance or environmental volatility, respectively (see Supplementary Figure 1 for model simulations with various combinations of parameter sets).

To make sure the three parameters correspond to distinct behavioral patterns we looked at parameter recovery and found high correlations between simulated and retrieved parameters under low noise assumption, with 480 trials ($r>0.96$). However, as decision noise increases, or as trial numbers decrease, parameter recovery becomes harder (Supplementary Figure 2).

Next, we aimed to determine under which conditions the reliability of computational parameters in an experiment would be sufficient. In other words, when can a task generating process (a jumping Gaussian in our case), coupled with a model, produce computational parameters that show good intra-participant reliability across testing sessions (Figure 2a). ICCs were dependent on trial numbers and decision noise and were generally high for the tonic belief volatility parameter and more varied for the environmental and variance volatilities. Yet, with sufficient trials and low decision noise, parameters are retrievable even across tasks with different trial numbers (Supplementary Figure 3). In contrast, when looking at simulated ICC scores for binary data with binary outcomes, using a binary version of the HGF often used modelling behavior in reversal learning tasks, we found that even with high trial numbers and low noise, environmental volatility does not have satisfactory ICC scores (Figure 2b).

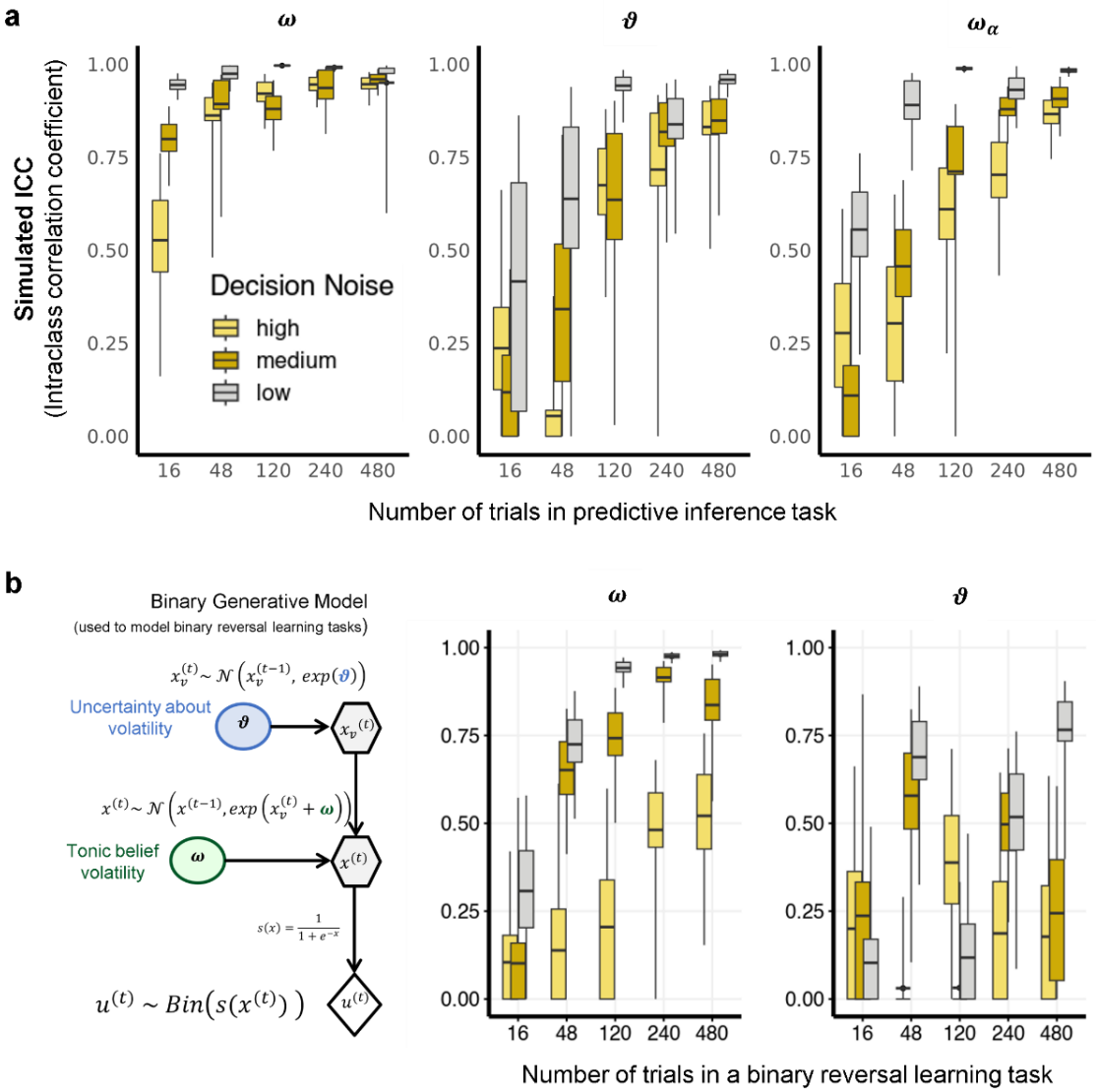


Figure 2. Simulated intraclass correlations. a, Continuous predictive inference task simulation with decision noise η (1, 16, or 48). **b**, Binary reversal one-armed bandit task, where the trajectories (correct/incorrect) are generated with the same probability of reversal as in (a) and the probability of correct in each block being either 0.2, 0.8, or 0.5. The decision noise is parametrized through the softmax inverse temperature set as 48, 8, or 1 (from low to high noise).

Experiment 1

Model and behavior

We first wanted to make sure that participants in our task tend to adapt their learning rate according to higher variances in outcomes. As above, we define the learning rate as the ratio between the prediction update and the prediction error. We found that participants on average reduced their learning rate in high compared to low outcome variance blocks (b =-0.167, SE = 0.011, d= 0.168, t(139.34) = -15.761, P<0.001, Figure 3a). We compared the model with all three parameters to models that would (correctly) assume that the environmental volatility does not change. We found that the model with all three parameters outperformed the other two models in all sessions (Supplementary Table 3), suggesting that on average participants tended to adjust their beliefs about the volatility of the mean throughout the task. In line with this, we found that the (model implied) mean beliefs about volatility were higher for high compared to low outcome variance blocks (b = 0.108, SE = 0.015, d =

0.190, t(55.245) = 7.199, P < 0.001, Figure 3b), as were mean beliefs about noise (b = 0.14519, SE = 0.01588, d = 0.265, t(52.504) = 9.141, P<0.001, Figure 3c).

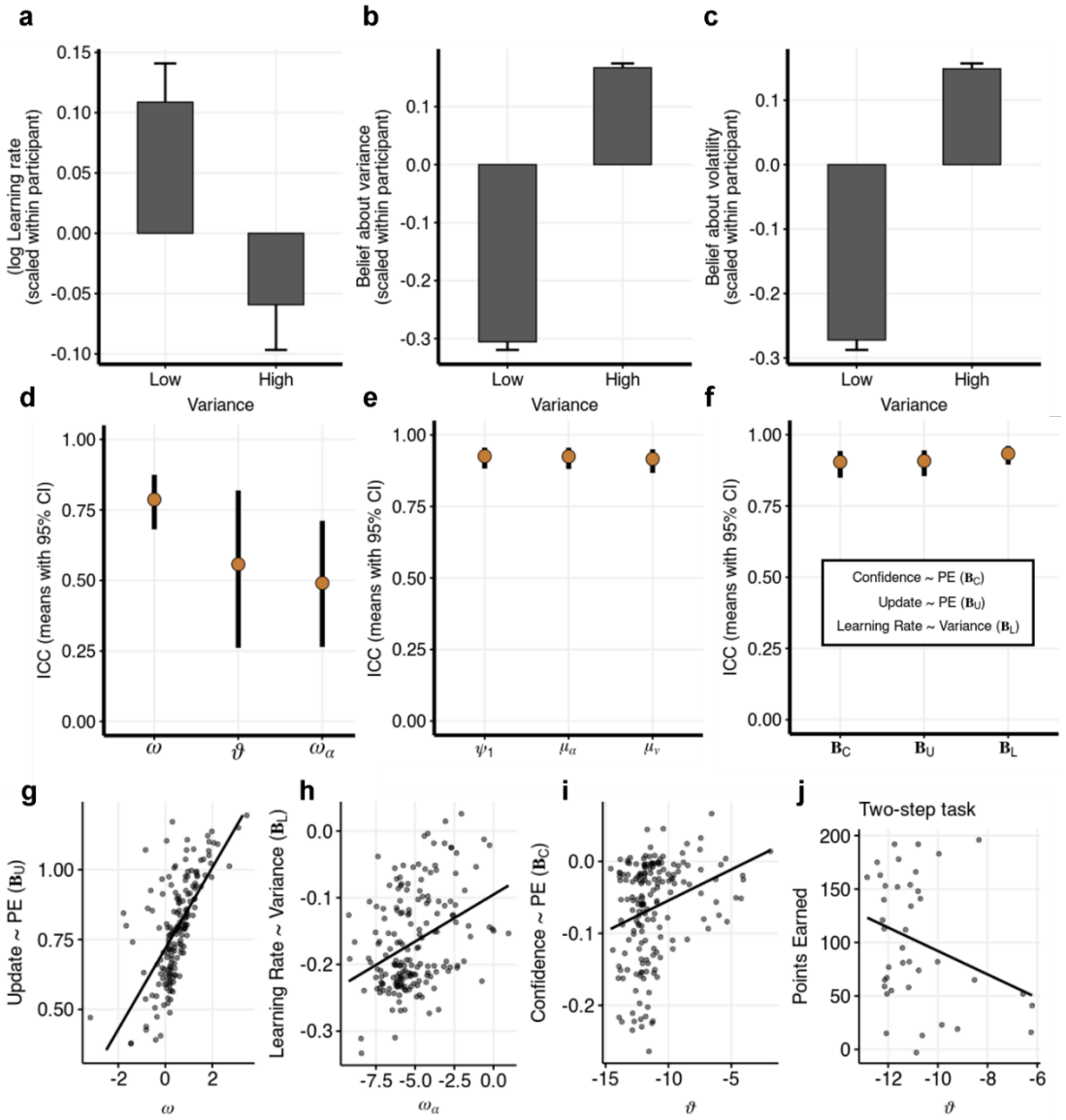


Figure 3. Model performance in Experiment 1. a – c, Learning rate adaptation (a), beliefs about variance (b) and volatility (c) across the variance blocks. All dependent variables were scaled within participant, then averaged within participant and variance block and then across participants. Bar plots and error bars represent means and standard errors across participants (with N = 45). **d**, Empirical ICC scores derived from the hierarchical model (means and 95% CrI). **e**, ICCs (means and 95% CI) of model derived trial-by-trial precision-weighted learning rates (ψ_1), beliefs about outcome variance (μ_α) and volatility (μ_v) meaned within participant. **f**, ICCs (means and 95% CI) plotted for the participant-level random slope of prediction error effect on confidence on the next trial ($\boldsymbol{B}_C$), singed prediction error on signed update ($\boldsymbol{B}_U$), and the effects of variance on learning rate ($\boldsymbol{B}_L$). **g – i**, Relationship of random slopes $\boldsymbol{B}_U$, $\boldsymbol{B}_L$, and $\boldsymbol{B}_C$ on model parameters. **j**, Correlation between the performance in the Two-step task and the environmental volatility parameter ϑ in session 1 (N = 40).

Test-retest

ICC scores (derived from the hierarchical models as ratios of posterior distributions of between-participant variance and total variance) were the following: for ω the mean was 0.787 (95% CrI [0.681, 0.874]), for ϑ the mean was 0.557 (95% CrI [0.261, 0.819]) and for ω_α the mean was 0.491 (95% CrI [0.264, 0.711], Figure 3d), indicating at least moderate ICCs. However, when looking at the ICC scores of averaged trial-by-trial trajectories calculated with a separate linear model (Figure 3d), we found excellent test-retest scores for the precision-weighted learning rates ψ_1 (ICC = 0.93, 95% CI [0.88, 0.96], for mean beliefs about volatility μ_v (ICC = 0.92, 95% CI [0.87, 0.95]), as well as noise μ_α (ICC

= 0.92, 95% CrI [0.88, 0.96]). The high ICCs are likely because hierarchical estimation pulls estimates within participants closer together. Note that the ICC scores are lower (in particular for the μ_v) when parameter estimation was done without pooling, or when looking at only two sessions at a time (Supplementary Figures 4, 5 and 6).

In the next step, we looked at how behavior in the task related to model parameters, and how stable it was across sessions. We defined three participant-level behavioral patterns as participant-level regression slopes of 1) the model predicting updates on the next trial from signed prediction error ($\boldsymbol{B_U}$); 2) the model estimating how the behavioral learning rate is affected by outcome variance ($\boldsymbol{B_L}$); and 3) the model predicting confidence ratings on the next trial from absolute (unsigned) prediction errors ($\boldsymbol{B_C}$). We found that the all three participant-level random slopes have excellent ICC (all mean ICC > 0.9, Figure 3e). Looking at how the three model parameters predict each of the three behavioral patterns, we found B_U was most strongly predicted by the belief volatility parameter ω (b = 0.795, SE = 0.053, t = 14.884, P < 0.001, Figure 3g, Supplementary Table 4), B_L correlated only with the variance belief volatility parameter ω_α (b = 0.438, SE= 0.115, t = 3.822, P < 0.001, Figure 3h, Supplementary Table 5), and B_C correlated ϑ (b = 0.356, SE = 0.111, t = 3.202, P = 0.002 r = 0.280, t(178) = 3.888, P < 0.001, Figure 3i, Supplementary Table 6).

The environmental volatility parameter represents the amount of uncertainty participants have about how volatile the environment is and is conceptually related to the ability to form and maintain high-order models about the world. This is comparable to the concept of model-based learning in the reinforcement learning framework. In support of this, we found some evidence that the performance of participants in the two-step task (points earned) correlated negatively to ϑ ($r_{session1}$ = -0.371, P = 0.048, $r_{session2}$ = -0.268, P = 0.097, $r_{session3}$ = -0. 317, P = 0.049, $r_{session4}$ = -0.226, P = 0.101, N=40 Figure 3j), but not with ω_α (all P >0.4), nor ω (all P > 0.16).

Schizotypal traits correlate with the environmental volatility parameter

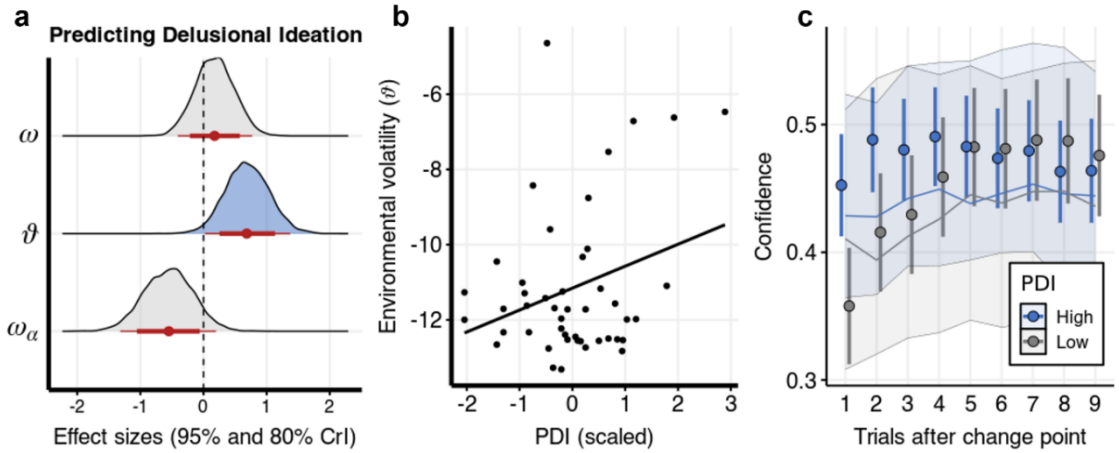


Figure 4. Model parameters and schizotypal traits. a, Posterior distributions of effect sizes of the effect of each model parameter (meaned per participant) predicting PDI, over means and CrIs. **b,** Correlation plot of ϑ and PDI. PDI was log transformed and scaled. **c,** Plotting confidence ratings after change points. Points and error bars represent means and standard errors (across participants). Line and shade represent the posterior predictive mean and 95% Predictive interval of a linear regression model predicting confidence from prediction errors. PDI: Peters et al. Delusional Inventory.

When looking at how the mean parameters across all four sessions predict PDI scores (Figure 4 a)we found that higher PDI was related to increased environmental volatility parameter ϑ (effect size d = 0.687, 95% CrI [0.027, 1.373], P(d<0) = 0.02, Figure 4 b), and no evidence for a decrease in the noise volatility parameter ω_α (d = -0.548, 95% CrI [-1.317, 0.198], P(d>0) = 0.072) nor on the tonic volatility parameter ω (d = 0.176, 95% CrI [-0.407, 0.771], P(d<0) = 0.286, for session-by-session analysis, and results of the analysis with parameters estimated without pooling, see Supplementary Table 7).

We then investigated how these effects are mirrored in the behavioral patterns described above. Based on the results from the previous section we expected that PDI scores would decrease the effect of prediction errors on the confidence rating in the next trial (B_C). However, we found no robust evidence of PDI's effect on confidence behavior ($b_{PE:PDI}$ = 0.009, 95% CrI [-0.011, 0.029], P(b<0) = 0.186). Plotting the posterior predictions of the model we see that the model poorly captures confidence ratings after change points with a high uncertainty in predictions (Figure 4 c).

Experiment 2

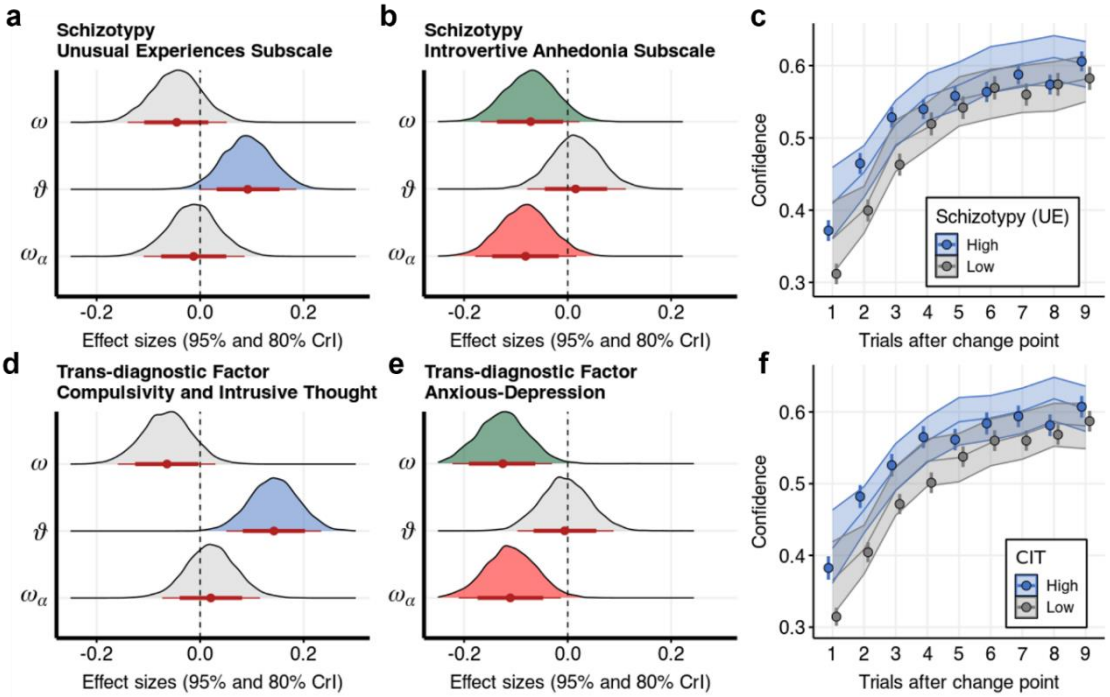


Figure 5. Modelling results from an online sample of healthy participants. a-b, Two subscales of the Short Scale for measuring Schizotypy were predicted by the three parameters controlling for age, gender and IQ. Plotting posterior distributions of effects of each model parameter, over means and CrIs of effect sizes. **c,** Plotting confidence ratings after change points for the median split of the Unusual Experiences scale. Points and error bars represent means and standard errors (across participants). Line and shade represent the posterior predictive mean and 95% Predictive interval from a linear regression model predicting confidence from prediction errors. **d-e**, CIT and AD predicted by model parameters controlling for age, gender and IQ in two separate models. Plotting posterior distributions of effects of each model parameter over means and CrIs of effect sizes. **f,** Plotting confidence ratings after change points. Points and error bars represent means and standard errors (across participants). Line and shade represent the posterior predictive mean and 95% Predictive interval of a linear regression model predicting confidence from prediction errors.

We analyzed the data from (Seow & Gillan, 2020) with the same computational model as in Experiment 1 and investigated how the model parameters relate to schizotypal traits. Predicting the total scores of the Short Scale for Measuring Schizotypy from model parameters, controlling for age, gender, and IQ we found positive correlations with the environmental volatility parameter ϑ (d = 0.14, 95% CrI [-0.007, 0.287], P(b<0) = 0.032, Supplementary Figure 8a), but not with ω_α (P(d>0) = 0.45), nor ω (P(d>0) = 0.171). We then further explored how our parameters predict two subscales of schizotypal traits: "Unusual Experiences" (UE) and "Introvertive Anhedonia" (IA, see results for the other two subscales in the Supplementary Figure 8b). Predicting the scores in the UE subscale (Figure 5a, Supplementary Table 8) we found positive correlations with the environmental volatility parameter ϑ (d = 0.092, 95% CrI [-0.001, 0.186], P(d<0) = 0.026), but not with ω_α (d = -0.013, 95% CrI [-0.109, 0.086], P(d>0) = 0.401), nor ω (d = -0.045, 95% CrI [-0.14, 0.051], P(d>0) = 0.171). In contrast, the IA subscale (Figure 5b, Supplementary Table 9) did not correlate reliably with ϑ (d = 0.029, 95% CrI [-0.066, 0.123], P(d<0) = 0.286), but we found weak evidence of a negative correlation with ω_α (d = -0.081, 95% CrI [-0.179, 0.017], P(d>0) = 0.053) and ω (d = -0.071, 95% CrI [-0.168, 0.023], P(d>0) =

0.071). When investigating how both subscales modulated the effect of prediction error on confidence ratings on the next trial (Figure 5c, Supplementary Table 10), we found a significant interaction effect of UE (d = 0.035, 95% CrI [0.015, 0.055], P(d<0) < 0.001), but not IA (d = 0.01, 95% CrI [-0.009, 0.029], P(d<0) = 0.154). Interestingly, we also found that higher UE scores predicted higher confidence ratings on average (d = 0.151, 95% CrI [0.102, 0.198], P(d<0) = 0.001), whereas IA scores predicted lower confidence on average (b = -0.051, 95% CrI [-0.096, -0.001], P(b>0) = 0.023).

We then investigated to what degree these effects are specific to traits related to schizotypy. We looked into how our model parameters predicted the first two factors of the factor analysis, named as "compulsive behavior and intrusive thought" (CIT), and "anxious-depression" (AD). In a model that predicted the CIT factor from model parameters, controlling for age, gender and IQ, we found a positive effect of the ϑ parameter (d = 0.143, 95% CrI [0.05, 0.234], P(d<0) = 0.001, Figure 5d, Supplementary Table 11), but not ω (d = --0.064, 95% CrI [-0.159, 0.029], P(d>0) = 0.088), nor ω_α (d = 0.021, 95% CrI [-0.073, 0.116], P(d<0) = 0.331). In contrast, the AD factor was negatively related to ω (d = -0.125, 95% CrI [-0.222, -0.03], P(d>0) = 0.006, Figure 5e, Supplementary Table 12), as well as ω_α (d = -0.111, 95% CrI [-0.21, -0.013], P(d>0) = 0.013), but not ϑ (d = 0.006, 95% CrI [-0.097, 0.089], P(d<0) = 0.457).

We also found that the factors modulated the effect of prediction error on confidence ratings on the next trial (Figure 5f, Supplementary Table 13), with CIT increasing confidence on average (d = 0.22, 95% CrI [0.168, 0.271], P(d<0) < 0.001), as well as having a positive interaction with the PE (d = 0.058, 95% CrI [0.037, 0.08], P(d<0) < 0.001), and AD decreasing confidence on average (d = -0.098, 95% CrI [-0.154, -0.045], P(d>0) < 0.001), but did not modulate the effect of the prediction error (d = 0.008, 95% CrI [-0.015, 0.032], P(d<0) = 0.256).

Experiment 3

To see to what extent these results extend to the clinical domain of positive and negative symptoms of schizophrenia we analyzed the behavior of 108 patients with schizophrenia (previously published in Nassar et al., 2021). The behavior in the task was analyzed with the same model as in Experiments 1 and 2, whereby we estimated the parameters for the two incentivization conditions (reward seeking and loss avoiding) separately. We first looked at how well the parameters of the model predict positive symptoms, controlling for age, gender, and IQ. Because patients might display a high degree of noise in their behavior, we also included decision noise as a predictor in the models.

Using separate parameters for reward seeking and loss avoiding conditions, we found that the environmental volatility parameter ϑ estimated in the reward seeking condition was a significant predictor of positive symptoms (d = 0.187, 95% CrI [-0.019, 0.403], P(d<0) = 0.039, Figure 6a, Supplementary Table 14), but we found no credible support for ω (d = -0.082, 95% CrI [-0.301, 0.137], P(d>0) = 0.235) nor ω_α (b = -0.168, 95% CrI [-0.392, 0.057], P(b>0) = 0.07). We also found no evidence for an association between the parameters estimated in the loss avoidance condition on positive symptoms (all P>0.326). When predicting negative symptoms, we found that in the loss avoiding condition, there was a negative association between ω_α (d = -0.298, 95% CrI [-0.523, -0.088], P(d>0) = 0.003, Figure 6b, Supplementary Table 15) and ω (d = -0.223, 95% CrI [-0.442, 0.003], P(d>0) = 0.026), but not ϑ (d = -0.012, 95% CrI [-0.238, 0.216], P(d>0) = 0.457). We found no evidence for an association of negative symptoms with the model parameters in the reward seeking condition (all P > 0.214).

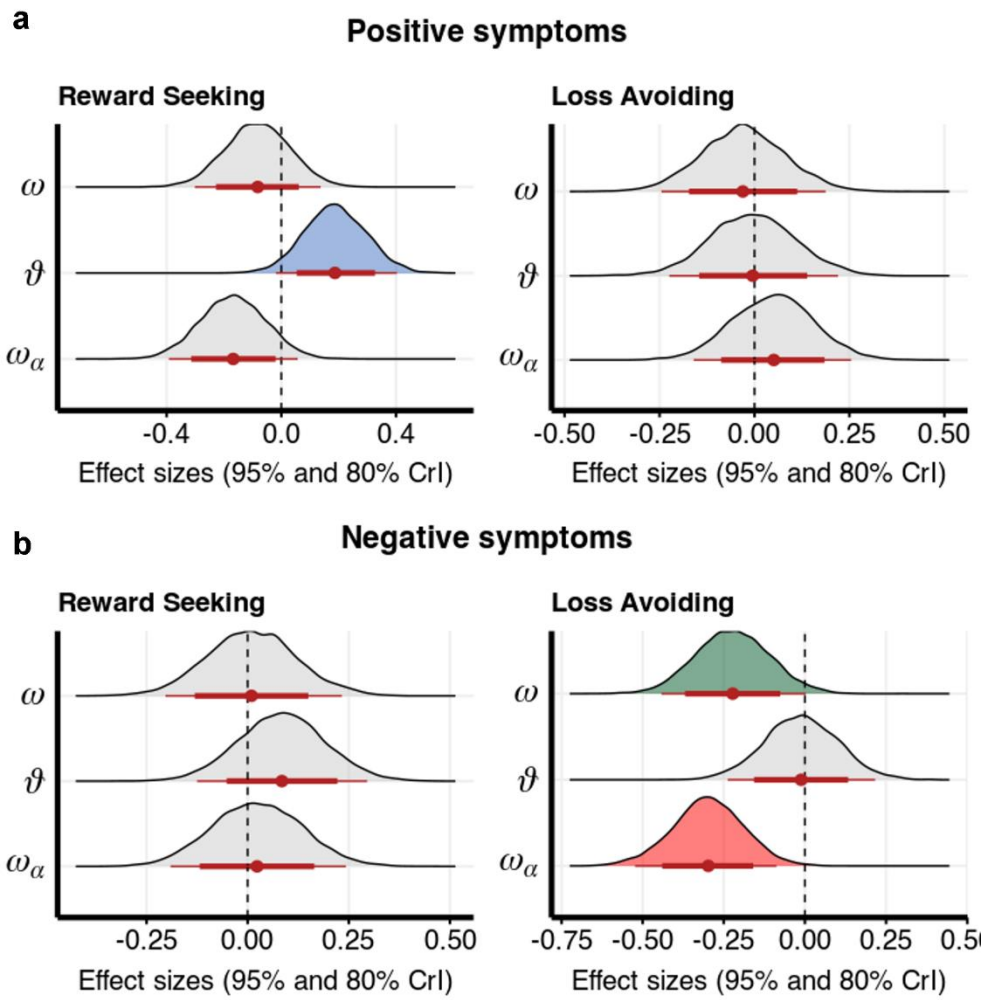


Figure 6. Modelling results from a sample of patients with chronic schizophrenia in Experiment 3. A total of eight parameters (three parameters of the belief model as well as a noise parameter η (on a log scale) for both reward seeking and loss avoiding condition) and age, gender, and IQ were used as predictor variables in two models predicting Positive symptoms measured by the Brief Psychiatric Rating Scale (**a**) and negative symptoms measured by the Scale for the Assessment of Negative Symptoms (**b**). Plotting posterior distributions of effects of each model parameter, over means and CrIs of effect sizes across two incentivization condition.

Discussion

In this paper, we present a computational model that describes how individuals with schizophrenia and schizotypal traits learn in environments where both observational noise and the volatility of environmental events can change. We demonstrate that schizotypal traits and positive symptoms in schizophrenia were related to higher uncertainty about the task volatility, while depressive-anxious traits and negative symptoms were associated with rigid beliefs about the underlying mean and outcome variance.

Previous work does suggest that schizotypal traits correlate with volatility estimates across both clinical and nonclinical samples (Deserno et al., 2019; Reed et al., 2020), however, not accounting for participants' perception of outcome reliability limits the conclusions that can be drawn from these studies (Piray & Daw, 2021; Pulcu & Browning, 2019). In tasks with binary outcome variables, outcome variance is determined from the probability of outcome and therefore indistinguishable from prior beliefs. Importantly, even in stable task environments (without latent changes), patients with schizophrenia do not down-weight their updating in response to higher outcome variance (Haarsma et al., 2020) and there are several known examples of patients failing to attenuate sensory signals (Adams et al., 2013; Schmack, Schnack, Priller, & Sterzer, 2015; Sterzer et al., 2018; Weilnhammer et al., 2020).

According to our results, in tasks with latent changes, positive symptoms in chronic patients and schizotypal trats are related to an inability or unwillingness to extract and utilize higher-order statistical features of the environment – processes that are encoded at higher levels of neural computation such as the hippocampus and the prefrontal cortex (Heinz et al., 2019; Lisman & Grace, 2005; Sterzer et al., 2018).

In support of this disrupted "model" learning or utilization, we found some evidence that the uncertainty about the volatility parameter is related to reduced performance in the Two-step task, a task commonly used to assess model-based behavior and prospective decision-making. These results link behavior in reversal learning tasks of individuals with schizotypal traits or positive symptoms to previous work showing impaired model-based control (Culbreth et al., 2016), or reduced goal-directed behavior and cognitive control in patients more generally (J. D. Cohen, 1996; Collins, Albrecht, Waltz, Gold, & Frank, 2017; Gold, Waltz, & Frank, 2015).

It is worth noting however, that higher uncertainty about the task volatility can have adaptive value since it allows for a rapid reduction of uncertainty and higher confidence when faced with unpredictability (Johnson & Fowler, 2011). In our data, higher confidence despite poor predictions was related to higher volatility uncertainty, as well as to schizotypal traits, in line with findings that patients with schizophrenia display overconfidence across a range of social, cognitive and learning tasks (Hahn, Moritz, Elmers, & Scheunemann, 2021; Hoven et al., 2019; Köther et al., 2012; Moritz et al., 2014; Rossi-Goldthorpe, Leong, Leptourgos, & Corlett, 2021). Modelling work has shown how having a wide prior over alternative hypotheses can serve as a "protective belt" when faced with surprising outcomes (Gershman, 2018b), which allows for explaining away unpredictable events with a new hypothesis (e.g., a new change has occurred), while being in line with the former belief being held with high confidence (Erdmann & Mathys, 2022). Within this framework, patients can display reduced belief updating (Waltz & Gold, 2007) and a bias against confirmatory evidence (Adams, Napier, Roiser, Mathys, & Gilleen, 2018; Doll et al., 2014), despite displaying sudden shifts in beliefs (Nassar et al., 2021), thereby reconciling the rigidness and confidence with which delusional beliefs are held, and the little evidence necessary to form those beliefs.

Our finding that rigid beliefs about outcome and outcome variance are related to the anxious-depression transdiagnostic psychiatric dimension is an important addition to previous modelling and empirical work (Pulcu & Browning, 2019). In a binary reversal learning task where shocks are paired with stimuli, one study showed that participants with higher trait anxiety were shown to less strongly adjust their learning rate (and pupillary response) to task volatility changes (Browning, Behrens, Jocham, O'Reilly, & Bishop, 2015) and another found a positive correlation between chronic stress and uncertainty about task volatility (De Berker et al., 2016). Recent modelling work with a joint model for outcome variance (stochasticity) and volatility suggests that the relationship between anxiety and learning rate adaptations is best explained by the degree to which surprising events are attributed to volatility vs variance belief updates (Piray & Daw, 2021). Our results provide empirical support for this claim, whereby individuals high on the anxious-depression dimension had more rigid beliefs about outcome variance, suggesting they were more likely to attribute surprising events to environmental changes. Similarly, we found the same pattern of reduced learning about variance and over all reduced learning being related to negative symptoms of patients in schizophrenia, but only when learning to avoid losses with no credible evidence for an effect when learning about gains. This is in accordance with the idea that negative symptoms arise from a reduced neural response to surprising events (Deserno, Heinz, et al., 2015; Maia & Frank, 2017), but seems to be at odds with a range of studies showing that this pattern is specific for reward and not avoidance learning (Gold et al., 2013, 2015; Strauss et al., 2011). This discrepancy could be due to most studies using binary choice outcomes, where loss avoidance is conceptualized as absence of action. This contrasts with our study, where behavior of participants reflects their inference about the underlying cause.

One important contribution of this study is the conceptual framework within which a task and model pair can be optimized for good test-retest reliability scores. We showed with simulations that

intra class coefficients of parameters can sometimes be improved with higher trial numbers, and that, in some cases, it not feasible to expect high parameter reliability, even with high trial numbers, such as estimating environmental volatility parameters using binary outcomes and behavioral inputs. This has three relevant implications. First, for reliable intra session measurement, tasks where participants learn about continuous input are preferred over binary inputs; second, simulations can be used to optimize the task generating process (for example, by restricting trajectories); and third, model selection can be contingent on the test-retest reliability of the computational parameters given the data. In line with previous work (Zech et al., 2022), we showed that empirical ICCs are improved with hierarchical estimation, which suggests that more sessions per individual will give more precise ICC estimates, as would standardization of experimental designs and models and extending to response models to include several data sources (e.g. confidence, or physiological data).

Despite the consistency of our main findings across the three studies, the effect sizes are low in the two higher powered studies (Experiment 2 and 3). Meta-analytic evidence indicates that this is usually the case when considering specific associations of belief updating patterns with positive symptoms (Dudley, Taylor, Wickham, & Hutton, 2016; McLean, Mattiske, & Balzan, 2017). These measures can nevertheless have a clinically meaningful impact (Funder & Ozer, 2019; Meyer et al., 2001), especially considering the possibility of repeated measurement that current standardized questionnaires are not suited for. However, to explain the heterogeneity of experiences across the psychosis continuum several other processes beyond belief updating (at least as they are operationalized within such experimental tasks) will have to be considered, such as the subjective nature, the content and the context of such experiences (Feyaerts, Henriksen, Vanheule, Myin-Germeys, & Sass, 2021).

In conclusion, we show that mechanistic descriptions of how individuals with psychotic symptoms and psychotic-like traits process information when faced with various sources of uncertainty have the potential to become reliable, transdiagnostic and clinically meaningful computational phenotypes. In future work, the robustness and reliability of these measures can be improved through larger standardization studies and extending response models to other data modalities. Clinical usability, such as predictive validity, can be explored by embedding the computational phenotypes within interventional clinical designs.

Supplementary Material

Supplementary Tables

Supplementary Table 1. Hierarchical Gaussian Filtering of the generative model returns inferred beliefs with update equations for each agent, given the free parameters ω, ω_α, and ϑ.

Beliefs	Inferred beliefs	Update equations
Mean	$x^{(t)} \sim N\left(\mu^{(t)}, \sigma^{(t)}\right)$	$\Delta\sigma = \exp\left(\mu_v^{(t)} + \boldsymbol{\omega}\right)$ $\pi^{(t)} = 1/\sigma^{(t)}$ $\hat{\pi}^{(t)} = \pi^{(t)} + \pi_u^{(t)}$ $\pi_u^{(t)} = 1/\exp\left(\mu_\alpha^{(t)}\right)$ $\psi_1 = \frac{\pi_u^{(t)}}{\hat{\pi}^{(t)}}$ $\Delta\mu = \psi_1\delta$
Volatility	$x_v^{(t)} \sim N\left(\mu_v^{(t)}, \sigma_v^{(t)}\right)$	$\Delta\sigma_v = \exp(\boldsymbol{\vartheta})$ $\pi_v = 1/\sigma_v$ $\Delta\mu_v \propto \frac{1}{\pi_v}\ \delta_v$
Outcome variance	$x_\alpha^{(t)} \sim N\left(\mu_\alpha^{(t)}, \sigma_\alpha^{(t)}\right)$	$\Delta\sigma_\alpha = \exp\left(\boldsymbol{\omega_\alpha}\right)$ $\pi_\alpha = 1/\sigma_\alpha$ $\Delta\mu_\alpha \propto \frac{1}{\pi_\alpha}\ \delta_\alpha$

Supplementary Table 2. Priors for the three HGF models used to analyze data from the Test-retest study (N=45).

	HGF 1	HGF 2	HGF 3
	Model with all uncertainty parameters, but fixed initial variance and volatility belief	Model with a fixed environmental volatility parameter, but free initial volatility estimate	Model with a fixed environmental volatility parameter, and fixed initial variance and volatility belief
Priors	$\omega \sim N(-3,5)$ $\omega_\alpha \sim N(-3,5)$ $\vartheta \sim N(-3,5)$ $\mu_v^{(0)} = 4$ $\mu_\alpha^{(0)} = 5$	$\omega \sim N(-3,5)$ $\omega_\alpha \sim N(-3,5)$ $\exp(\vartheta) = 0$ $\mu_v^{(0)} \sim N(4,5)$ $\mu_\alpha^{(0)} = 5$	$\omega \sim N(-3,5)$ $\omega_\alpha \sim N(-3,5)$ $\exp(\vartheta) = 0$ $\mu_v^{(0)} = 4$ $\mu_\alpha^{(0)} = 5$

Supplementary Table 3. Bayesian model comparison in Experiment 1. Exceedance probabilities across the four sessions for the three models described in Supplementary Table 2.

	HGF 1	**HGF 2**	**HGF 3**
Session 1	0.9863	0.0131	0.0006
Session 2	0.8608	0.0270	0.1123
Session 3	0.8166	0.1782	0.0053
Session 4	0.5117	0.4686	0.0197

Supplementary Table 4. The effect of model parameters in Experiment 1 on the participant-level random slope from the model predicting singed updates from signed prediction errors. Values of N = 45 participants and all four sessions were included. All variables were scaled.

$B_U \sim \omega + \vartheta + \omega_\alpha$	Estimate	Std. Error	t value	Pr(>|t|)
$(Intercept)$	0	0.038	0	0.999
ω	0.795	0.053	14.884	<0.001
ϑ	0.54	0.059	9.108	<0.001
ω_α	-0.008	0.065	-0.121	0.904

Supplementary Table 5. The effect of model parameters in Experiment 1 on the participant-level random slope from the model predicting behavioral learning rates from variance block (factor). Values of N = 45 participants and all four sessions were included. All predictor variables were scaled.

$B_L \sim \omega + \vartheta + \omega_\alpha$	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.068	0.083	-0.814	0.417
ω	0.146	0.107	1.371	0.172
ϑ	-0.22	0.105	-2.097	0.037
ω_α	0.438	0.115	3.822	<0.001

Supplementary Table 6. The effect of model parameters in Experiment 1 on the participant-level random slope from the model predicting confidence from absolute prediction errors. Values of N = 45 participants and all four sessions were included. All variables were scaled.

$B_C \sim \omega + \vartheta + \omega_\alpha$	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0	0.071	0	0.999
ω	0.216	0.1	2.155	0.032
ϑ	0.356	0.111	3.202	0.002
ω_α	-0.058	0.121	-0.483	0.63

Supplementary Table 7. Effects of PDI on model parameters across sessions in Experiment 1. To check the robustness or our results across sessions, we ran three models each predicting one model parameter with PDI and a factor for session as a fixed and random effect. We repeated the same process for parameters estimated without pooling as well as parameters estimated by one hierarchical model.

$\vartheta \sim PDI * session + (session|ID)$

Par. estimation	Hierarchical				No pooling			
	Est.	2.5%	97.5%	P	Est.	2.5%	97.5%	P
$B_{PDI\_session1}$	0.486	0.01	0.964	0.024	0.39	-0.062	0.851	0.05
$B_{PDI\_session2}$	0.393	-0.141	0.932	0.066	0.401	-0.127	0.94	0.068
$B_{PDI\_session3}$	0.358	-0.157	0.901	0.086	0.181	-0.294	0.64	0.234
$B_{PDI\_session4}$	0.864	0.227	1.508	0.005	0.792	0.283	1.27	0.002

$\omega \sim PDI * session + (session|ID)$

Par. estimation	Hierarchical				No pooling			
	Est.	2.5%	97.5%	P	Est.	2.5%	97.5%	P
$B_{PDI\_session1}$	-0.202	-0.44	0.038	0.951	-0.508	-0.891	-0.126	0.996
$B_{PDI\_session2}$	-0.215	-0.456	0.031	0.96	-0.389	-0.798	0.006	0.974
$B_{PDI\_session3}$	-0.117	-0.339	0.113	0.853	-0.105	-0.359	0.174	0.787
$B_{PDI\_session4}$	-0.128	-0.403	0.149	0.819	-0.101	-0.491	0.278	0.7

$\omega_\alpha \sim PDI * session + (session|ID)$

Par. estimation	Hierarchical				No pooling			
	Est.	2.5%	97.5%	P	Est.	2.5%	97.5%	P
$B_{PDI\_session1}$	0.231	-0.167	0.669	0.131	0.369	-0.085	0.822	0.054
$B_{PDI\_session2}$	-0.026	-0.578	0.552	0.535	-0.27	-0.931	0.403	0.786
$B_{PDI\_session3}$	0.007	-0.527	0.579	0.489	-0.016	-0.502	0.482	0.526
$B_{PDI\_session4}$	-0.006	-0.61	0.625	0.508	0.039	-0.548	0.642	0.439

Supplementary Table 8. In Experiment 2, effects of model parameters in the "Unusual Experiences" subscale of the Short Scale for Measuring Schizotypy. All continuous variables were scaled. Effect sizes reported in the main text obtained by dividing the effect of interest by residual standard deviation σ.

$SCZ_{UE} \sim \omega + \vartheta + \omega_\alpha + Age + IQ + Gender$				
	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.031	0.072	-0.112	0.169
ϑ	0.09	0.047	-0.001	0.181
ω_α	-0.012	0.048	-0.106	0.083
ω	-0.045	0.048	-0.137	0.049
Age	-0.211	0.047	-0.307	-0.119
IQ	-0.065	0.047	-0.159	0.029
Gender	-0.057	0.097	-0.243	0.131

Supplementary Table 9. In Experiment 2, effects of model parameters in the "Introvertive Anhedonia" subscale of the Short Scale for Measuring Schizotypy. All continuous variables were scaled. Effect sizes reported in the main text obtained by dividing the effect of interest by residual standard deviation σ.

$SCZ_{IA} \sim \omega + \vartheta + \omega_\alpha + Age + IQ + Gender$				
	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.089	0.071	-0.046	0.226
ϑ	0.016	0.047	-0.077	0.112
ω_α	-0.08	0.049	-0.176	0.017
ω	-0.071	0.049	-0.166	0.023
Age	-0.176	0.048	-0.271	-0.082
IQ	0.01	0.046	-0.078	0.1
Gender	-0.16	0.095	-0.346	0.023

Supplementary Table 10. In Experiment 2, effects of two subscales of schizotypal traits: "Unusual Experiences" (UE) and "Introvertive Anhedonia" on confidence rating s on the next trial predicted by the confidence rating on the previous trial and prediction error on the previous trials and its interaction with the group level variables. Prediction error variable was scaled both within group and participant (id), group level variables were scaled within group. Confidence ratings were shrunk by 5% and logit transformed. Total variance was calculated by the sum of all variance component. Effect sizes reported in the main text are obtained by dividing the effect of interest by the squared root of the total variance.

$Confidence^{(t+1)}_{LogOdds} \sim Confidence^{(t)}_{LogOdds} + PE^{(t)}(SZC_{UE} + SCZ_{IA}) + (PE^{(t)} \mid id)$				
Total Variance = 1.213	Estimate	Est.Error	Q2.5	Q97.5
$Intercept$	0.187	0.028	0.128	0.241
$Confidence^{(t)}_{LogOdds}$	0.471	0.002	0.467	0.476
$PE^{(t)}$	-0.381	0.012	-0.405	-0.358
SZC_{UE}	0.183	0.03	0.124	0.241
SCZ_{IA}	-0.06	0.03	-0.117	-0.001
$PE^{(t)}SZC_{UE}$	0.042	0.012	0.018	0.066
$PE^{(t)}SZC_{IA}$	0.012	0.012	-0.011	0.036

Supplementary Table 11. In Experiment 2, effects of model parameters in the "compulsive behavior and intrusive thought" factor. All continuous variables were scaled. Effect sizes reported in the main text obtained by dividing the effect of interest by residual standard deviation σ.

$CIT \sim \omega + \vartheta + \omega_\alpha + Age + IQ + Gender$				
	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-0.004	0.068	-0.137	0.131
ϑ	0.133	0.043	0.048	0.217
ω_α	0.019	0.045	-0.069	0.108
ω	-0.06	0.044	-0.149	0.027
Age	-0.318	0.045	-0.407	-0.231
IQ	-0.207	0.045	-0.294	-0.118
Gender	0.007	0.093	-0.179	0.194

Supplementary Table 12. In Experiment 2, effects of model parameters in the "anxious-depression" factor. All continuous variables were scaled. Effect sizes reported in the main text obtained by dividing the effect of interest by residual standard deviation σ.

$AD \sim \omega + \vartheta + \omega_\alpha + Age + IQ + Gender$				
	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.155	0.07	0.018	0.293
ϑ	-0.005	0.046	-0.095	0.086
ω_α	-0.106	0.048	-0.2	-0.013
ω	-0.122	0.048	-0.216	-0.029
Age	-0.221	0.047	-0.313	-0.128
IQ	-0.01	0.047	-0.1	0.082
Gender	-0.281	0.094	-0.465	-0.1

Supplementary Table 13. In Experiment 2, effects of three factors of the factor analysis on confidence rating s on the next trial predicted by the confidence rating on the previous trial and prediction error on the previous trials and its interaction with the group level variables. Prediction error variable was scaled both within group and participant (id), group level variables were scaled within group. Confidence ratings were shrunk by 5% and logit transformed. Total variance was calculated by the sum of all variance component. Effect sizes reported in the main text are obtained by dividing the effect of interest by the squared root of the total variance.

$Confidence_{LogOdds}^{(t+1)} \sim Confidence_{LogOdds}^{(t)} + PE^{(t)}(CIT + AD + SW) + (PE^{(t)} \mid id)$				
Total Variance = 1.201	Estimate	Est.Error	Q2.5	Q97.5
$Intercept$	0.186	0.026	0.133	0.238
$Confidence_{LogOdds}^{(t)}$	0.471	0.002	0.467	0.476
$PE^{(t)}$	-0.381	0.011	-0.404	-0.359
CIT	0.264	0.031	0.202	0.327
AD	-0.118	0.034	-0.185	-0.054
SW	-0.035	0.036	-0.107	0.035
$PE^{(t)}CIT$	0.07	0.013	0.044	0.096
$PE^{(t)}AD$	0.009	0.014	-0.018	0.038
$PE^{(t)}SW$	-0.033	0.015	-0.063	-0.004

Supplementary Table 14. In Experiment 3, effects of model parameters for both avoid and seek sessions on positive symptoms measured by the Brief Psychiatric Rating Scale. All continuous variables were scaled. Effect sizes reported in the main text obtained by dividing the effect of interest by σ.

$BPRS \sim \eta_{avoid} + \omega_{avoid} + \vartheta_{avoid} + \omega_{\alpha,avoid} + \eta_{seek} + \omega_{seek} + \vartheta_{seek} + \omega_{\alpha,avoid} + Age + IQ + Gender$

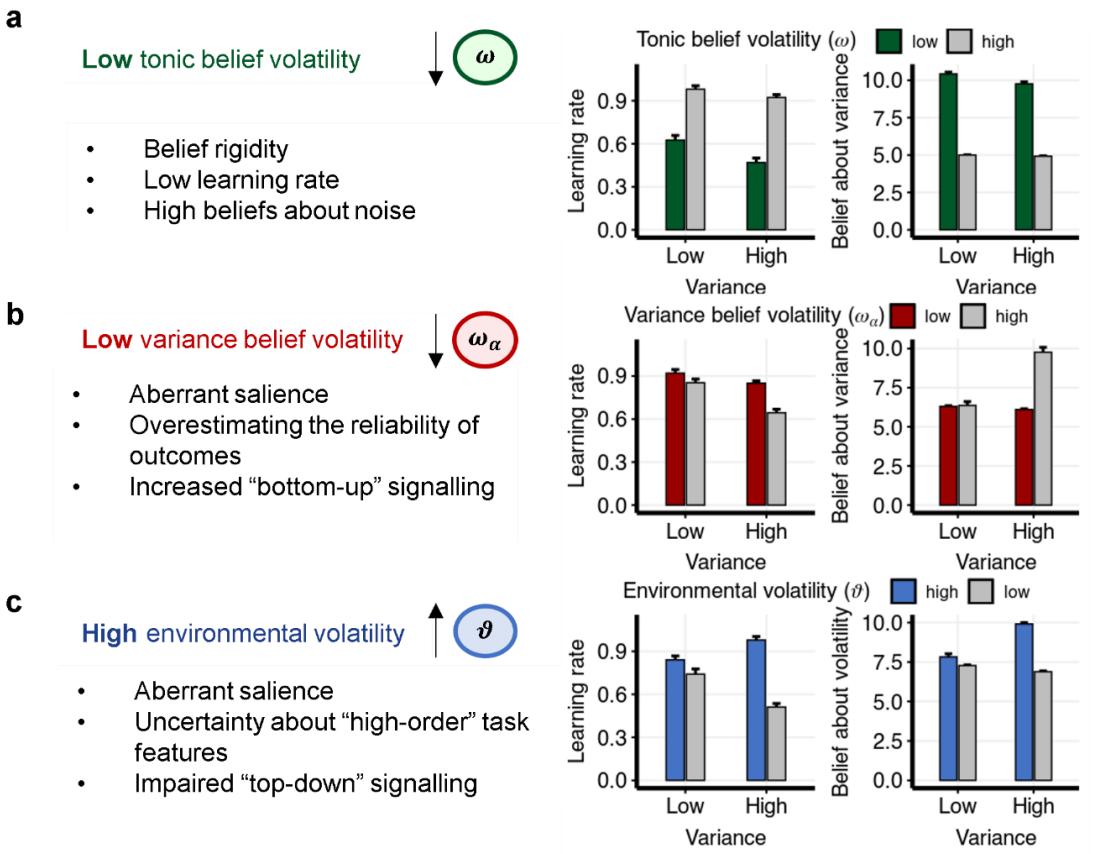
$\sigma = 1.20$	Estimate	Est.Error	Q2.5	Q97.5
Intercept	2.114	0.124	1.87	2.354
Age	0.094	0.125	-0.153	0.344
IQ	0.086	0.126	-0.166	0.334
Gender	-0.048	0.257	-0.562	0.466
η_{avoid} (log noise)	0.022	0.126	-0.224	0.264
ω_{avoid}	-0.035	0.133	-0.29	0.232
ϑ_{avoid}	-0.006	0.135	-0.268	0.259
$\omega_{\alpha,avoid}$	0.058	0.128	-0.193	0.31
η_{seek}(log noise)	0.145	0.131	-0.106	0.406
ω_{seek}	-0.098	0.134	-0.363	0.165
$\boldsymbol{\vartheta_{seek}}$	0.224	0.127	-0.023	0.476
$\omega_{\alpha,seek}$	-0.2	0.137	-0.467	0.068

Supplementary Table 15. In Experiment 3, effects of model parameters for both avoid and seek sessions on negative symptoms measured by the Scale for the Assessment of Negative Symptoms. All continuous variables were scaled. Effect sizes reported in the main text obtained by dividing the effect of interest by σ.

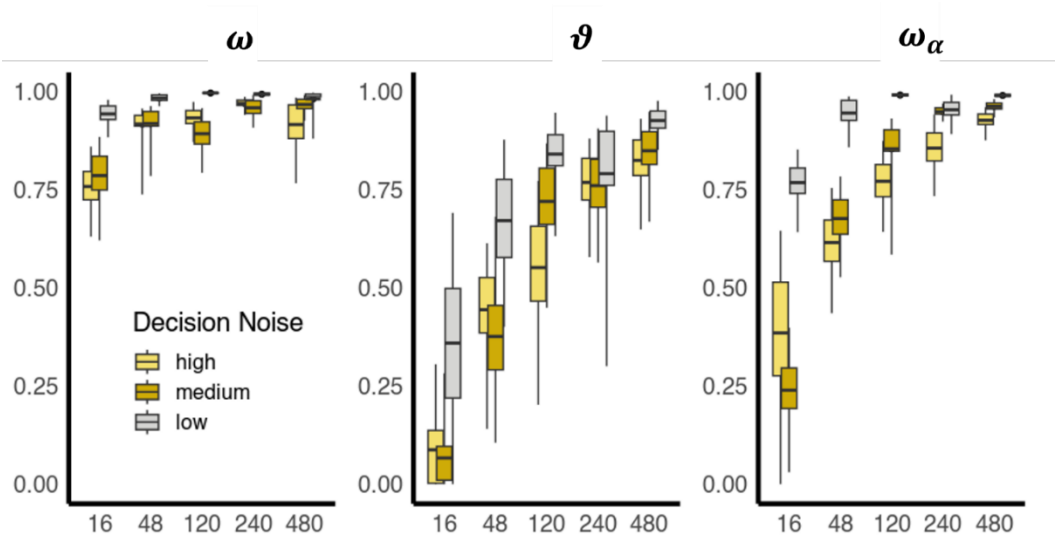
$BPRS \sim \eta_{avoid} + \omega_{avoid} + \vartheta_{avoid} + \omega_{\alpha,avoid} + \eta_{seek} + \omega_{seek} + \vartheta_{seek} + \omega_{\alpha,avoid} + Age + IQ + Gender$

$\sigma = 0.62$	Estimate	Est.Error	Q2.5	Q97.5
Intercept	1.425	0.065	1.297	1.553
Age	0.225	0.066	0.097	0.353
IQ	0.063	0.063	-0.062	0.188
Gender	-0.136	0.139	-0.408	0.138
η_{avoid} (log noise)	-0.009	0.065	-0.137	0.118
$\boldsymbol{\omega_{avoid}}$	-0.137	0.07	-0.274	0.002
ϑ_{avoid}	-0.007	0.071	-0.148	0.133
$\boldsymbol{\omega_{\alpha,avoid}}$	-0.184	0.066	-0.316	-0.055
$\boldsymbol{\eta_{seek}}$(log noise)	-0.14	0.068	-0.273	-0.002
ω_{seek}	0.006	0.068	-0.127	0.143
ϑ_{seek}	0.052	0.066	-0.079	0.184
$\omega_{\alpha,seek}$	0.015	0.069	-0.118	0.149

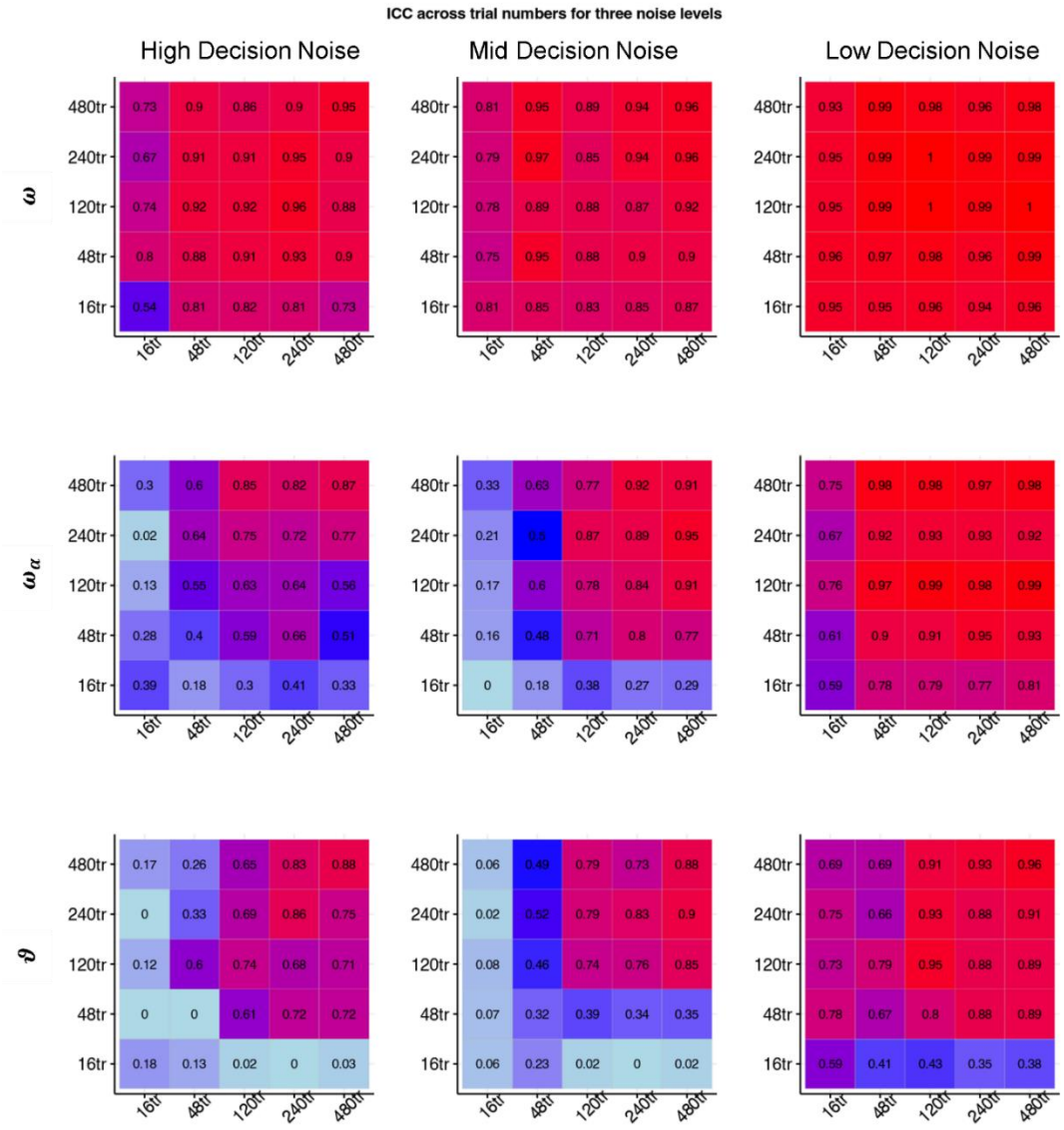
Supplementary Figures



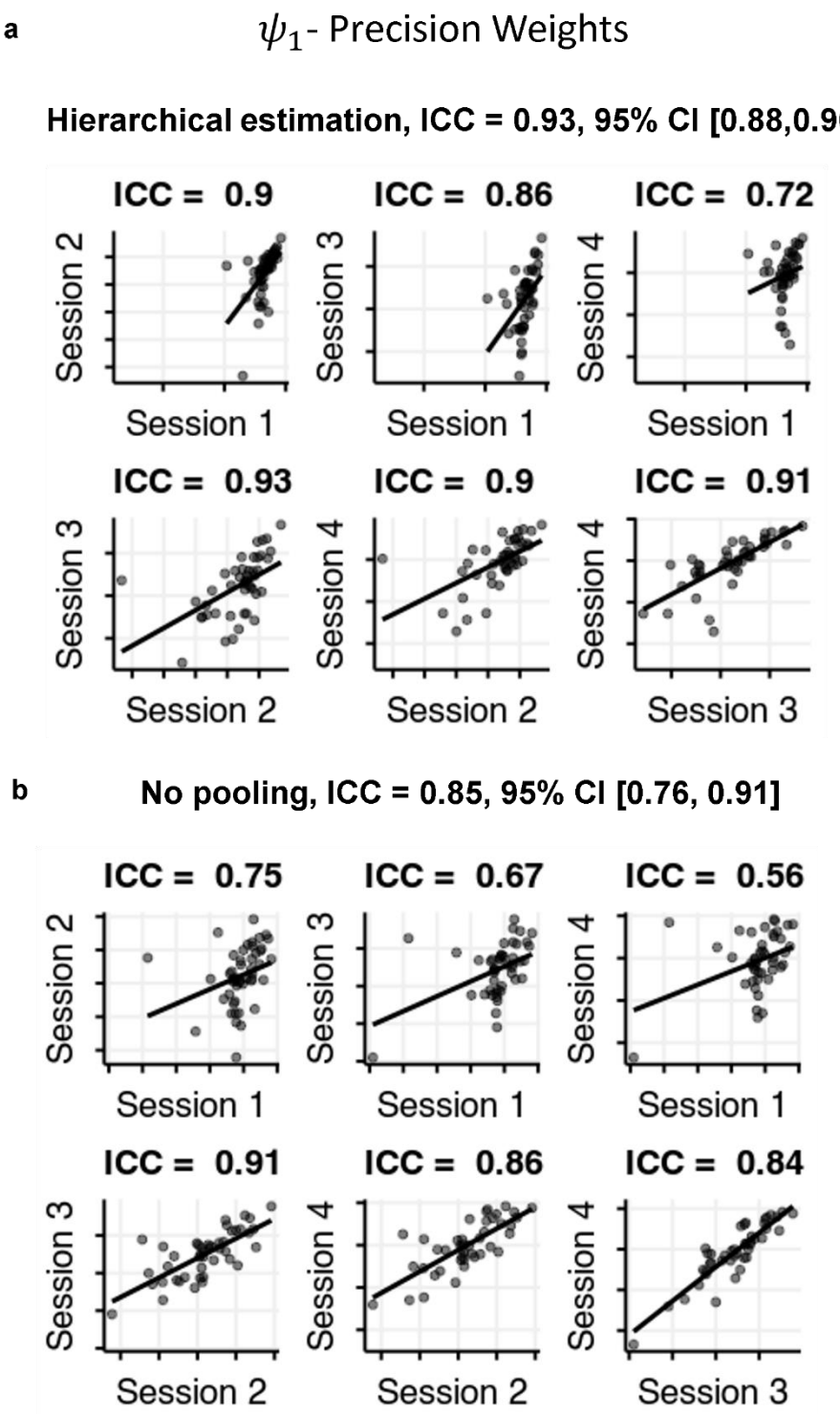
Supplementary Figure 1. Model simulations. Model simulations of two agents with different parameter sets. **a**, Different ω parameters and fixed and low values for ω_α and ϑ. **b**, Different ω_α and fixed ω and ϑ. **c**, Different ϑ and fixed ω and ω_α. Learning rate refers to the ratio of action update and signed prediction error. Bar plots are means across blocks for each agent with error bars representing standard errors. Simulations are done with 480 trials for two agents (one agent per parameter set). Through this, we can define several computational patterns of belief updating. When comparing simulated behavior and belief trajectories of an agent with lower (compared to higher) tonic belief volatility the model implies lower prior precision and therefore lower precision-weighted learning rate (ψ_1). Therefore, the behavioral learning rate is reduced across blocks with both low and high outcome variance. Furthermore, large deviations from expected outcomes are explained as noise, leading to high beliefs about outcome variance. In contrast, in the case of a lower outcome variance volatility simulations show an increased learning rate, particularly in blocks with high variance, resulting from a failure to update beliefs about the noise in high variance blocks. When environmental volatility is higher , higher outcome variance is interpreted as increased volatility, leading again to higher learning rates in the high variance block. Both latter computational patterns can be perceived as related to aberrant salience attribution; however, they might be implemented in distinct neural circuits.



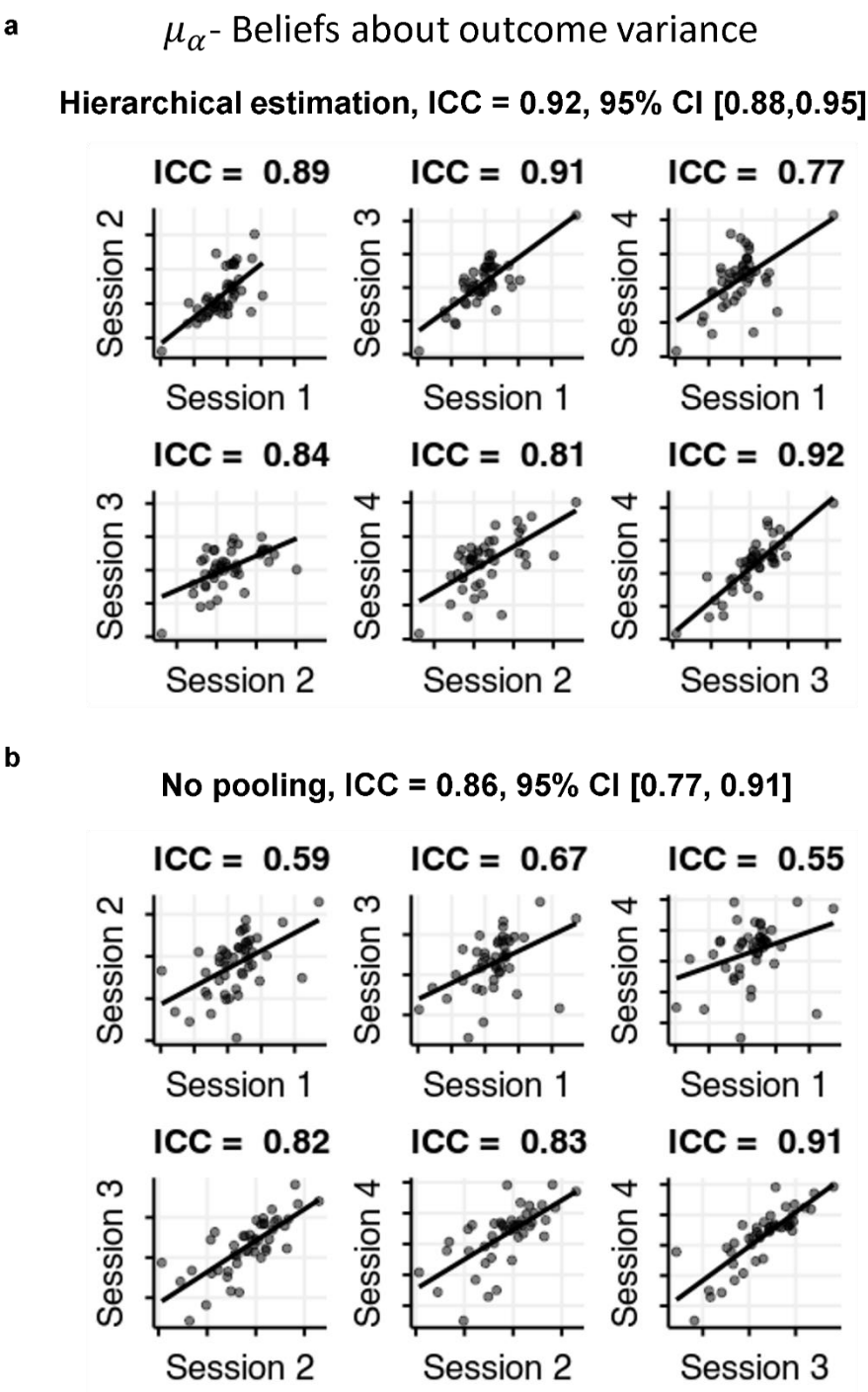
Supplementary Figure 2. Parameter recovery. We simulated behavior for n = 50, with parameters drawn from the prior distribution and re-estimated the parameters for various trial lengths and three decision noise levels.



Supplementary Figure 3. Simulating ICCs across trial numbers for three different levels of decision noise. tr – Trials

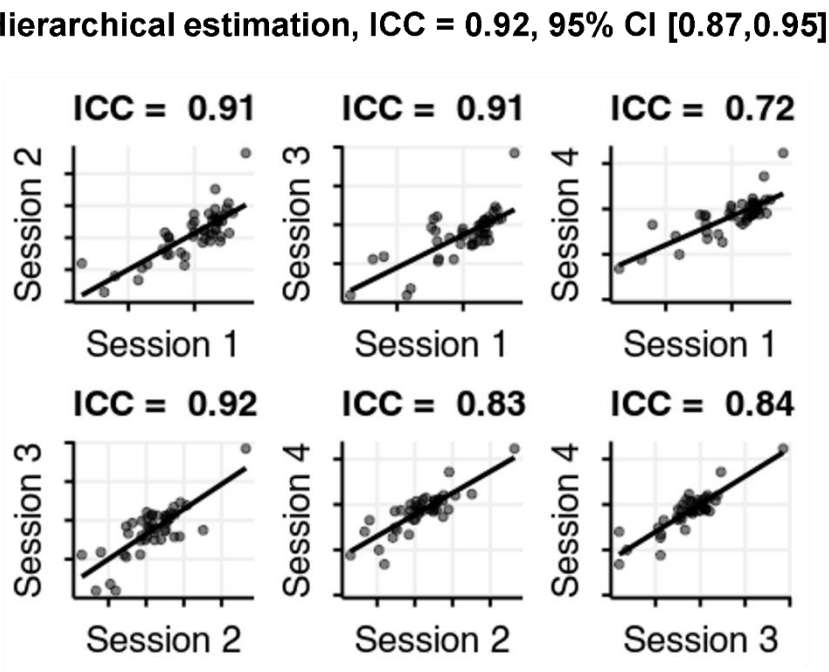


Supplementary Figure 4. ICCs and regression lines for average precision weights for parameter estimation with (a) and without (b) pooling.

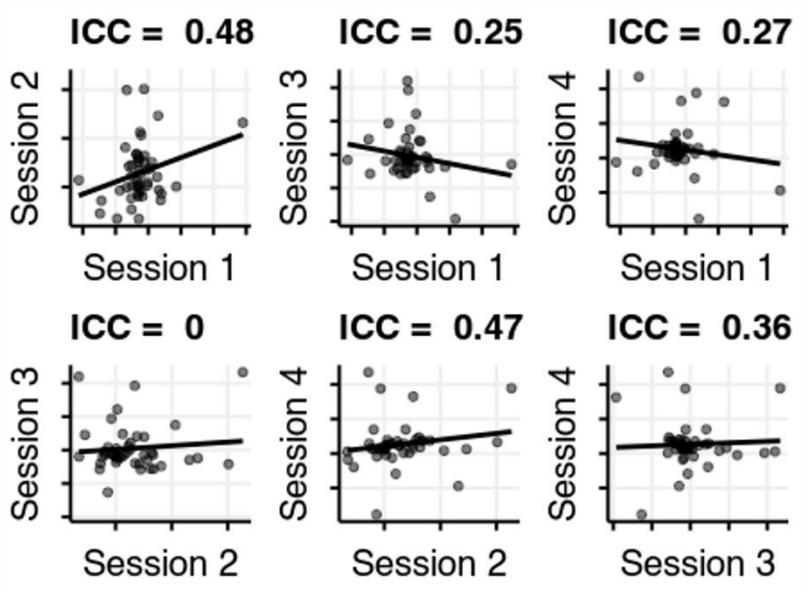


Supplementary Figure 5. ICCs and regression lines for average beliefs about variance for parameter estimation with (a) and without (b) pooling.

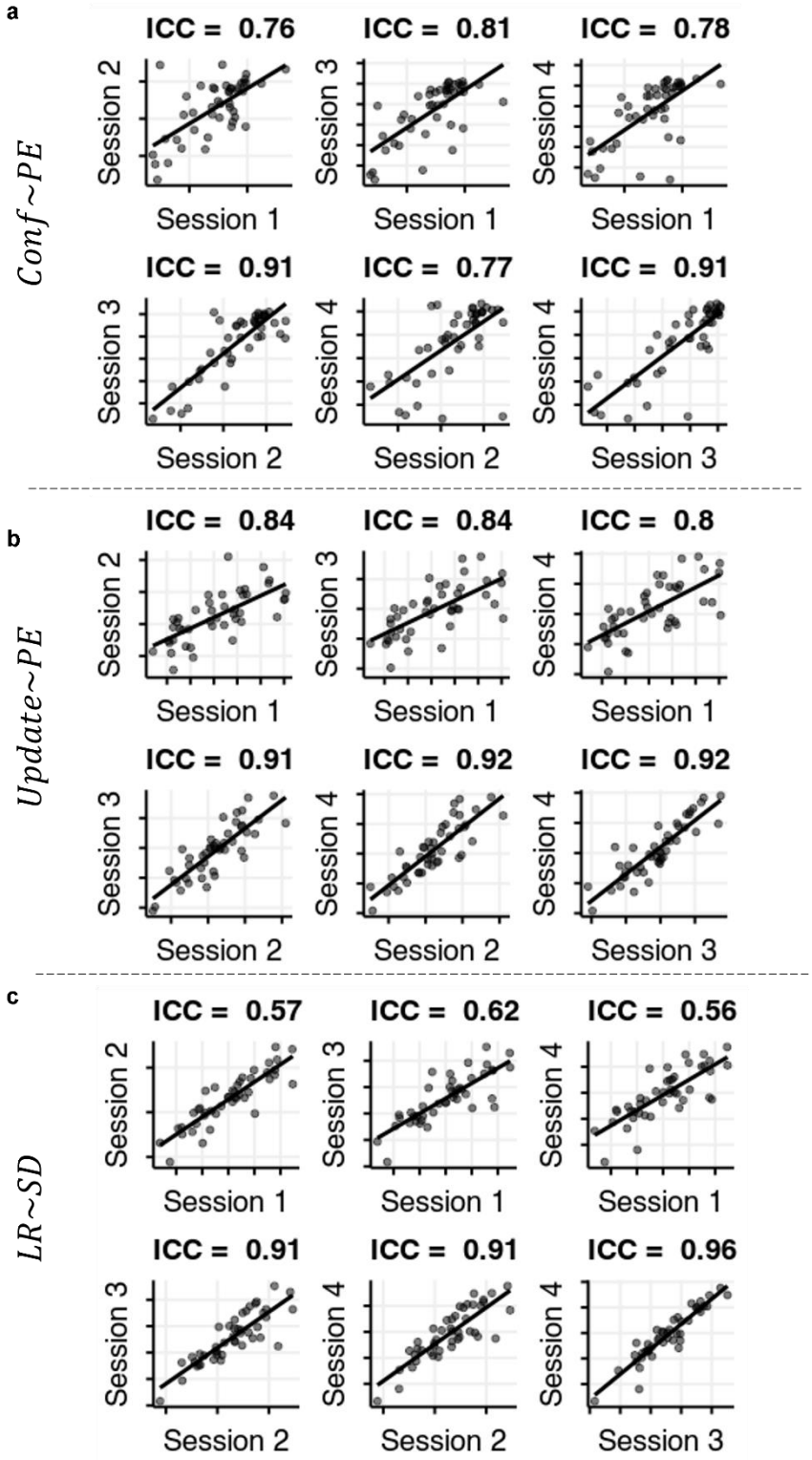
a μ_v - Beliefs about environmental volatility



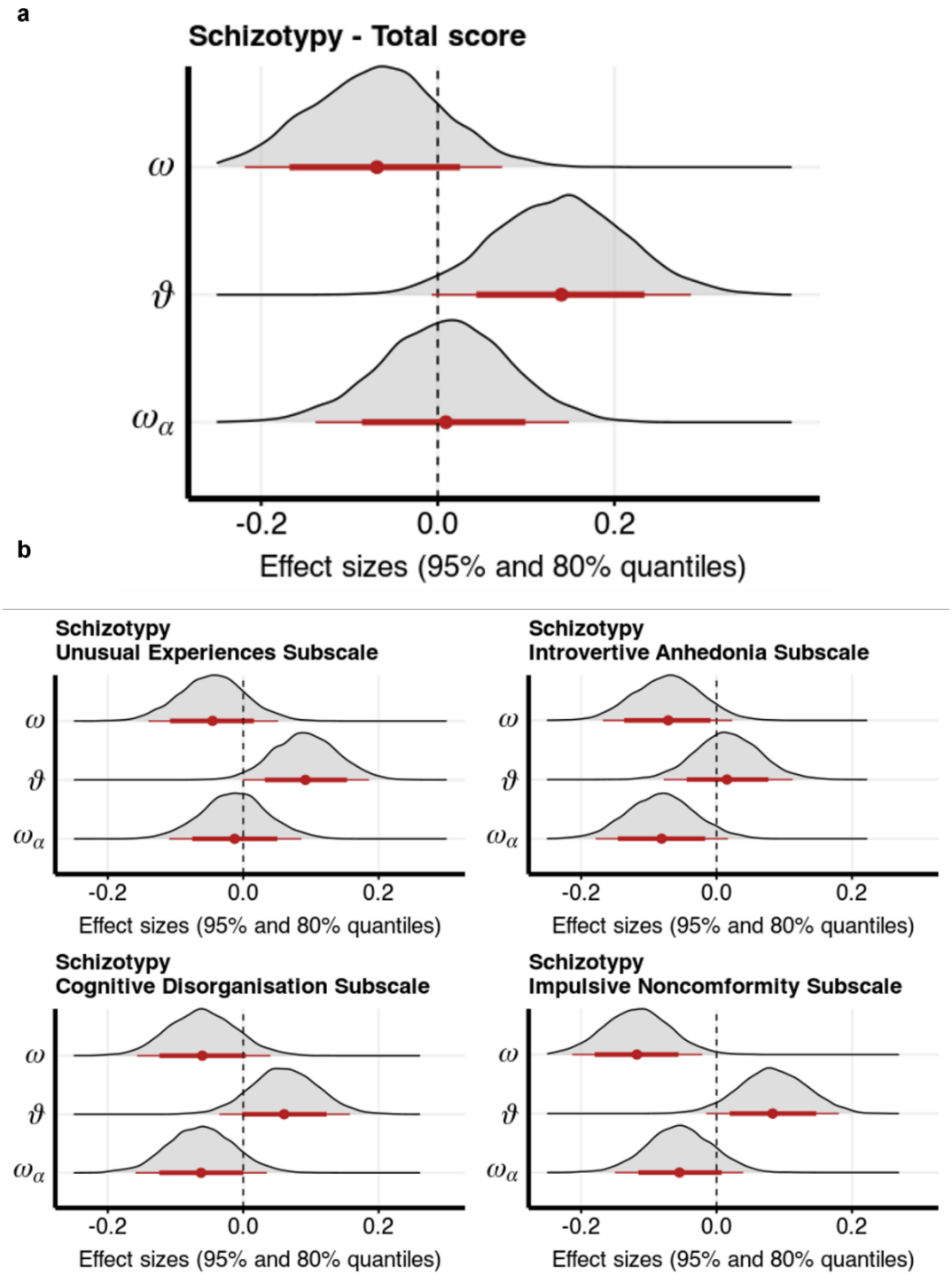
b



Supplementary Figure 6. ICCs and regression lines for average beliefs about volatility for parameter estimation with (a) and without (b) pooling across sessions.



Supplementary Figure 7. ICCs and regression lines for behavioral patterns (a) B_c, (b) B_U, and (c) B_L defined by the random slopes of three models, estimated per session.



Supplementary Figure 8. Modelling results Seow and Gillan (2020). a Total score of the Short Scale for measuring Schizotypy predicted by the three parameters controlling for age, gender and IQ. **b**, All subscales of the Short Scale for measuring Schizotypy were predicted by the three parameters controlling for age, gender and IQ. Plotting posterior distributions of effects of each model parameter, over means and CrIs of effect sizes.

Supplementary Methods

Behavioural Analysis

Variables that were only positive were translated by 1 (to avoid dividing by zero), and log-transformed (e.g. learning rate, absolute (unsigned) update, and PDI). Where possible, all within subject variables were used as random effects. In all linear models, all continuous variables were scaled. For Bayesian models we report Credibility intervals (CrI) and the probability of the posterior interval that lies above (or below) zero. Effect sizes for linear models with random effects were calculated by dividing the regression coefficient of interest with the square of the total variance, calculated as a sum of residual variance and the variance of all random effects (Nalborczyk, Batailler, Loevenbruck, Vilain, & Bürkner, 2019).

Parameter estimation

Initial parameter estimation and simulations were done for each individual and session separately using maximum a posteriori estimates and the HGF toolbox in Matlab (Mathworks) and is available at www.translationalneuromodeling.org/tapas as open-source code (GPLv3).

In the test-retest study (Experiment 1), we first analysed behavior without pooling. We compared three different models using the Random-Effects Bayesian Model Selections (Stephan et al., 2009). See Supplementary Table 2 for the priors of all three models.

Because hierarchical inference can improve estimation of parameters across individuals and at the group level (Waltmann et al., 2022), we also analyzed our data within a hierarchical Bayesian framework. Models were implemented in custom code in Stan (Carpenter et al., 2017) using R as the programming language and RStudio as the integrated development environment for R (see the full model definition below). Models were estimated using Markov Chain Monte Carlo sampling, with four independent chains and 3000 iterations (1000 warm-up). Convergence of sampling chains was estimated through the Gelman-Rubin $\hat{R}$ statistic (Gelman & Rubin, 1992), whereby we considered $\hat{R}$ values smaller than or equal to 1.01 as acceptable.

Data from Experiments 2 and 3 were analyzed without pooling with the same priors and procedures as in Experiment 1.

The intraclass correlation (ICC) for the parameters was defined as the ratio of between-participant variance and total variance and estimated either with custom code (for hierarchical models) or using the *icc* function from the R package psych (using bootstrapping with 1000 samples to obtain distributions).

Hierarchical Bayesian Model definition

Recent work has shown that hierarchical inference can improve reliability scores (Waltmann et al., 2022), because it reduces overfitting (McElreath, 2018), pools information across different levels (testing sessions and participants) and allows for estimation of both participant parameters as well as group-level relevant statistics (such as the ICC scores) in one inferential step.

The participant level parameters for session 1 (ω, ω_α, ϑ, and η), were modelled as correlated (multivariate) effects with

$$\begin{pmatrix} \omega_{\mathrm{id}} \\ \omega_{\alpha,\mathrm{id}} \\ \vartheta_{\mathrm{id}} \\ \eta_{\mathrm{id}} \end{pmatrix} \sim \mathrm{MVNormal}(\boldsymbol{\mu}, \boldsymbol{S}), \quad (1)$$

$$S = \boldsymbol{\sigma^b R \sigma^b},$$

where $\boldsymbol{S}$ is the covariance matrix and $\boldsymbol{\mu}$ is the vector of mean values for each parameter. The matrix $\boldsymbol{S}$ was factored into a diagonal matrix with between subject standard deviations of the parameters (σ^b) and the correlation matrix $\boldsymbol{R}$ (Bürkner, 2017; McElreath, 2018). For each parameter (e.g. ω) for participant id and session s we therefore defined

$$\omega_{id,s} = \mu_\omega + b_s + r_{id} + r_{id,s},$$

where μ_ω is the overall mean, b_s is the overall session (fixed) effect, r_{id} is the participant level intercept, and $r_{id,s}$ is the participant level effect of session. With the following prior structure:

$$r_{id,s} \sim N(0, \sigma_\omega^w),$$
$$b_s \sim N(0,1)$$
$$\sigma_\omega^w \sim HalfCauchy(0,2),$$
$$\sigma_\omega^b \sim HalfCauchy(0,2),$$
$$R \sim LKJcorr(2),$$
$$\mu_\omega, \mu_\eta \sim N(0,1),$$
$$\mu_{\omega_\alpha}, \mu_\vartheta \sim N(-6,1)$$

The intraclass correlation (*ICC*) for the parameters was defined as the ratio of between-subjects variance (σ_b^2) versus total variance:

$$ICC = \frac{\left(\sigma^b\right)^2}{(\sigma^b)^2 + (\sigma^w)^2}.$$

Predictive Inference task

Experiment 1

After a change point the new mean was drawn from a uniform distribution on the [10,90] interval and the standard deviation could be either high (15) or low (5). Participants are required to express their confidence in their choice by holding the response button for longer (for a max of 1.4 seconds). Participants were reminded several times throughout the task to pay attention to the confidence ratings, with a text message flashed on the screen during the prediction phase of the next trial.

The task was implemented in Psychtoolbox in Matlab (Mathworks). We embedded the task within a story and incentivized accuracy. Participants were told to predict the amount of gold (indicated by a yellow bar) an alien will bring back from a mine on each trial with their arrow keys. Whereas the alien would usually bring back similar amounts of gold from one mine, it might also change mines once and a while. Participants were trained in a sample of mines that differed both in the mean amount of gold they produced as well as the variance. The participants were told that the amount of money they earned on each trial depended on the size of the yellow bar and how close they were to the actual outcome. A full bar was indicative of € 1 of potential gain. Participants first underwent a 20-minute training session where they would experience low and high variance blocks as well as several mean shifts, where the changes of mines were made obvious through visual cues.

In each session participants played 240 trials of the task. Four different trajectories were used for the four sessions and the trajectories were the same for all participants. The trajectories were selected randomly but with a fixed number of reversals. Furthermore, we conditioned the selection of the trajectories on the model. To maximize test-retest reliability, we selected trajectories where the parameters of an ideal Bayesian observer (given the priors in Table 2) were not too far away from each other (i.e. the Mahalanobis distance was below the 5th percentile).

Experiment 2:

In contrast to the task used in Experiment 1, the variance of responses did not change, but the volatility did. In high volatility blocks of trials, the probability for a change on each trial was 0.125, and in low volatility blocks it was 0.025. In this version of the task participants were told to catch a particle flying from an invisible canon from the center to the edge of a large circle. Participants would gain points when they caught the particle and lose points otherwise. All participants played 300 trials of the task, with one volatility condition across 150 trials. In contrast to Experiment 1, each participant's trajectory was randomly generated. The standard deviation around the mean bucket position was 12 from a 360-point scale (which compares to a standard deviation of 3.33 on a scale from 1 to 100 used in Experiment 1).

Experiment 3

The version of the task was comparable to that in Experiment 2 and to other studies published previously (Nassar et al., 2010). Participants were told, they need to catch bags that are dropped from a helicopter, whereby they cannot see the helicopter and need to infer its position based on the dropped bags (McGuire, Nassar, Gold, & Kable, 2014). The probability of the helicopter changing its position on each trial was 0.125 (as in the high volatility block of Experiment 2 and comparable to the volatility in Experiment 1), and the standard deviation was 20 on a 300-point scale (which compares to a standard deviation of 6.67 on a scale from 1 to 100 used in Experiment 1). Contrary to both Experiment 1 and 2, participants' performance was incentivized in two different ways across two different task sessions, each lasting for 100 trials. In the appetitive condition, participants could earn points and in the aversive condition, participants would have points taken away from their initial endowment in a performance contingent way.

Two-step Task

The task was based on the version from Kool, Cushman, & Gershman (2016), and was slightly adapted (Mikus, Korb, et al., 2022). Participants are required to earn points in a dynamically changing environment whereby each trial consists of two steps or stages. In the first stage, participants see one of two possible starting scenarios, each featuring a pair of spaceships that fly deterministically to one of two second stages that feature planets, where they encounter an alien that gives them points ranging from -4 to +5. How many points each alien gives changes throughout the experiment according to a Gaussian random walk. The deterministic transition between the spaceships and the planets stays the same throughout the experiment and participants are explicitly instructed about it. This represents the participants' "knowledge about the regularities in the world". In this version of the task, referring to this mapping when making decisions leads to higher payoff in the task (Kool et al., 2016).

Chapter 3: Blocking D2/D3 dopamine receptors in male participants increases volatility of beliefs when learning to trust others

Original manuscript as published in Nature Communications

DOI: 10.1038/s41467-023-39823-5

Nace Mikus[1, 2, *], Christoph Eisenegger[1, 3, §] Chris Mathys[2, 4, 5], Luke Clark[6, 7], Ulrich Müller[3, 8], Trevor W. Robbins[3], Claus Lamm[1, +, *], Michael Naef[9, +, *]

[+] These authors jointly supervised this work: Lamm, Claus, Naef, Michael

[§] Deceased on 27th February 2017

[*] Correspondence to:

Nace Mikus, nace.mikus@univie.ac.at,

Claus Lamm, claus.lamm@univie.ac.at, or

Michael Naef, michael.naef@durham.ac.uk

Affiliations:

[1] Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Austria

[2] Interacting Minds Centre, Aarhus University, Aarhus, Denmark

[3] Behavioural and Clinical Neuroscience Institute and Department of Psychology, University of Cambridge, Cambridge, UK

[4] Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zurich and ETH Zurich, Zurich, Switzerland

[5] Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy

[6] Centre for Gambling Research at UBC, Department of Psychology, University of British Columbia, Vancouver, BC, Canada

[7] Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, BC, Canada

[8] Adult Neurodevelopmental Services, Health & Community Services, Government of Jersey

[9] Department of Economics, University of Durham, Durham, UK

https://www.nature.com/articles/s41467-023-39823-5

Abstract

The ability to learn about other people is crucial for human social functioning. Dopamine has been proposed to regulate the precision of beliefs, but direct behavioural evidence of this is lacking. In this study, we investigate how a high dose of the D2/D3 dopamine receptor antagonist sulpiride impacts learning about other people's prosocial attitudes in a repeated Trust game. Using a Bayesian model of belief updating, we show that in a sample of 76 male participants sulpiride increases the volatility of beliefs, which leads to higher precision weights on prediction errors. This effect is driven by participants with genetically conferred higher dopamine availability (Taq1a polymorphism) and remains even after controlling for working memory performance. Higher precision weights are reflected in higher reciprocal behaviour in the repeated Trust game but not in single-round Trust games. Our data provide evidence that the D2 receptors are pivotal in regulating prediction error-driven belief updating in a social context.

Introduction

Knowing whom to trust with our money, information, or health is central to our personal well-being (Meyer-Lindenberg & Tost, 2012). The ability to form beliefs about other persons' attitudes from their actions is therefore pivotal for successfully navigating our social world. Inflexible beliefs, particularly about intentions of others, can lead to thoughts of persecution or even paranoid delusions - a hallmark symptom of psychotic disorders (Diaconescu, Wellstein, Kasper, Mathys, & Stephan, 2020; Gromann et al., 2013; Wellstein et al., 2020). Understanding the neurocomputational substrates of social inference is therefore essential for informing pharmacological treatments of psychotic symptoms.

When learning whether to trust another person, we often do so by observing their behaviour across repeated interactions. How behaviours of others affect our overall beliefs about their trustworthiness largely depends on how certain we are about the attitudes that presumably drive others' actions (FeldmanHall & Shenhav, 2019a). For instance, if we believe someone will be hostile or friendly towards us with high certainty, any gesture from them will not much change our belief about them. On the other hand, that same gesture from someone whose intentions we are unsure of will likely strongly shift what we think about them. This process of belief updating under uncertainty has been formalized within the Bayesian Inference framework, where beliefs are represented as probability distributions and the degree to which new information affects the updating of beliefs is modulated by the precision (the inverse of uncertainty) of those beliefs (Mathys et al., 2011). As in similar computational frameworks (Sutton & Barto, 2017), the belief update is proportional to the deviation of the prediction from the actual outcome, termed a prediction error (PE), weighted by the precision of prior beliefs. On top of this, new information also reduces uncertainty about the outcome. When prior beliefs are highly uncertain, the weight on the PE will be high and beliefs will be highly volatile. Conversely, if beliefs are held with high precision, this leads to a down-regulation of the influence of PE on learning and lowers belief volatility. Inflexibility in forming beliefs about others proportionally to their actions can result from high precision of prior beliefs about others' attitudes. Yet, the neurocomputational and neurochemical mechanisms regulating the uncertainty of beliefs are poorly understood. In this study, we examined the effects of the antipsychotic drug sulpiride, a D2/D3 dopamine receptor antagonist, on the uncertainty of beliefs about another person's trustworthiness.

Seminal studies in animals have established that mesolimbic dopaminergic circuits carry PE signals that drive belief updating in various contexts (Matsumoto & Hikosaka, 2009; Montague, Dayan, & Sejnowski, 1996; Schultz, 1998). However, dopaminergic midbrain neurons have been shown to be involved in various probabilistic computations that go well beyond phasic signalling of surprising rewarding events. Dopamine responses scale with outcome variance (Schultz et al., 2008; Tobler et al., 2005) and reflect temporal and perceptual precision (De Lafuente & Romo, 2011; Fiorillo et al., 2008; Fiorillo, Tobler, & Schultz, 2005). Several computational accounts of brain function suggest that uncertainty or precision coding is the main unifying feature of dopamine's involvement in belief updating (Friston et al., 2014; Gershman, 2018a; Gershman & Uchida, 2019; Mikhael & Bogacz, 2016).

Through encoding of uncertainty of beliefs about the world and what action to perform, dopamine receptors are thought to adjust the weights on PEs and control action selection (Adams et al., 2013; Babayan, Uchida, & Gershman, 2018). But while there is evidence for the involvement of dopamine receptors in processing uncertainty in action selection (Adams et al., 2020; Eisenegger et al., 2014, 2013; Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015), evidence for their causal role in regulating the uncertainty of social beliefs and adjusting weights on PEs is lacking.

Dopamine receptors within the corticostriatal circuitry are ideally positioned to regulate PE-related signal propagation and encode precision (Friston, 2008; W. D. Yao, Spealman, & Zhang, 2008). One possible neurobiological substrate of precision is proposed to be the post-synaptic gain of neuronal populations reporting PEs, where synaptic gain refers to the amplification or attenuation of the pre-synaptic signal on the post-synaptic cell (Adams et al., 2013; Friston et al., 2012). Post-synaptic D1 and D2 type dopamine receptors in the striatum have complementary effects on synaptic gain (Frank, 2005; W. D. Yao et al., 2008). D1-like receptors increase the excitability of post-synaptic neurons, whereas D2-type attenuate signal propagation and decrease synaptic gain (Frank, 2005). A prediction from this is that dopamine binding to D1 receptors would promote PE propagation and increase belief updating. In contrast, D2 receptor stimulation would reduce post-synaptic responses and attenuate changes in beliefs, leading to belief rigidity (Adams, Huys, & Roiser, 2016; Adams, Vincent, Benrimoh, Friston, & Parr, 2021). In line with this reasoning, when learning about others, blocking D2 receptors should increase the volatility (or rate of change) of beliefs.

Although some studies indeed showed that blocking D2-type receptors enhanced learning from positive feedback (Frank & O'Reilly, 2006), led to pronounced PE-related activity in the striatum (Iglesias et al., 2021), and enhanced performance (Eyny & Horvitz, 2003; Jocham, Klein, & Ullsperger, 2011), there is also evidence for attenuated PE coding and greater variability in choice selection (Eisenegger et al., 2014; Jocham, Klein, & Ullsperger, 2014; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). The inconsistencies of these findings raise several important considerations. First, when alternative choices are available, it is often unclear whether increased switching between available choice options arises from drug effects on belief updating per se or from the effects on decision-making strategies (see for instance (L. Zhang, Lengersdorff, Mikus, Gläscher, & Lamm, 2020)). Second, D2 dopamine receptors have a higher affinity for dopamine (Richfield, Penney, & Young, 1989) and doses of D2 antagonists commonly used in studies with healthy participants might not be sufficiently high to block the D2 receptor driven regulation of the PE signal (Bressan et al., 2003). Third, administration of compounds binding to dopamine receptors can have different and even opposing effects on learning and decision-making, depending on genetic variation in baseline dopamine function (M. X. Cohen, Krohn-Grimberghe, Elger, & Weber, 2007; Cools & D'Esposito, 2011; Eisenegger et al., 2013). Fourth, integrating novel information with prior experiences also relies on working memory (Collins & Frank, 2012), which is also affected by drugs that target dopamine receptors (Mehta, Sahakian, McKenna, & Robbins, 1999). And finally, beyond the methodological limitations of previous work, most studies with dopamine receptor antagonists have examined learning about abstract stimulus-outcome associations using secondary rewards, which makes the translation to more complex social interactions questionable.

In light of these considerations, the present study administered a relatively high dose of the selective D2/D3 receptor antagonist sulpiride (800 mg) or placebo in a randomized double-blind parallel group design to 78 male participants, preselected based on their Taq1a polymorphism. The drug dose was chosen to maximise the blockade of postsynaptic dopamine D2 receptors while still being safe (Takano et al., 2006). Most previous work used doses of 400 mg which leads to an occupation of approximately 30% of D2 receptors (Mehta, Montgomery, Kitamura, & Grasby, 2008). Using 800 mg leads to more than 60% occupancy and increases the likelihood of sufficiently blocking the effect of D2 receptors. Furthermore, as mentioned above, the effect of D2 antagonists often interacts with baseline variation on dopamine function (M. X. Cohen et al., 2007; Westbrook et al., 2020). Taq1a polymorphism is one of the most widely investigated genetic variations of the D2 receptor. Individuals

with at least one A1 minor allele have been shown to have higher presynaptic dopamine availability (Laakso et al., 2005) and reduced D2 receptor density in some subdivisions of the striatum (Gluskin & Mickey, 2016; Smith et al., 2017). Blocking D2 receptors might therefore have a stronger effect on belief updating in that genetic subgroup.

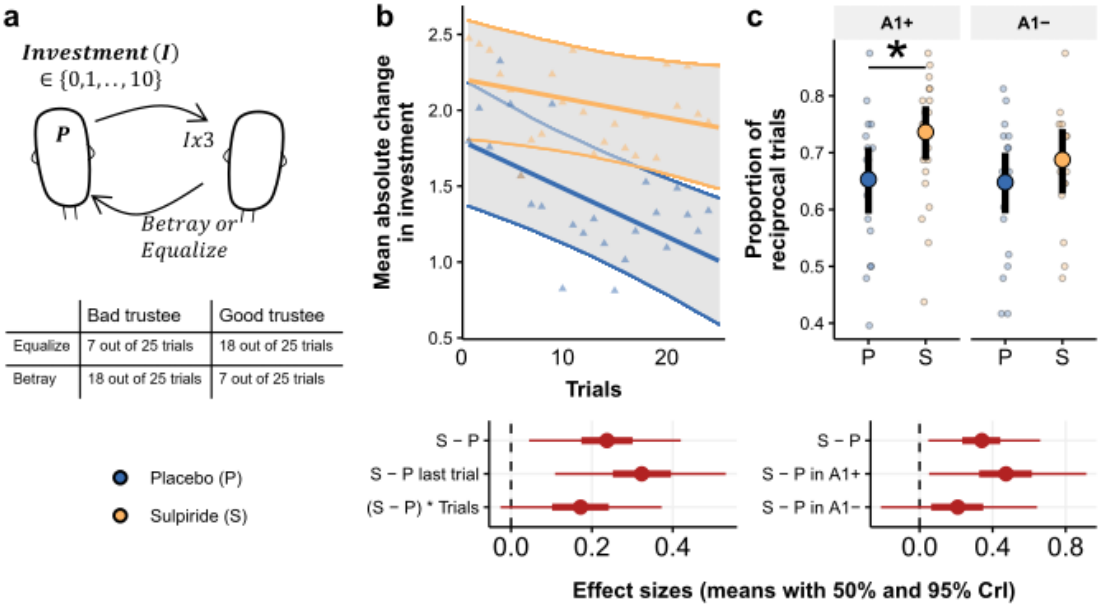


Figure 1. Effects of sulpiride on investment updates in the repeated Trust Game. a, The participants played 25 trials with two trustees. Each trial started with an endowment of 10 points to both players. On each trial they could invest any integer between 0 and 10. The trustee received a tripled amount of the investment and could decide to either equalize payoff or betray the other player and keep all the points for himself. The trustees were pre-programmed to be either "good" or "bad". **b**, Mean and 95% CrI of absolute change of investment from one trial to the next for both treatment groups based on a Bayesian multilevel model, plotted over raw means for each treatment group (△), obtained over n = 76 participants, n = 38 in each drug group, and with 2x25 trials per participant. Corresponding effect sizes with means, 50% and 95% CrI, for the main effect of sulpiride (S-P), for the effect of sulpiride in the last trial and for the interaction of the drug with Trial variable. **c**, Mean and 95% CrI of reciprocal trials (defined as trials where investment was increased following positive feedback, or decreased following negative feedback) based on a Bayesian logistic multilevel model, plotted over raw proportion of reciprocal trials with standard errors for each participant (sample sizes as in **b**). Effect sizes in log-odds shown for the main effect of sulpiride as well as the effect of sulpiride within each genotype group.

We investigated social learning by asking the participants to learn about other players' trustworthiness through a repeated Trust game (Figure 1a). In the Trust game the investor may choose to transfer any portion of their monetary endowment to the trustee (Berg, Dickhaut, & McCabe, 1995). The transferred points are then multiplied by the experimenter before being passed on to the trustee. The trustee can then either reciprocate in a way that equalizes the payoff of the two players or betray and keep everything. Participants in our study played 25 rounds of the Trust game as investors against two other players that were preprogramed to mostly equalize ("good trustee") or mostly betray ("bad trustee"). Importantly, we told the participants that the other players had given their answers weeks before the study day. Therefore, their decision to equalize or betray did not depend on the participant's investment. With this procedure, we increased the likelihood that their investments reflected the degree of uncertainty they had about the other player's response and were not confounded by strategic investment strategies, or exploratory action policies. By asking the participants to learn about a stable feature, we also ensure that participants' behaviour did not reflect differences in beliefs in the task volatility (the likelihood that the other person changed their mind), which might have obscured more basic processes related to forming beliefs about others.

The main goal of the study was to test the hypothesis that blocking D2-type receptors increases belief updates by reducing the precision of beliefs about others, whereby we also hypothesized that this

effect would be more pronounced in participants with genetically conferred higher endogenous dopamine levels. The Results section of the paper is structured as follows: we first looked at how sulpiride affected investment updates and how this effect was moderated by the Taq1a genotype. We then examined how the updates related to the back-transfer of the trustee, by examining the effects of the drug and genotype on reciprocal behaviour. We then turned to computational modelling to determine how sulpiride affects the course of each participant's uncertainty around the other player's trustworthiness. To evaluate to what degree the effects of sulpiride were due to actions on working memory we included data on spatial working memory performance into our parameter estimation process. Finally, to control for effects of sulpiride on sensitivity to social feedback unrelated to learning, we surveyed data from two single-round social interaction tasks, targeting positive and negative reciprocal behaviour.

With this we show that sulpiride has a profound effect on how healthy males update their investment more from one trial to the next this effect is driven by increased uncertainty around the other player's actions. Results from both behavioural analysis and computational modelling support the claim that the drug effect on belief updating was more pronounced in participants with higher endogenous striatal dopamine levels (indexed by the Taq1a polymorphism).

Results

D2/D3 receptor antagonism increases investment updates

We employed a Bayesian multi-level linear model predicting absolute change in investment from the previous trial, including variables for Treatment (sulpiride or placebo), Trial and their interaction as predictors (refer to supplementary material for outcomes of alternative models). Figure 1b shows that, following sulpiride administration, participants on average updated their investments more than participants in the placebo group ($b = 0.633$, 95% Credibility Interval (CrI) [0.117, 1.115], proportion of the posterior distribution of the regression coefficient below 0 being $P(b<0) = 0.005$), with an effect size $d = 0.239$ (95% CrI [0.045, 0.42]). The difference in investment updates was most apparent in the last trial of the task ($b = 0.863$, 95% CrI [0.289, 1.411], $P(b<0) = 0.002$, $d = 0.325$, 95% CrI [0.109, 0.531]) and we also found a small effect size on the Trial*Treatment interaction ($b = 0.457$, 95% CrI [-0.069, 0.99], $P(b<0) = 0.047$, $d = 0.172$, 95% CrI [-0.026, 0.373]). As participants learned about the trustees, changes of investments from one to the next trial reduced, and this decrease across time was less pronounced in the sulpiride group.

To examine whether the effects of the drug were moderated by the Taq1a polymorphism we ran another model including a variable for Taq1a-specific genotype and Trustee as predictors with the four-way interaction between the two new variables, Treatment and Trial, including a random intercept and slope for the Trustee (Supplementary Figure 1a, Supplementary Table 4). Again, we found a main effect of treatment ($b = 0.595$, 95% CrI [0.112, 1.098], $P(b<0) = 0.008$), and a significant three-way interaction between Treatment, Genotype and Trial number ($b = 0.053$, 95% CrI [0.01, 0.098], $P(b<0) = 0.007$), but found no credible evidence of a two-way interaction Treatment x Genotype ($b = -0.284$, 95% CrI [-1.266, 0.708], $P(b>0) = 0.287$). These analyses suggest that on average sulpiride affected investment updates comparably across both genotype groups, but in contrast to the A2 homozygotes, the effect in the A1+ group was time dependent.

D2/D3 receptor antagonism increases sensitivity to social feedback in the A1+ group in the repeated Trust game

To further understand how investment updates related to back-transfer from the trustee, we defined reciprocal trials as trials where participants either increased investments (or repeated the maximal investment of 10 points) following positive feedback and decreased investments (or repeated an investment of 0 points) following a betrayal (Figure 1c, for exact definition see Supplementary Note

2). We found that sulpiride led to a higher proportion of reciprocal trials ($b_{logodds}$= 0.339, 95% CrI [0.048, 0.661], P($b_{logodds}$ <0) = 0.012).This effect was significant in the A1+ group ($b_{logodds}$= 0.469, 95% CrI [0.052, 0.914], P($b_{logodds}$<0) = 0.015) but we found no credible evidence for an effect in the A1- group ($b_{logodds}$= 0.209, 95% CrI [-0.212, 0.643], P($b_{logodds}$<0) = 0.162); however, we also found no credible evidence that there was a difference of drug effects between the two genotype groups ($b_{logodds}$ = -0.263, 95% CrI [-0.867, 0.329], P($b_{logodds}$>0) = 0.186). Furthermore, we found some support for a dose response effect, whereby sulpiride serum levels in the blood correlated with reciprocal trials in the A1+ group (b = 0.185, 95% CrI [-0.04, 0.41], P(r<0) = 0.05), but found no credible evidence for a correlation in the A1- group (Supplementary Table 1). Similar, albeit weaker, effects were found when we examined to what extent the signed investment update was dependent on the back-transfer and how this differed across the drug and genotype groups (Supplementary Figure 1b).

No credible evidence of an effect of D2/D3 receptor antagonism on average investment behaviour or overall performance

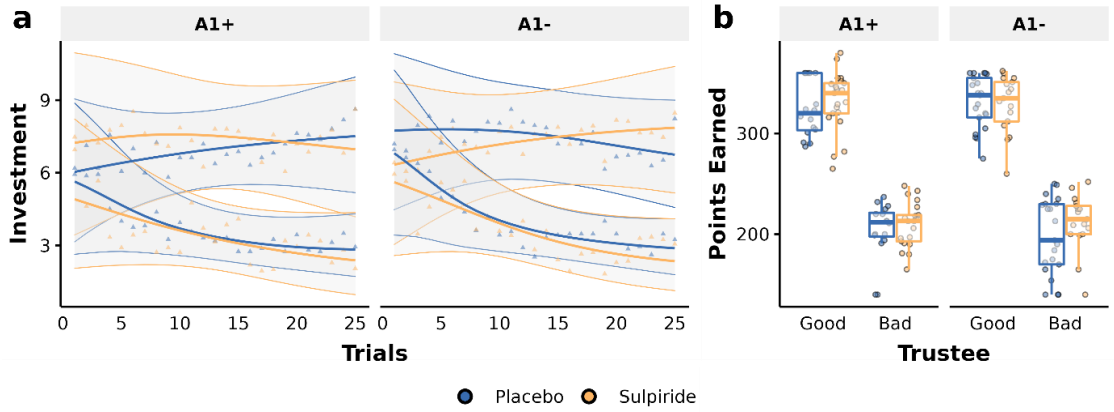


Figure 2. Effects of sulpiride on average investments in the repeated Trust Game. a, Average investment behaviour grouped for each trustee. Lines depict model predictions (means with 95% CrI as the shaded area), plotted over raw means for each drug group (△), obtained for 2x25 trials per participant, n = 76 participants (in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride). We found no credible evidence for an effect of sulpiride in either genotype group on average investment behaviour regardless of the trustee. For statistics, refer to the main text and Supplementary Figure 2. **b**, Overall points earned in the task grouped for each trustee. Dots are average points earned for each participant. Boxplots with centre lines as medians, box bounds as 25th and 75th percentiles, and whiskers terminating at maxima/minima (a distance of 1.5 times the IQR away from the 25th and 75th percentiles). Sample sizes as in **a**.

Next, we investigated whether this higher change of investments from one trial to the next is reflected in average investment patterns (Figure 2a, for more detailed plots see Supplementary Figure 2). Using an ordinal logistic model predicting investments from Treatment, Genotype, Trustee and Trial variables, with a random slope for Trustee and Trial we found no credible evidence of a difference between sulpiride and placebo on average investment behaviour either in the A1+ group (b_{good} = -0.011, 95% CrI [-1.95, 1.796], P(b_{good}>0) = 0.494, b_{bad} = 0.089, 95% CrI [-2.256, 2.47], P(b_{bad}<0) = 0.47), nor in the A1- group (b_{good} = -0.496, 95% CrI [-2.353, 1.367], P(b_{good}>0) = 0.297, b_{bad} = 0.821, 95% CrI [-1.508, 3.257], P(b_{bad}<0) = 0.244). We also found no credible evidence of a difference in initial investments across the four drug and genotype groups (Supplementary Figure 2). The overall initial investment was estimated to be, on average, 6.33 (95% CrI [3.129, 9.687]), suggesting that most participants expected a positive back-transfer initially. In line with this, the slope when playing against the good trustee was positive (b = 0.086, 95% CrI [-0.008, 0.187], P(b<0) = 0.036), but not as pronounced as the slope when playing against the bad trustee (b = -0.153, 95% CrI [-0.284, -0.027], P(b>0) = 0.009). While we found no credible evidence of a difference between slopes across the drug groups in the A1+ participants (b_{good} = -0.058, 95% CrI [-0.184, 0.065], P(b_{good}>0) = 0.179, b_{bad} =

0.044, 95% CrI [-0.125, 0.21], P(b_{bad}<0) = 0.297), we did observe an increase in the slope following sulpiride administration in the A1- group when playing against the bad trustee (b_{bad} = 0.159, 95% CrI [-0.007, 0.336], P(b_{bad}<0) = 0.03) but not when playing against the good trustee (b_{good} = 0.069, 95% CrI [-0.06, 0.19], P(b_{good}<0) = 0.14). Similarly, we found no credible evidence of a difference across drug and genotype groups regarding how many points they earned when playing against either trustee (Figure 2b, Supplementary Table 13).

In summary, we found no credible effect of sulpiride on investing behaviour on average, but we do find some evidence in support of sulpiride increasing sensitivity to social feedback when learning about others. To determine whether and how this behavioural pattern relates to the uncertainty of participants' beliefs about the other persons' trustworthiness, we explicitly modelled the participants' trial-by-trial evolution of beliefs with a Bayesian belief model.

Computational framework

The belief model uses a hierarchical Gaussian filter (HGF) to generate trial-wise sequences of participants' beliefs about the trustworthiness of two trustees as well as the uncertainty (or precision) surrounding those beliefs (, (Mathys et al., 2011, 2014)). We estimated a participant-specific parameter ω, called *belief volatility*, that describes how each participant's precision of beliefs evolved over time and consequently determined the relative rigidity (or malleability) of beliefs. More specifically, on each trial, we approximate the latent belief about the trustworthiness of the other player as a gaussian distribution with a specific mean and variance. Higher belief volatility ω implies higher variance (or lower precision) of trial-by-trial belief estimates. Importantly, the dynamic learning rate (ψ_t) on the PE is proportional to the expected variance or inversely proportional to the precision of beliefs and is therefore referred to as a "precision-weight". Low precision of prior beliefs leads to higher precision-weighted learning rates and stronger shifts in beliefs throughout the task (see two example belief trajectories with different ω values in Figure 3b).

The beliefs about trustworthiness are mapped on to probability of positive or negative feedback with an inverse logistic function. Because D2 receptor activity is linked to choice uncertainty and action variability (Adams et al., 2020; Eisenegger et al., 2014), we also included another parameter called choice precision parameter γ that determined the nonlinear mapping from beliefs to the investments. Higher choice precision implies an investment distribution centred around extremes (i.e., investing 0 and 10), and lower values imply a more dispersed investment distribution and more uncertainty or stochasticity in action selection. It thus mirrors the stochastic aspect of the inverse temperature parameter in the softmax equation often used in non-ordinal (e.g., binary) choice tasks. Finally, how beliefs about the probability of a positive back-transfer affect investment behaviour is determined through an ordinal logistic likelihood function. The degree to which inferred trustworthiness correlates with investments is determined by another parameter called the trustworthiness slope (η). Crucially, the computational parameters of the model represent distinct behavioural patterns and can be recovered reliably (Figure 3c). To determine how noisy trials are represented in the model, we defined mistake trials as trials where participants either decreased their investment after a positive back-transfer or increased their investment after a betrayal (for exact definition see Supplementary Note 1). Importantly, we observed that belief volatility ω correlates with reciprocity (r = 0.277, t = 2.476, df = 74, p = 0.016, Figure 3d) confirming that higher trial-by-trial uncertainty of beliefs lead to a higher chance of reciprocal behaviour. The log-transformed choice uncertainty parameter γ correlates negatively with the proportion of mistake trials (r = -0.592, t = -6.3254, df = 74, p < 0.001, Figure 3d) implying higher randomness in investment selection. We also predicted data from the posterior distributions of parameters. We confirmed that the model captures the crucial aspects of behaviour (Figure 3e,f) and plotted average beliefs about the other player's trustworthiness, grouped for each trustee. (Figure 3g).

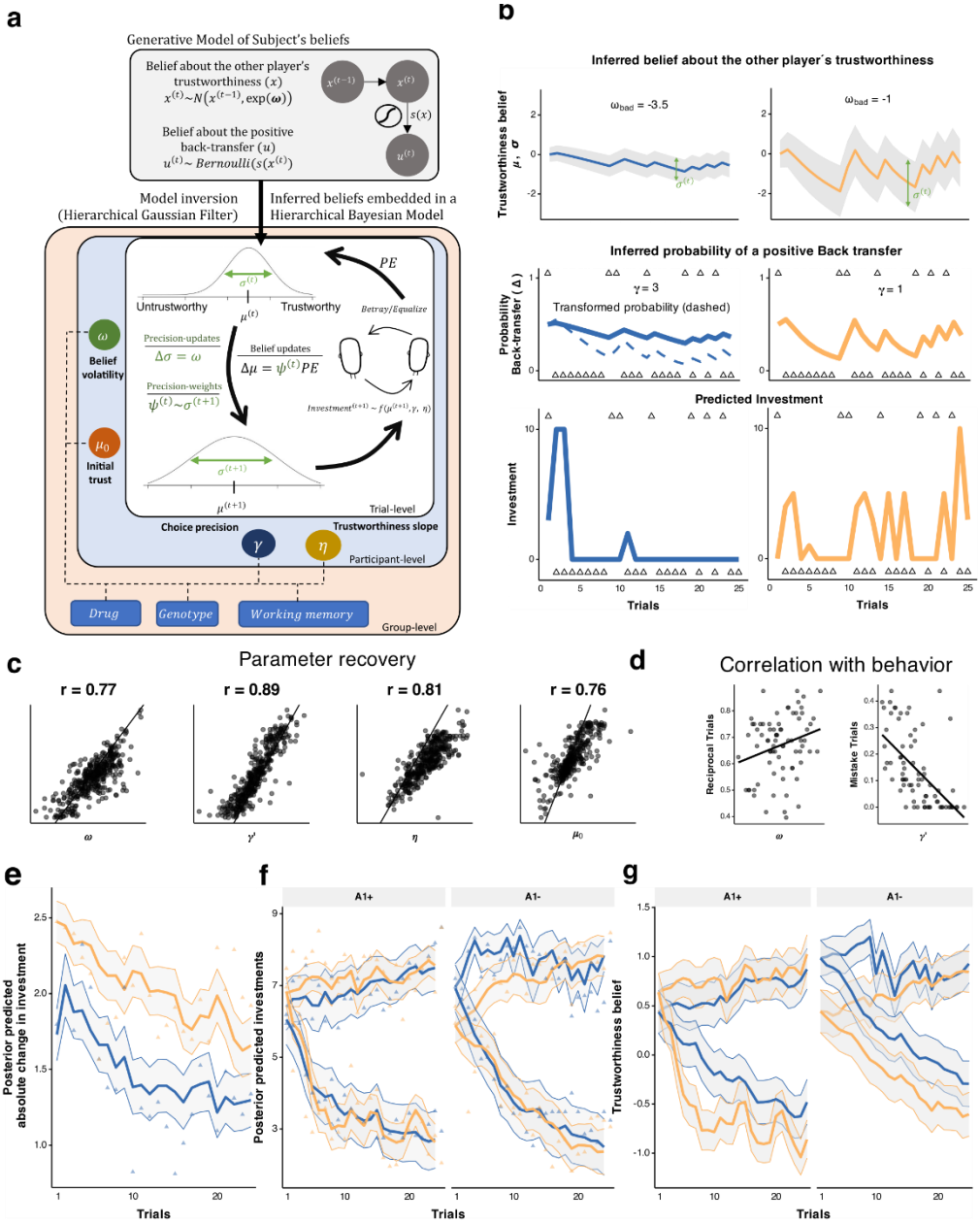


Figure 3. Computational modelling. a, We defined a generative model that describes the evolution of participants' beliefs about the other person's trustworthiness as a Gaussian random walk with the step size of ω. The hierarchical Gaussian filter (HGF) inverts this model and provides trial-level estimations of participants' beliefs about the trustworthiness of others as Gaussian variables with mean $\mu^{(t)}$ and standard deviation $\sigma^{(t)}$. The evolution of $\sigma^{(t)}$ is determined by the belief volatility parameter ω. The precision-weights $\psi^{(t)}$ are proportional to $\sigma^{(t)}$ and serve as dynamic learning rates when updating beliefs about the trustworthiness of the other player. We also estimate initial trustworthiness belief per participant (μ_0). The ordinal logistic link function governs how beliefs about others' trustworthiness map to investments with two additional subject-level parameters: choice uncertainty (γ) and the slope (η). The parameter estimation is done through hierarchical Bayesian inference, where we estimate all individual and group-level parameters in one inferential step. **b**, Two example belief trajectories portray the different behaviours that the model can capture, depicted as mean ($\mu^{(t)}$, line) and standard deviation of beliefs ($\sigma^{(t)}$, error band). The participants have different belief volatilities for the good (ω_{good}) and the bad trustee (ω_{bad}). Higher ω implies more uncertainty surrounding the trustworthiness beliefs ($\sigma^{(t)}$), which in turn leads to stronger belief shifts. **c**, For each participant, we randomly draw parameters from their individual posterior distribution, simulate data, and re-estimate them five times. Relative high correlations indicate that the model parameters are well-defined. **d**, The two main parameters of interest, belief volatility and choice uncertainty, correlate with distinct behavioural features, obtained for all participants (n=76). **e-f**, Posterior predictive for (**e**) absolute investment change from one trial to the next and (**f**) for the average investment behaviour. Plotted over raw means per trial per group and with standard deviations of predictions in the shaded area. **g**, Lines depict average beliefs about the trustworthiness across participants for each trials, with error bands depicting

average uncertainty around the investment (σ). All plots in **e-g** obtained with the following sample sizes: in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride.

We compared this model to an HGF model without the γ parameter and a Rescorla-Wagner (RW) model. The HGF model without γ has been used previously when modelling both social (Siegel, Mathys, Rutledge, & Crockett, 2018) and non-social (Adams et al., 2018) learning. In this model the nonlinear mapping from beliefs to probabilities is modulated by a coupling parameter κ (see Methods for details). The RW model is a simple Q-learning model with separate static learning rates for gains (positive outcomes) and for losses (negative outcomes). All models used the same ordinal-logistic likelihood function and were compared based on their trial-by-trial predictive accuracy through leave-one-out cross-validation information criterion (LOOIC) and expected log predictive density (ELPD). We found that the HGF model with the choice precision parameter γ outperforms both models (Supplementary Figure 3a). We also compared the models across trials and across trustees with the LOOIC and by looking at the correlations of predicted investments with actual behaviour (Supplementary Figure 3b, c). Interestingly, performance of both models varies similarly across trials with the HGF performing better across the whole task, particularly for investments against the good trustee.

D2/D3 receptor antagonism increases belief volatility

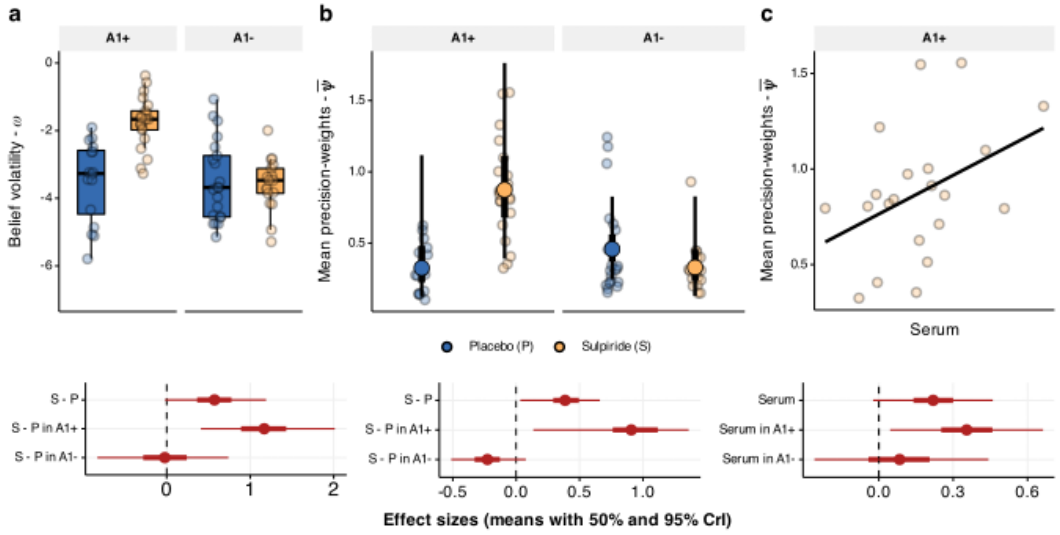


Figure 4. Effects of sulpiride on belief volatility and precision weights. a, Belief volatility boxplots over individual means of posterior distributions. Boxplots with centre lines as medians, box bounds as 25th and 75th percentiles, and whiskers terminating at maxima/minima (a distance of 1.5 times the IQR away from the 25th and 75th percentiles). Sample sizes in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride. Belief volatility is higher in the sulpiride group, and this effect is driven by the A1+ group (50% and 95% CrI of effect sizes below). **b** Precision-weights on PEs. Scattered points are meaned precision weights across all trials for each participant (sample sizes as in **a**). Overlayed group level medians with 50% and 95% CrI. The effect sizes were calculated from posterior distributions of differences in means across four groups. **c**, Precision weights correlate with log transformed serum levels in the blood, plotted for participants in the A1+ group that received sulpiride (n = 21). The effect sizes with means (and 50% and 95% CrI) depict correlations between serum levels and median precision weights for the sulpiride group (n=38), and then separately for the A1+ genotype group (n=21) and for the A1- genotype group (n=17).

For parameter estimation, we embedded the HGF derived equations in a hierarchical Bayesian model which allowed us to estimate the drug and genotype effects on all computational parameters in one inferential step (Kruschke, 2014; McElreath, 2018). Through this analysis, we found a main effect of sulpiride on volatility of beliefs (b = 0.831, 95% CrI [0.115, 1.533], P(b<0) = 0.01, d = 0.65, 95% CrI [0.088, 1.283], Figure 4a), and an interaction effect of sulpiride with the genotype (b = -1.506, 95% CrI [-2.649, -0.411], P(b>0) = 0.004, d = -1.175, 95% CrI [-2.238, -0.306]). In fact, the effect of sulpiride on belief volatility is driven by the A1+ allele carriers (b = 1.598, 95% CrI [0.727, 2.465], d = 1.25, 95%

CrI [0.533, 2.119]) while we found no credible evidence of an effect in the A1- group (b = 0.076, 95% CrI [-0.874, 0.985], d = 0.06, 95% CrI [-0.683, 0.783]).

The key consequence of higher belief volatility is that it leads to lower precision of prior beliefs and therefore of predictions, which has a direct effect on the learning rates. Indeed, we founnd credible evidence that participants under sulpiride have higher average precision-weights (d = 0.452, 95% CrI [0.081, 0.704], P(d<0) = 0.008, Figure 4b), particularly in the A1+ group (d = 1.042, 95% CrI [0.225, 1.424], P(d<0) = 0.003), but little credible evidence for an effect in the A2 homozygotes (d = -0.202, 95% CrI [-0.482, 0.103], P(d>0) = 0.089) with a significant interaction effect (d = 1.244, 95% CrI [0.335, 1.714], P(d<0) = 0.001). Importantly, in the A1+ group, this effect of sulpiride on precision-weighting correlated with the degree of serum levels in the blood (b = 0.356, 95% CrI [0.045, 0.663], P(b<0) = 0.013, Figure 4c).

Looking at potential asymmetries when dealing with uncertainty around beliefs about trustworthy or untrustworthy partners, we find that, on average, the volatility of beliefs about the bad trustee were more volatile (d = 0.412, 95% CrI [0.031, 0.827], P(d<0) = 0.018). When examining the drug effects, we observed that in the A1+ group, the difference in ω between placebo in sulpiride is apparent in interactions with both trustees (d_{bad} = 1.62, 95% CrI [0.731, 2.644], P(d_{bad} <0) < 0.001, d_{good} = 0.689, 95% CrI [-0.089, 1.575], P(d_{good} <0) = 0.042, but is higher for the bad trustee ($d_{good-bad}$ = -0.923, 95% CrI [-1.754, -0.121], P($d_{good-bad}$ >0) = 0.01, Supplementary Figure 4b, c). Interestingly, this analysis also showed that in the A1- group, there is a significant interaction of sulpiride and trustee effects ($d_{good-bad}$ = -1.453, 95% CrI [-2.529, -0.51], P($d_{good-bad}$ >0) = 0.001, Supplementary Figure 4b, c), whereby we find little credible evidence that the effects of sulpiride on belief volatility are higher for the bad trustee (d_{bad} = 0.711, 95% CrI [-0.21, 1.669], P(d_{bad} <0) = 0.066), and even negative for the good trustee (d_{good} = -0.742, 95% CrI [-1.74, 0.091], P(d_{good} >0) = 0.042). At this point, we also note that our model suggests that participants expected the trustee to reciprocate (d = 0.717, 95% CrI [0.406, 1.023], P(d<0) = 0.001) with initial inferred probability of reciprocation being 0.67 (95% CrI [0.60, 0.73]). However, we found no credible evidence of a difference between treatment groups in initial beliefs (μ_0) about the trustworthiness either overall (d = -0.166, 95% CrI [-0.697, 0.37], P(b>0) = 0.272, Supplementary Figure 4), in the A1+ group (d = 0.111, 95% CrI [-0.585, 0.802], P(d<0) = 0.383), or in the A1- group (d = -0.441, 95% CrI [-1.164, 0.287], P(d>0) = 0.112).

We then compared the results from the HGF model to those of the RW model (Supplementary Figure 5). The evidence points in same direction, with some evidence for sulpiride leading to higher learning rates overall (d = 0.315, 95% CrI [-0.087, 0.729], P(d<0) = 0.064, Supplementary Figure 5a), an effect that the A1+ participants drove (d = 0.593, 95% CrI [0.03, 1.16], P(d<0) = 0.02), with no credible evidence of an effect in the A1- participants (d = 0.037, 95% CrI [-0.524, 0.605], P(d<0) = 0.453), and little credible evidence of a difference between the effect (d = -0.553, 95% CrI [-1.343, 0.235], P(d>0) = 0.079). Further, the effect of the drug in the A1+ group was observed both when learning about positive outcomes (d = 0.697, 95% CrI [0.029, 1.369], P(d<0) = 0.021, Supplementary Figure 5b) as well as negative outcomes (d = 0.488, 95% CrI [-0.059, 1.044], P(d<0) = 0.039, Supplementary Figure 5c). However, the difference across the two types of learning rates was not as pronounced as in the HGF model.

D2/D3 receptor antagonism increases choice uncertainty

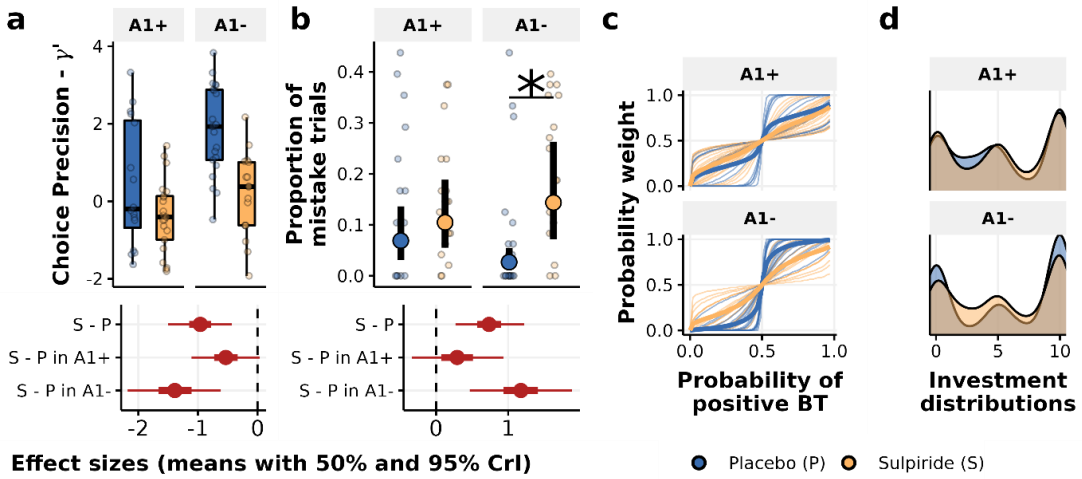


Figure 5. Effect of sulpiride on choice precision. a, Effects of sulpiride on choice uncertainty, estimated in log – space (hence the prime). Dots are participant-level posterior means. The 95% and 50% CrI of effect sizes show a main effect of sulpiride, driven by the A1- group. Boxplots with centre lines as medians, box bounds as 25th and 75th percentiles, and whiskers terminating at maxima/minima (a distance of 1.5 times the IQR away from the 25th and 75th percentiles). Sample sizes in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride. **b**, Mean proportion of mistake trials (with SEM), with samples sizes as in **a**. Means and 95% quantiles of posterior distributions across the four groups are plotted based on a logistic regression model. Corresponding effect sizes below. **c**-**d**, The choice uncertainty parameter determines the probability weight (**c**) and therefore the investment distribution (**d**). Higher values for the placebo group (the A1 group in particular) indicate more extreme investment choices and higher belief inflexibility.

In addition to the effect on belief volatility, sulpiride also increases choice uncertainty by decreasing the choice precision parameter γ (b = -1.049, 95% CrI [-1.6, -0.502], P(b<0) < 0.001, d = -0.979, 95% CrI [-1.535, -0.455], Figure 5a), with smaller effects in the A1+ group (b = -0.646, 95% CrI [-1.272, -0.033], P(x>0) = 0.02, d = -0.608, 95% CrI [-1.206, -0.031]) and more prominent effects in the A2 group (b = -1.44, 95% CrI [-2.261, -0.639] , P(b<0) < 0.001, d = -1.351, 95% CrI [-2.133, -0.601]). Since lower values of γ correlated with higher proportion of mistake trials we examined how sulpiride affected the proportion of mistake trials and found that it on average increased the number of mistake trials ($b_{logodds}$ = 1.172, 95% CrI [0.443, 1.992], P($b_{logodds}$ <0) < 0.001, Figure 5b), an effect driven by the A1- group ($b_{logodds}$ = 1.876, 95% CrI [0.781, 3.032], P($b_{logodds}$ <0) < 0.001) with no credible evidence for an effect in the A1+ group ($b_{logodds}$ = 0.468, 95% CrI [-0.535, 1.537], P($b_{logodds}$ <0) = 0.184), and some evidence for an interaction effect ($b_{logodds}$ = 0.885, 95% CrI [-0.041, 1.857], P($b_{logodds}$ <0) = 0.03). The effect of sulpiride on the proportion of mistake trials in the A1- group was proportional to the blood serum levels ($b_{logodds}$ = 0.607, 95% CrI [0.089, 1.142], P($b_{logodds}$ <0) = 0.011) with no credible evidence of a correlation of the A1+ group ($b_{logodds}$ = -0.328, 95% CrI [-0.818, 0.143], P($b_{logodds}$ >0) = 0.085). This parameter also determines the skew in the distribution of investments, whereby higher values make extreme investments more likely (Figure 5c, d). From a perspective of an expected utility maximising agent, extreme investments are most optimal (Supplementary Note 2). Individuals with higher γ therefore behave more as rational agents and take the uncertainty of the outcome less into consideration when choosing investments. Sulpiride also increased the η parameter (b = 1.459, 95% CrI [0.532, 2.42], P(b<0) < 0.001, d = 0.941, 95% CrI [0.331, 1.58]), further advocating for the assertion that sulpiride increased the degree to which beliefs about trustworthiness influenced participants' investments. In sum, the overall results from the computational modelling suggest that sulpiride treatment led to higher choice uncertainty (lower choice precision), which was related to increased mistakes in the in the A1- group specifically. We found strong support for sulpiride increasing belief volatility and precision-weights on PEs, an effect that was driven by the A1+ group, whereas we found no credible evidence of an effect in the A1- group.

Effects of sulpiride on belief updating remain after accounting for working memory performance

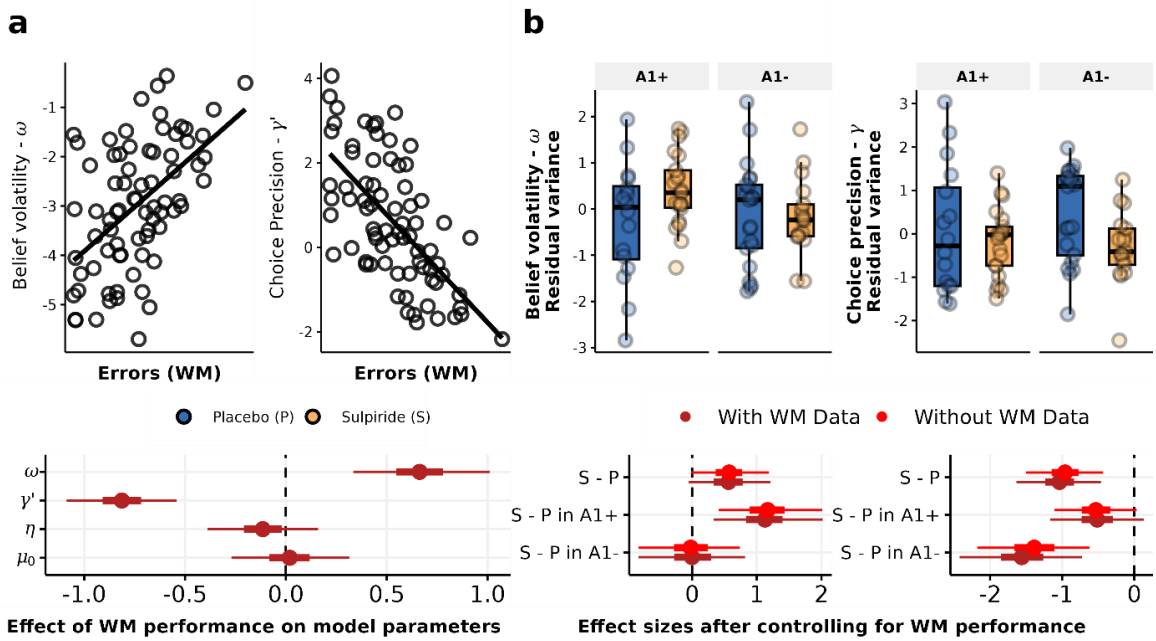


Figure 6. Working memory performance and computational parameters. a, We reran the parameter estimation with a multilevel model that only included working memory data (number of errors) at the group level effect (and was agnostic about drug and genotype groups). Poorer performance in the spatial working memory task correlated positively with belief volatility ω and negatively with choice precision γ and did not affect noise or initial trustworthiness (obtained from sample size n = 75). Effect sizes depicted with means, 50% and 95% CrIs. **b**, Residual variances after accounting for working memory data from the model that is agnostic about drug and genotype data, for parameters ω and γ. Boxplots with centre lines as medians, box bounds as 25th and 75th percentiles, and whiskers terminating at maxima/minima (a distance of 1.5 times the IQR away from the 25th and 75th percentiles), with the following samples: in A1+ group, n=16 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride. In the second step, the parameters were estimated with working memory data and drug and genotype variables at the group level. The results of this analysis are shown below as effect sizes with means, 50% and 95% CrIs. The analysis is compared to that with the model that does not include working memory data.

In the repeated Trust game, participants must remember the trustees' responses to previous trials. Higher choice stochasticity could therefore be due to poorer working memory. Furthermore, the inability to remember outcomes of past trials might increase the reliance on the previous trial and thereby cause increased learning rates and belief volatility. To determine to what degree our findings were influenced by the possible effects of sulpiride on working memory, we included data from a spatial working memory (WM) task performed in the same sample and previously published (Naef et al., 2017). In the spatial WM task, participants uncover 'tokens' from sets of boxes, whereby they need to remember which boxes were previously searched and what the outcomes of those searches were (see Methods for details). As Naef et al. report (Naef et al., 2017), sulpiride had a detrimental effect on working memory performance, whereby participants in both genotype groups performed more errors (opened boxes previously already opened) in more challenging task trials (trials with 10 or 12 boxes).

In the present study, we first investigated whether the model parameters are influenced by WM performance. To do so, we re-estimated the hierarchical model that included only WM data at the group level without the drug and genotype variables. As can be seen from Figure 6a, the belief volatility parameter ω was associated with a higher number of errors in the WM task (d = 0.658, 95% CrI [0.334, 1.01], P(d<0) < 0.001), and the Choice precision parameter γ negatively correlated with the number of errors (d = -0.811, 95% CrI [-1.087, -0.541], P(d>0) < 0.001). This implies that poorer working memory performance is related to higher choice and belief uncertainty. Importantly, however, when plotting the residuals of the model parameters (unexplained variance after accounting for WM effects), the impact of sulpiride on belief volatility in the A1+ group can be seen still to be present (Figure 6b). To obtain posterior distributions of drug and genotype effects after accounting for WM data, we re-

estimated the parameters of the model this time including WM data as well as drug and genotype variables. We found that including WM information in the hierarchical model only slightly changed the inference about the effect of sulpiride on belief updating. The main effect of sulpiride on ω was now somewhat less certain with the 95% CrI including values below 0 (d = 0.56, 95% CrI [-0.052, 1.211], P(d<0) = 0.036), but the effect in the A1+ group was still present (b = 0.852, 95% CrI [0.116, 1.61], P(b<0) = 0.014, d = 0.694, 95% CrI [0.093, 1.369]). Similarly, posterior intervals of sulpiride effects on γ after including WM data were comparable to those without WM data, with the main effect remaining negative (d = -1.034, 95% CrI [-1.634, -0.461], P(d>0) < 0.001), and the evidence for the effect is substantial in the A1- group (d = -1.562, 95% CrI [-2.423, -0.722], P(d>0) < 0.001) and less so in the A1+ group (d = -0.512, 95% CrI [-1.168, 0.133], P(d>0) = 0.056).

An important final step was to exclude the possibility that this increase in updating was due to increased sensitivity to social feedback in general, or due to decreased desire to maximise outcomes. For this, we turned to data from single-round social interaction games that measure learning-independent positive and negative reciprocity.

No credible evidence for an effect of D2/D3 receptor antagonism on single-round reciprocal behaviour

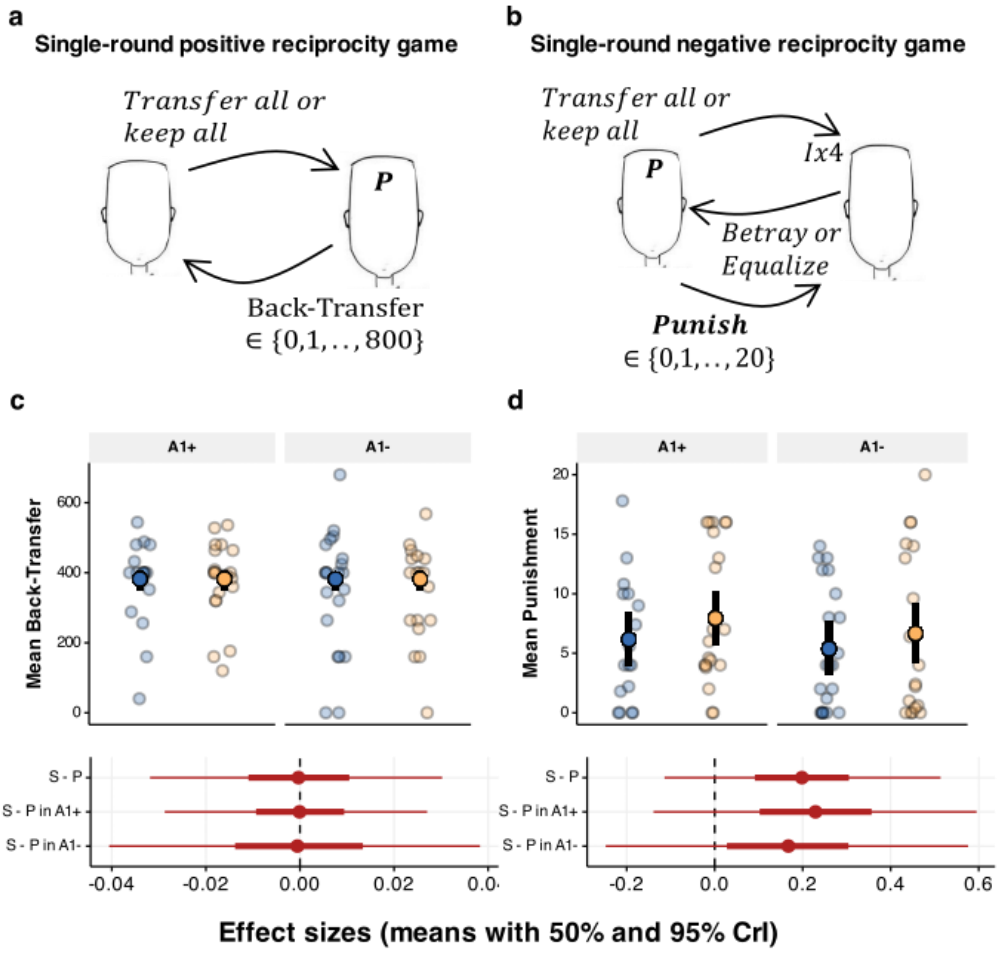


Figure 7. Single-round reciprocity games. a, In the positive reciprocity task, participants played as trustees were paired with 7 other players. The investor in this version of the game received 800 points and could either keep everything or give everything to the trustee, who could then decide how to split the points. The investors were pre-programmed so that 5 out of 7 transferred everything to the trustee. **b**, In the single-round negative reciprocity game, the participants played the investor. In the beginning of the round both players were given 10 points. The investor could then decide to transfer everything or nothing. The transferred investment got multiplied by a factor of four and the trustee could them decide to either equalize or betray. Crucially, after the choice of the trustee, both players received another 20 points and the investor could use his points to punish the trustee, with a factor of three. **c**, **d**, Mean and 95% CrI of Back-transfer (**c**) and punishment (**d**) across the four

groups, plotted over raw means (+/- SEM) per participant. Means, 50%, and 95% CrIs of effect sizes shown below, with the following sample sizes: in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride.

In the single-round interaction games the participants played a slightly modified versions of the Trust game. In the positive reciprocity game, they played the trustee and could reward the investor for their decision (Figure 7a). In the negative reciprocity game, they played as investor and could punish the trustee (Figure 7b). We found no credible evidence of a difference between sulpiride and placebo, neither in the amount of reward (Back-transfer) in the positive reciprocity game (b = -0.023, 95% CrI [-6.605, 6.263], d = 0.000, 95% CrI [-0.032, 0.03], Figure 7c) nor in punishing behaviour in the negative reciprocity game (b = 1.552, 95% CrI [-0.903, 3.98], P(x<0) = 0.106, d = 0.2, 95% CrI [-0.114, 0.513], Figure 7d). This implies that the drug effect on reciprocal behaviour in the Repeated Trust Game was not due to higher sensitivity to social-feedback, or to less rational behaviour.

Discussion

Inferring attitudes of others is fundamental to our social functioning, but the neurocomputational mechanisms of the updating of beliefs about others are not well understood. We show that blocking D2/D3 dopamine receptors by sulpiride has a profound effect on how healthy participants process uncertainty in a social context. When playing as investors in the Repeated Trust Game, participants given a high dose this D2/3 receptor antagonist changed their investment more from one trial to the next. Using a hierarchical Gaussian filter to explicitly model the evolution of participants' beliefs about the trustworthiness of the trustees, we show that sulpiride increased belief volatility. This implies that for the participants under sulpiride, the beliefs about the trustworthiness of others were held with less precision (i.e., with higher uncertainty), which in turn caused increased precision weights on PEs. This effect was more pronounced in participants with at least one minor A1 allele of the Taq1a polymorphism, associated with higher endogenous striatal dopamine levels. The increase in precision weights on PEs in that genetic subgroup scaled with the sulpiride serum levels in the blood. As a consequence, sulpiride led to higher reciprocal behaviour (increased investment after positive back-transfer and decreased investment after negative back-transfer), but only in the repeated Trust game, whereas we found no credible evidence for an effect in single-round interactions. The effect on the repeated Trust game was present even after controlling for working memory performance. Moreover, sulpiride decreased the value of the parameter of the model that codes for deterministic action selection policies (γ), implying higher uncertainty about investment selection. The effect was present in both genotype groups.

On the neurophysiological level, it has been proposed that precision is encoded through the post-synaptic gain (i.e. amplification or blunting of presynaptic neuronal input) of neurons that propagate PE signals (Friston et al., 2012). Our results are in accordance with the idea that dopamine binding to D1 receptors of the medium spiny neurons in the striatum increases the gain on PE signals, while binding to D2 receptors decreases gain through disinhibition of the so called indirect pathway (Frank, 2005; W. D. Yao et al., 2008). Within this framework increased precision-weights following D2 antagonism can be explained by more dopamine being available to bind to D1-like receptors, a claim that is further substantiated by the observation that the effect of sulpiride was stronger in participants with genetically conferred higher presynaptic dopamine availability and lower D2 receptor density. These findings extend previous studies highlighting the role of dopamine receptors in coding precision or uncertainty in various contexts, such as perceptual and risk-based decision making (Adams et al., 2021; Guggenmos, Wilbertz, Hebart, & Sterzer, 2016; Schwartenbeck et al., 2015). In particular, previous work has shown that sulpiride decreased the perceived precision of temporal expectations (Tomassini, Ruge, Galea, Penny, & Bestmann, 2016). In a task where participants were explicitly told about the variance of outcomes, they adapted their behaviour accordingly, which led to more optimal choice performance (Diederen, Spencer, Vestergaard, Fletcher, & Schultz, 2016). This behavioural pattern was accompanied by adaptive PE signals in the midbrain and the ventral striatum. Under 600 mg of sulpiride, both the PE scaling as well as the adaptive PE coding in the midbrain and partially in

the striatum were reduced (Diederen et al., 2017). This suggests that D2 receptors likely play a general role in uncertainty coding across various task modalities and contexts.

Our findings that blocking D2/D3 receptors increases learning rates may seem to be at odds with previous work showing that D2/D3 antagonists reduced performance in other learning tasks and attenuated prediction error signals in the striatum (Jocham et al., 2014; Pessiglione et al., 2006) as well as with studies showing no effect of D2/D3 antagonism on learning rates (Jocham et al., 2011, 2014), even when using similar computational frameworks (Iglesias et al., 2021; Marshall et al., 2016). It is thus important to note that the A1 is a minor allele of the Taq1a polymorphism, meaning that in most other studies participants were likely predominantly A2 homozygotes. We observed a more general effect of D2/D3 receptor antagonism on choice uncertainty that was more prominent in A2 homozygotes and was related to a higher number of mistake trials in that subgroup of participants, although the number of mistakes was not high enough to reduce investment on average. Furthermore, participants could invest on an ordinal 11-point scale, which allowed us to capture smaller belief shifts that might either be missed in learning tasks with categorical choice options or be attributed to a different choice selection policy. For example, the participants in our study also performed a standard probabilistic two-bandit task afterwards , where participants in the A1+ group under sulpiride compared to placebo continued to switch between choice options, which was explained by increased choice stochasticity, parametrized through the soft-max decision temperature (Eisenegger et al., 2014). Further, it is also plausible that the processing of uncertainty in a social context is different than in a non-social context. People might be inherently more motivated to reduce uncertainty about others, so that they can (for instance) classify them more definitely as being a friend or foe (FeldmanHall & Shenhav, 2019b). For example, in one study, stress increased the choice to gamble in a non-social context but decreased the likelihood to invest in one shot-Trust games (FeldmanHall, Raio, Kubota, Seiler, & Phelps, 2015). Furthermore, patients with basolateral amygdala damage show markedly impaired belief updating in a repeated Trust game, but seem to have no trouble learning about non-social rewards through a task matched in difficulty and reward size (Rosenberger et al., 2019). It is therefore plausible that the results we found are specific to the social context and might not translate to learning about non-social cues.

One fundamental distinction that separates risky decision-making under asocial compared to social conditions is an aversion to betrayal (Fehr, 2009). People are less risk-taking in social interactions and might be particularly sensitive to indications of untrustworthy interaction partners (Bohnet, Greig, Herrmann, & Zeckhauser, 2008). Using a similar model to ours, previous work has shown that belief volatility was higher when assessing (morally) bad agents (Siegel et al., 2018), an effect that was present in our data as well, with participants having higher belief volatility when playing against the bad trustee. Although the drug effects on belief volatility were present across both trustees, the effects were stronger for the bad trustee. One reason for this could be that because participants initially expected higher trustworthiness and higher rates of positive back transfers, there was more to learn when playing against the bad trustee and, therefore, more variance across investment behaviour. This asymmetric increase in sensitivity to negative outcomes would also be in line with the notion that D1 and D2 receptors in the striatum contribute to positive and negative outcome processing via the "Go" and "No-Go" pathways, respectively (Frank & O'Reilly, 2006; Frank et al., 2004). According to this circuit model, D2 antagonism mimics the dopaminergic 'dip' that occurs following negative reinforcement and therefore enhances learning to avoid action with a negative outcome. However, contrary to what we observed, this model also predicts that blocking post-synaptic D2 receptors should decrease positive prediction error propagation.

When interpreting our findings within the Go/No-Go framework, it should be noted that in the repeated Trust game in this study, participants have no agency over the valence of the outcome (positive or negative back-transfer), and investments are possible only on an ordinal scale. Similarly, the RW model we used in our study should also be interpreted with this in mind. Mirroring the effects in the HGF model, sulpiride increased learning rates in the RW model for both positive and negative outcomes. In multi-arm bandit tasks or Go/No-Go tasks where the reinforcement learning framework

is often used to explain choice selection, the learning rate reflects the "Law of Effect" whereby actions that lead to positive (negative) outcomes are more (less) likely to get repeated (Thorndike, 1911). A positive outcome following a specific investment choice in the repeated Trust game will lead to higher investment (if possible). A higher learning rate simply reflects the change in the expected response of the trustee. It is therefore related to the degree to which beliefs about trustworthiness change (on average across trials). Generally, the crucial distinction between RW and Bayesian models is that the latter assumes that agents consider the uncertainty of outcomes when updating beliefs. With this, Bayesian models such as the HGF or the Kalman filter can account for phenomena where non-Bayesian reinforcement learning models fail, such as latent inhibition and sensory preconditioning (Gershman, 2015, 2018a). Given the relative increase in the overall and trial-by-trial predictive performance of the HGF model that includes choice uncertainty over the RW, our data support the notion that uncertainty about the outcomes and which actions to take affects choice behaviour in the repeated Trust game.

One important factor that could confound increased belief volatility and learning rates following sulpiride administration is working memory. Previous work has shown that individual differences in working memory capacity contribute to behavioural variability in reinforcement learning tasks (Collins, Brown, Gold, Waltz, & Frank, 2014; Collins & Frank, 2012) whereby decreased memory capacity might lead to a higher salience of more recent outcomes and therefore higher learning rates (Collins, Ciullo, Frank, & Badre, 2017). We also find support for this notion in our data, whereby poorer WM performance was strongly linked to higher belief volatility and higher choice uncertainty. However, despite sulpiride decreasing WM capacity in our cohort, including WM data in the model did not affect inference about sulpiride's effect on belief volatility, nor on choice stochasticity. This increase in choice uncertainty or stochasticity under sulpiride is therefore likely not due to failures in working memory capacity. Instead it could have been due to participants being less motivated to maximise outcomes, and therefore less likely to behave as a rational "homo economicus" (Camerer, 2003). Were that the case in our study, one would expect a different behavioural pattern in the single shot Trust games. Participants under sulpiride should behave less as rational agents and therefore would be less likely to punish betrayals and reward trusting behaviour (Fehr, Fischbacher, & Gächter, 2002). Note that D2 receptors generally do play a role in motivation. For example, optogenetic excitation and inhibition of D2 receptors in the ventral striatum of rats is reported to respectively increase and decrease motivation (Soares-Cunha, Coimbra, David-Pereira, et al., 2016). It is possible that the increased action variability resulted from reduced motivation, or increased noise in belief updating and not in choice selection per se (Findling, Skvortsova, Dromnelle, Palminteri, & Wyart, 2019). What speaks against this interpretation is that the overall performance in the task was not reduced following sulpiride administration for either of the genetic subgroups, suggesting that the investment selection under sulpiride was not random and instead reflected uncertainty about which investment to choose when interacting with the other player.

Indeed, variability in investment selection following sulpiride administration is well in line with what we know about the role of dopamine receptors in action selection. Stimulation of D2 receptors through endogenous dopamine leads to inhibition of the indirect (No-Go) pathway and increases the probability of repeating the same action. Accordingly, blockade of postsynaptic D2 receptors increases the probability of performing competing actions and therefore promotes randomness in action selection (Sridharan, Prashanth, & Chakravarthy, 2006). For example, in macaques, microinfusion of D2 (but not D1) receptor antagonists into the dorsal striatum led to increased choice stochasticity (E. Lee, Seo, Dal Monte, & Averbeck, 2015) and a similar pattern was observed in D2 receptor knockout mice (Kwak et al., 2014). In humans, a recent positron emission tomographic imaging study showed that D2 receptor availability in the striatum correlated with deterministic decision-making strategies represented either through decision temperature within reinforcement learning as well as with policy precision within active inference (Adams et al., 2020).

The key idea of active inference models is to extend the Bayesian generative models of beliefs about the states of the world, to include beliefs about preferred states, therefore casting both action and perception as an inference problem (Friston et al., 2015, 2013). An active inference agent thus prefers actions that minimize the statistical distance (relative entropy) between the distributions of desired and predicted future states. The expected precision of a policy, in the context of our task, controls the confidence with which the participants selected a certain action, which can explain the more variable investment we observed in the sulpiride group. Within this framework, we can interpret the effects of sulpiride in our study as reflecting a more general role of D2 receptors in coding precision of both beliefs and action policies, thus extending previous theoretical and experimental work on the involvement of dopamine in modulating precision in predictive coding schemes (Friston et al., 2012; Nour et al., 2018; Schwartenbeck et al., 2015).

Our findings might be particularly relevant for understanding the effects of antipsychotic medication in patients with psychosis, a disorder characterized by rigid beliefs of persecution, underlined by a profound lack of trust in others (Freeman, 2016; Fuchs, 2015). Previous studies with repeated Trust games showed that patients with psychosis have lower initial trust and find it hard to change their beliefs (Fett et al., 2012; Gromann et al., 2013). Neurocomputational accounts of delusions suggest that hyperactivity of D2 receptors in patients leads to increased precision beliefs that result in rigid convictions held with high confidence (Adams et al., 2013; Sterzer et al., 2018) and a recent paper shows that higher belief instability in patients with schizophrenia predicts responses to psychotherapeutic treatment (Hauke et al., 2022). This suggests that decreasing belief rigidity through D2 antagonism could be an essential contributor to the success of adjunct psychosocial treatment. However, there are profound differences between the effects of repeated use of antipsychotics in patients and acute D2 antagonism in healthy participants. For example, rodent studies show that although in healthy animals D2 receptor antagonism increases the activity of midbrain dopaminergic cells, this pattern is reversed in an established animal model of schizophrenia (Valenti, Cifelli, Gill, & Grace, 2011). Furthermore, despite rapid receptor blockade of D2 receptor antagonists, the inhibition of excessive dopaminergic signalling proposed to underly the therapeutic effects develops only after weeks of treatment (Grace, Bunney, Moore, & Todd, 1997; Kapur, Agid, Mizrahi, & Li, 2006). Our data also suggest that the therapeutic effect could be larger in patients that are A1 allele carriers of the Taq1a polymorphism. Yet, there is no evidence for this (J.-P. Zhang, Lencz, & Malhotra, 2010), despite higher dopamine synthesis being the most likely biomarker of psychotic symptoms (Fusar-Poli & Meyer-Lindenberg, 2013; Howes et al., 2012) and a predictor of response to antipsychotic treatment (Jauhar et al., 2019). Translating our findings to clinical practice will require more work with targeted patient populations.

Several important limitations should be kept in mind when considering the generalizability of the findings in this study. First, the sample was limited to male participants. This restriction was initially motivated by the notion that including female participants would require more than doubling the sample size due to increased variance of dopamine availability across the ovarian cycle. However, recent work shows no support for this (Petersen et al., 2021). Given that there are important sex differences in responses to antipsychotics, both in terms of efficacy and side-effect profiles (Hoekstra et al., 2021), future work should prioritize studies in females. Second, despite our hypothesis-driven approach, the sample size of the genetic subgroups was small. The drug-gene interactions we report should be interpreted with this in mind. We note however that we did find a main effect of sulpiride on increased investment change and on belief volatility in a sample size comparable to other pharmacological studies with a between-subject design (Martins, Mehta, & Prata, 2017).

In conclusion, we show that blocking D2 dopamine receptors increases the flexibility of beliefs when learning about others. This finding importantly contributes to our understanding of how the brain infers the attitudes of other people. By mapping out the connection between alterations in the dopaminergic system with specific computational substrates this study not only contributes to the

advancement of our knowledge of how the brain performs inference, but also to our understanding of when it fails to appropriately do so.

Methods

Participants

The study was performed in accordance with the Declaration of Helsinki and approved by the National Research Ethics Committee of Hertfordshire (11/EE/0111). Data were collected from 78 male participants, aged between 19 and 44 years (mean= 32.1), recruited from a large panel of participants, that were genotyped and screened for mental and physical health (Cambridge BioResource). Only participants with no history of neurological or psychiatric disorder were included in the study. Participants were stratified based on the Taq1A genotype into two groups: participants carrying at least one A1 allele (A1+), and A2 allele homozygotes (A1-).

Procedure

After arrival participants underwent another psychiatric screening and an alcohol test to exclude alcohol consummation on the study day. After an assessment of general intelligence (National Adult Reading Test) participants signed an informed consent before they were administered a single oral dose of either 800 mg of sulpiride or placebo in a randomized, double-blind fashion. We used the parallel group design, because complex behavioural tasks (like the Trust game) have practice (repetition) effects that can confound the results of within group pharmacological experiments. Sulpiride maximal plasma concentration is expected to peak after 3 h, with a plasma half-life of about 12 h (Bressolle, Bres, & Fauré-Jeantis, 1992; Wiesel, Alfredsson, Ehrnebo, & Sedvall, 1982). Before behavioural testing participants waited for 3 h in a quiet room, where they were allowed to read a newspaper. To monitor the effects of the pharmacological manipulation, blood pressure and heart rate and mood and drug effects were assessed prior to drug administration and after the 3h waiting period. Similarly, blood samples to determine the serum levels were taken at both time points. After the blood draws (at around 3h 20') the behavioural testing started with the social interaction tasks presented here, which included a repeated Trust game, and positive and negative reciprocity tasks, and were followed by a working memory task and an instrumental learning task, both published elsewhere (Eisenegger et al., 2014; Naef et al., 2017). Two participants were excluded from the analysis: one felt uncomfortable in the room, and one did not sufficiently understand the instructions of the social interaction tasks. This led to the following group distributions: 17 A1 allele carriers received placebo, and 21 received sulpiride, and 21 A2 homozygotes received placebo, and 17 received sulpiride. Participants were matched across the four groups for age, body mass index, general and verbal intelligence (Table 1, all $p>0.30$). Participants received a monetary compensation of £50 plus the extra money earned in the behavioural tasks.

Table 1. Demographic information

Genotype	Treatment	N	IQ	sd	Verbal IQ	sd	BMI	sd
A1+	Placebo	17	120.2	5.333	120.335	5.913	26.098	3.125
A1+	Sulpiride	21	120.21	7.334	120.355	8.133	26.595	5.74
A1-	Placebo	21	120.05	6.089	120.17	6.76	24.604	4.415
A1-	Sulpiride	17	117.659	7.494	117.518	8.318	25.064	5.393

sd, standard deviation

Sulpiride Serum Concentration Measurements

The level of serum sulpiride was determined by high-performance liquid chromatography. This method utilizes fluorescence endpoint detection with prior solvent extraction. The excitation and emission wavelengths were 300 and 360 nm, respectively. Both intra- and inter-assay coefficients of variation (CVs) were 10% and the limit of detection was 5–10 ng/ml.

Prolactin level assessment

The prolactin level was measured using a commercial immunoradiometric assay (MP Biomedicals, Santa Ana, CA, USA), 3h after capsule ingestion. Prolactin levels were expected to increase with blocking postsynaptic D2 receptors (Wiesel et al., 1982). The intra- and inter-assay coefficients of variation were 4.2% and 8.2%, respectively, and the limit of detection was 0.5 ng ml−1 (Eisenegger et al., 2014). We found that sulpiride administration significantly increased blood plasma prolactin levels (Δ = 33.1 mg/ml, $p < 0.001$), and this increase was significantly higher ($p < 0.001$) than the changes in the placebo group (Δ = -0.91 mg/ml, Mann Whitney test for differences $p < 0.001$). Data for three participants were excluded due to blood contamination.

Side-effects and mood assessments

Side effects were assessed with a neurovegetative list (C. R. Rush, Stoops, Hays, Glaser, & Hays, 2003), 3 h after drug intake. Mood was assessed with a visual analogue scale at baseline and 3 h after drug intake. Items in the visual analogue scales (VAS) were alert/drowsy, calm/excited, strong/feeble, muzzy/clear-headed, well coordinated/clumsy, lethargic/energetic, contented–discontented, troubled–tranquil, mentally slow/quick-witted, tense/relaxed, attentive/dreamy, incompetent/proficient, happy/sad, antagonistic/amicable, interested/bored and withdrawn/gregarious. The factors "alertness", "contentedness", and "calmness" were calculated from these items (Eisenegger et al., 2010). Data from one participant were excluded due to technical issues. We found no credible evidence of drug effects on mood, heart rate, blood pressure or self-reported side-effects (for details see Supplementary Material).

Repeated Trust game

In the Trust game (Berg et al., 1995) an investor (Player A) decides on how much money they want to transfer to the other player, called the trustee (Player B). The trustee receives the investment that is however tripled by the experimenter and decides on how to split the acquired sum. We used a multi-round version of the task (King-Casas, 2005), where the interchange between the investor and the trustee repeated across 25 trials. In the beginning of each trial both players were endowed with 10 points, to avoid investments motivated by inequality aversion. Each point converted to two pence at the end of the experiment. The participants could invest points on a scale from 0 to 10 and the trustee could respond in a binary fashion, by either equalizing the payoff, or defecting by keeping all the points in the trial for themselves. Participants played as investors against two pre-programmed trustees: one defected in 7 out of 25 trials (the good agent) and the other defected in 18 out of 25 trials (the bad agent). The feedback was pseudorandomized separately for each participant and was interleaved whereby only two consecutive trials with the same trustee were allowed. To increase ecological validity, the participants were led to believe that they play against two actual people that have already given their answers in advance several weeks before the testing, and that their decision will impact the payoff of these participants. All paradigms were programmed in Visual Basic.

Positive and negative reciprocity games

In the positive reciprocity game, the two players need to distribute 800 points. First, player A is offered a distribution whereby they get 800 points and player B gets 0 points. They can decide to either keep all the points or delegate the decision on how to divide the points to player B. If the decision was to delegate, to player B can decide on any point distribution between the two players. Participants in our study played as player B sequentially against 7 different people playing as player A. The negative reciprocity game is like a Trust game in which defecting behaviour of the trustee can be punished by the investor. Both players are first endowed with 10 points. Player A then decides to either transfer his endowment (all the 10 points) all transfer nothing. The transfer of player A is quadrupled by the experimenter. Player B can then decide to either keep everything to themselves or to equalize the payoff. Following the decision of player B, both players get endowed with another 20 points and player A can spend each of these 20 points to penalize player B's outcome, whereby each penalty point of

player A spent this way deducts three times as many points from player B's outcome. Participants in our study played as player A against 7 different people playing as player B. The actions of people playing player B were pre-programmed so that 5 out of 7 defected.

In both games the participants were told that the players have given their answers already days before the testing. Each point converts to 0.2 pence for the positive reciprocity game and 4 pence for the negative reciprocity game.

Spatial working memory task

In the Spatial working memory task the participants were required to search through a spatial array of coloured squared boxes for a hidden 'token' (Naef et al., 2017) using a tablet. Participants have to touch the box to open it in order to reveal whether the token is in the box or not. When a token is found, the search starts again, whereby no token will be hidden in a box twice in the same trial. The performance measure is the number of search errors defined as errors committed when participants choose a box that has already had a token in that trial. There are 3, 4, 6, 8, 10 or 12 boxes in each trial, making the task progressively harder. The three boxes search were used as practice trials and a successful completion of them was a requirement for progressing onto the main test. The other difficulties each appeared three times. One subjects did not complete the working memory task and was removed from the analysis.

Behavioural analysis

Behavioural analysis was done with Bayesian multilevel (generalized) linear regression (McElreath, 2018), fitted with the brms package in R (Bürkner, 2017) through RStudio. All models were run with 4 chains, 3000 iterations each with 800 warmup. The quality of chain convergence was inspected visually based on trace plots of main fixed effects, and a threshold on Gelman-Rubin $\hat{R}$ Statistic for each parameter was set to 1.01(Gelman & Rubin, 1992). Throughout the behavioural analysis we z-scored the dependent variables (across the whole group), coded the Treatment variable as 0 (placebo) and 1 (sulpiride), Back-transfer as 1 (equalize) and -1 (betray), Genotype as 0 (A1-) and 1 (A1+) and centred the trustee variable (0.5 for good, and -0.5 for bad trustee). All random effects were modelled as a multivariate normal distribution, thereby evaluating the correlation between the effects as well as pooling information across the effects. Priors used are depicted in Table 2. The effect sizes where calculated by dividing the regression coefficients with the square root of summed variances of the residuals and of all random effects (Nalborczyk et al., 2019). All models were redone also in the lme4 package (Bates, Mächler, Bolker, & Walker, 2014) or nlme package (Pinheiro, Bates, DebRoy, Sarkar, & Team, 2007), and the results of those models are reported in the supplementary material.

Table 2: Prior distributions for the analysis of investment changes

Standard Deviations	$\boldsymbol{\sigma \sim HalfCauchy(0,2)}$
Regression Coefficients	$\beta \sim N(0,3)$
Prior for the correlation matrix	$R \sim LKJcorr(2)$

Analysing investment behaviour

All model summary tables are in the supplementary material. The effect of sulpiride on absolute change from one trial to the next was evaluated with a model predicting effects on absolute change of sulpiride and trials with random intercepts for each participant. We also rerun the model including a participant-level slope for the trial and found that it does not affect inference about the main effect, but does increase the uncertainty around the interaction term. Next, the Genotype and Trustee as group level predictors were included as well as a random slope for the Trustee for each participant. Since the dependent variable is bounded at 0, the same analysis was done again with the dependent variable shifted by 1 and log transformed. This did not affect the conclusion of the model.

To analyse relative changes in investment the z-scored relative change from one trial to the next was predicted from the variable for Back-transfer (coded as -1 and 1), Treatment, Genotype, Trustee as well as their interactions, with a participant-level random intercept and slope for the Trustee.

The reciprocal and mistake trials were analysed with a multilevel logistic regression model including predictors Treatment, Genotype, Trustee, and their interaction, again with a random intercept and slope for the effects of the Trustee for each participant.

To analyse average investments, we used a multilevel ordinal-logistic regression model, with Treatment, Genotype, Trustee, Trial and their interaction, with a random intercept and slope for the effects of the Trustee for each participant.

Models analysing the single-round reciprocity tasks predicted Punishment (negative reciprocity) and Back transfer (positive reciprocity) from Treatment, Genotype and their interaction, including a random intercept per participant.

Computational modelling

We first defined a generative model of the evolution of beliefs about the other players' trustworthiness as a Gaussian random walk. The belief volatility parameter ω describes the degree to which these beliefs can change from one trial to the next. We then used the Hierarchical Gaussian Filter (HGF) to invert this model (Mathys et al., 2011).

Generative Model

The generative model describes the evolution of beliefs about the other person's trustworthiness as a Gaussian random walk with a step size of $\exp(\omega)$. In particular, at trial t the belief on the other player's trustworthiness is defined as

$$x^{(t)} \sim N\left(x^{(t-1)}, \mathrm{e}^{\boldsymbol{\omega}}\right), \tag{2}$$

where ω is a participant level parameter. The mapping from the trustworthiness beliefs to the probability of a positive Back transfer (BT) occurs through a sigmoid transform $s(.)$. So, at trial t we define:

$$BT^{(t)} = Bernoulli(s\left(x^{(t)}\right), \tag{3}$$

$$(x) = 1/(1 + e^{x}). \tag{4}$$

Model inversion and update equations

To define the inferred participant level belief trajectories the generative model is inverted using Hierarchical Gaussian Filtering (Mathys et al., 2011, 2014). The HGF approximates full Bayesian inference using variational Bayes to derive at trial level update equations that resemble those of a Kalman filter (Gershman, 2015; Kakade & Dayan, 2002). In particular, the weights (learning rates) on the PEs are determined by the precision of prior beliefs as well as the uncertainty about the outcome. The HGF provides inferred posterior distributions of participants' belief trajectories as Gaussians through the mean $\mu^{(t)}$ and variance ${\sigma^{(t)}}^2$ or its inverse, the precision $\pi^{(t)}$ in the update equations for both time series:

$$PE = BT^{(t)} - s\left(\kappa\mu^{(t)}\right), \tag{5}$$

$$\mu^{(t+1)} = \mu^{(t)} + \psi_t PE, \tag{6}$$

$$\pi^{(t+1)} = 1/(\frac{1}{\pi^{(t)}} + \mathrm{e}^{\omega}), \tag{7}$$

$$\hat{\pi}^{(t)} = \pi^{(t+1)} + \left(s\left(\mu^{(t)}\right) * \left(1 - s\left(\mu^{(t)}\right)\right)\right), \tag{8}$$

$$\psi_t = 1/\hat{\pi}^{(t+1)}, \tag{9}$$

where PE is the prediction error, ψ_t is the precision weight (learning rate) that is determined by the expected precision of prediction ($\hat{\pi}^{(t)}$), and ω and κ are free parameters in the model. This is a so-called recognition or perceptual model and describes our beliefs about the belief of the participants. To map the beliefs about the trustworthiness of the other person ($\mu^{(t)}$)on to the probability of feedback we used two versions of non-linear mapping from beliefs to probabilities:

$$\mu_p^{(t)} = pw(s(\kappa\mu^{(t)}), \gamma), \tag{10}$$

$$pw(x, \gamma) = \frac{x^\gamma}{x^\gamma + (1-x)^\gamma}, \tag{11}$$

where $pw(.)$ is a probability weighting function on the unit interval, with another free parameter γ that determines the skew of the function. We estimated either γ or κ and fixed the other parameter.

We then mapped the probability of feedback to behaviour of the participant with a response model is defined through a likelihood function. Because investments occur on an ordinal scale we used the ordered logistic link function(Bürkner & Vuorre, 2019):

$$P(I^{(t)} = k) \sim \text{Ordered} - \text{Logit}(\eta\mu_p^{(t)} C), \tag{12}$$

Where C is a vector of intercepts and η is the noise parameter that (similarly to the inverse temperature in the softmax equation) determines to what degree belief about the other person's trustworthiness determines investment behaviour. The ordered-logit estimates 10 intercepts, that determine the mapping from the linear term to the ordinal investments. In an ideal case, we would estimate all 10 intercepts for each subject, which was not feasible with our data. We therefore estimate the 10 intercepts for all subjects, add one subject-level intercept in the linear term and assist the model in accounting for the various investment distributions with the non-linear mapping from beliefs to probabilities enabled either by κ or γ. In the winning HGF model, κ was fixed to 1 and γ was estimated. In the second HGF model, γ was fixed to 1 and κ was estimated as a free parameter.

We also compared both HGF models to a simple Rescorla-Wagner model (Rescorla & Wagner, 1972) with separate learning rates for positive and negative feedback. Given a learning rate α, and an action value Q of a chosen action on trial t, we defined the updating equation as:

$$Q^{(t+1)} = Q^{(t)} + \alpha PE, \tag{13}$$

$$PE = BT^{(t)} - Q^{(t)}, \tag{14}$$

while using the same likelihood function as for the HGF models. We estimated a different α for positive and negative outcomes.

Parameter Estimation

The model parameters were estimated in one hierarchical Bayesian model. This approach reduces overfitting (McElreath, 2018), pools information across different levels (drug groups, and participants) and allows us to estimate both participant and group level parameters in one inferential step. Meaning, we estimate the effects of our drug manipulation on all relevant computational parameters in one model, while at the same time, leading to more stable parameter estimates (Ahn, Haines, & Zhang, 2017). Models were implemented in Stan (Carpenter et al., 2017) using R as the programming language and RStudio as the integrated development environment for R. Each candidate model with four independent chains and 3000 iterations (800 warm-up). Convergence of sampling chains was estimated through the Gelman-Rubin $\hat{R}$ statistic (Gelman & Rubin, 1992), whereby we considered $\hat{R}$ values smaller than or equal to 1.01 as acceptable.

The intercepts from the response model, $c_k, k = 1, \dots, 10$, were estimated on the group level. This determined a general mapping from the probability to Investment. The participant level parameters (ω, $\Delta\omega$, γ, η and μ_0) were modelled as a multivariate Gaussian distribution:

$$\begin{pmatrix} \omega \\ \Delta\omega \\ \gamma' \\ \eta \\ \mu_0 \end{pmatrix} \sim \text{MVNormal}(\boldsymbol{\mu}, \boldsymbol{S}), \tag{15}$$

Where S is the covariance matrix and μ is the vector of means. The $\Delta\omega$ parameter denotes the modelled difference in ω between the good and the bad Trustee. The matrix S was factored into a diagonal matrix with standard deviations and the correlation matrix R (Bürkner, 2017; McElreath, 2018). The prime denotes the parameters in estimation space, whereby γ was estimated in log space, due to it being lower bound by 0. The vector μ included all group level regression coefficients for the drug, genotype, and their interaction. The priors for group-level means for non-transformed parameters were weakly informative, for γ, estimated in log-space, the prior was $N(0,0.5)$, the prior for group-level standard deviations were more regularizing, with $\sigma \sim HalfNormal(0,0.2)$, and the prior for the correlation matrix was $R \sim LKJcorr(2)$. The prior for group level η was set to something above 0, because chains that sampled from areas too close to 0 usually got stuck in that area. An overview of priors of parameters across the three models can be seen in Table 3.

Table 3. Priors for the three computational models

	HGF with γ	HGF (with κ)	RW
Hyperpriors (random effects)	$\kappa - fixed$ $\mu_\omega \sim N(-2,1)$ $\mu_{\Delta\omega} \sim N(0,1)$ $\mu_\eta \sim N(10,1)$ $\mu_\gamma \sim N(0.0.5)$ $\mu_{\mu_0} \sim N(0,1)$ $\sigma \sim HN(0,0.2)$ $R \sim LKJcorr(2)$	$\gamma - fixed$ $\mu_\omega \sim N(-2,1)$ $\mu_{\Delta\omega} \sim N(0,1)$ $\mu_\eta \sim N(10,1)$ $\mu_\kappa \sim N(0.1)$ $\mu_{\mu_0} \sim N(0,1)$ $\sigma \sim HN(0,0.2)$ $R \sim LKJcorr(2)$	$\mu_\alpha \sim N(0,1)$ $\mu_{\Delta\alpha} \sim N(0,1)$ $\mu_\eta \sim N(10,1)$ $\mu_{\mu_0} \sim N(0,1)$ $\sigma \sim HN(0,0.2)$ $R \sim LKJcorr(2)$
Priors of group level (fixed) effects	$\beta_* \sim N(0,1)$ $c_k \sim N(0,5)$		

To control to what degree working memory affects inference about the effects of drug and genotype we additionally estimated another model that included the number of search errors (z-scored) as a covariate on the group-level effect. To calculate the subject-level residuals after accounting for spatial working memory data, we ran another model that did not include drug data, meaning the only group level parameter affecting the group level means was the regression coefficient for te spatial working memory.

Model Validation and Comparison

For parameter recovery 5 parameter sets were drawn from each participant's mean and standard deviation and used to simulate data. Simulated data were then estimated with the same model and the re-estimated parameters were correlated with the simulated ones. Further, posterior distributions of parameters were used to simulate data and check whether the crucial aspects of behaviour are captured by the model. A trial based Leave-One-Out Information Criterion (LOOIC) was used to compare the three models (Vehtari, Gelman, & Gabry, 2017) using the loo package in R. The LOOIC approximate out-of-sample predictive accuracy of each trial, with lower LOOIC scores indicating better prediction accuracy out of sample. Additionally, the performance of the models was compared with the with the

"loo_compare" which return the difference (and the standard error) of each model to the best performing model in terms of expected predictive accuracy on a log scale (ELPD).

Data availability

All data of the experiment is available online (https://doi.org/10.5281/zenodo.7779029).

Code availability

The analysis scripts are available online (https://github.com/nacemikus/belief-volatility-da-trustgame.git).

Acknowledgements

This paper is dedicated to the memory of Christoph Eisenegger. This research work was funded by the Vienna Science and Technology Fund (WWTF) with a grant (VRG13-007) awarded to Christoph Eisenegger and Claus Lamm. Christoph Eisenegger was also supported by the Swiss National Science Foundation (PA00P1_134135). We are grateful for the participation of all NIHR Cambridge BioResource (CBR) participants and thank the Cambridge BioResource staff for their help with participant recruitment. We also thank members of the Cambridge BioResource SAB and Management Committee for their support given to our study and the National Institute for Health Research Cambridge Biomedical Research Centre for funding. The manuscript has significantly improved from the constructive and insightful remarks of the three reviewers. Open access funding was provided by University of Vienna and Durham University.

Author Contributions Statement

Study Design: CE, MN, UM, LC, & TWR; Computational modelling: NM, & CM; Software: MN, & NM; Data analysis: NM, & CM; Data Collection: CE, MN, & UM; Medical cover: UM; Resources: TWR, CE & CL; Data curation: CE, NM & MN; Writing – Original draft: NM; Writing – reviewing & editing: all authors; Visualization: NM; Supervision: MN, TWR, CM, & CL; Funding acquisition: CE, TWR & CL.

Corresponding authors:

Correspondence to Nace Mikus, Claus Lamm, or Michael Naef.

Competing interests

LC has received royalties from Cambridge Cognition Ltd. relating to neurocognitive testing. UM discloses consultancy for Janssen-Cilag, Lilly, Heptares and Shire, and educational funding from AstraZeneca, Bristol-Myers Squibb, Janssen-Cilag, Lilly, Lundbeck and Pharmacia-Upjohn. TWR discloses consultancy with Cambridge Cognition Ltd and a research grant with Shionogi Inc. The remaining authors declare no competing interests.

Supplementary Material

Supplementary Tables

Abs_change~Treatment*Trial + (1|ID)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-0.15	0.076	-0.295	-0.003
TreatmentSulpiride	0.309	0.106	0.108	0.518
Trial	-0.015	0.003	-0.021	-0.009
TreatmentSulpride:Trial	0.007	0.004	-0.001	0.015
Abs_change~Treatment*Trial + (Trial|ID)				
Intercept	-0.114	0.069	-0.246	0.025
Treatmentsulpiride	0.234	0.097	0.043	0.427
Trial_c	-0.012	0.005	-0.022	-0.001
Treatmentsulpiride:Trial_c	0.007	0.007	-0.008	0.022

Supplementary Table 2. Results of the model Bayesian mixed model predicting absolute changes in investment from one trial to the next including drug treatment and trial as predictors. Similar results are achieved when the slope for the trial is allowed to vary per participant. Treatment (coded as 1 sulpiride and 0 placebo), trial was used as a continuous variable and centred. Refer to the code about the effect in native space and the calculation of effect sizes.

Abs_change ~Treatment*Trial + (1|ID)	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.113	0.066	3570	-1.715	0.086
TreatmentSulpiride	0.235	0.093	74	2.515	0.014
Trial	-0.012	0.003	3570	-3.87	<10e3
TreatmentSulpiride:Trial	0.007	0.004	3570	1.626	0.104

Supplementary Table 3. The output of the non-Bayesian absolute change model. T-values all two-sided, p values uncorrected.

log(1+Abs_change) ~ Treatment*Trial + (1|ID)	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.149	0.073	3570	-2.045	0.041
Treatmentsulpiride	0.309	0.103	74	3.004	0.004
Trial	-0.015	0.003	3570	-5.084	<10e3
Treatmentsulpiride:Trial	0.007	0.004	3570	1.715	0.086

Supplementary Table 4. The non-Bayesian absolute change model reran again with a transformed outcome variable. Because the absolute change variable is positive, we use a log transform of a translation. T-values all two-sided, p values uncorrected.

Abs_change ~ Treatment*Genotype *Trustee *Trial + (Trustee |ID)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-0.109	0.069	-0.25	0.024
Treatmentsulpride	0.234	0.099	0.044	0.431
Trial	-0.013	0.003	-0.018	-0.007
Trustee	0.003	0.07	-0.133	0.143
Genotype	0.079	0.138	-0.195	0.348
Treatmentsulpride:Trial	0.007	0.004	-0.001	0.016
Treatmentsulpride:Trustee	-0.145	0.1	-0.339	0.046
Trial:Trustee	0.015	0.006	0.003	0.027
Treatmentsulpride:Genotype	-0.111	0.198	-0.494	0.278
Trial:Genotype	-0.011	0.006	-0.023	0
Trustee:Genotype	-0.002	0.137	-0.272	0.274
Treatmentsulpride:Trial:Trustee	-0.007	0.009	-0.024	0.01
Treatmentsulpride:Trial:Genotype	0.021	0.009	0.004	0.038
Treatmentsulpride:Trustee:Genotype	-0.051	0.194	-0.433	0.332
Trial:Trustee:Genotype	-0.016	0.012	-0.04	0.008
Treatmentsulpride:Trial:Trustee:Genotype	0.042	0.017	0.009	0.076

Supplementary Table 5: Bayesian mixed model predicting absolute changes in investment from one trial to the next including drug treatment, trial, genotype and trustee as predictors. Treatment coded as 1 sulpiride and 0 placebo, trial used as a continuous variable. Variables Trial, Genotype (0.5 A1+, -0.5 A1-) and Trustee (0.5 good, -0.5 bad) were centred.

Change ~Treatment*Genotype*Backtransfer*Trustee+ (Trustee |ID)	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.017	0.038	3560	0.436	0.663
Treatmentsulpiride	-0.002	0.051	72	-0.034	0.973
Backtransfer	0.204	0.038	3560	5.337	<10e3
GenotypeA1-	-0.033	0.051	72	-0.65	0.518
Trustee	-0.105	0.076	3560	-1.375	0.169
Treatmentsulpiride:Backtransfer	0.114	0.051	3560	2.218	0.027
Treatmentsulpiride:GenotypeA1-	0.03	0.072	72	0.42	0.676
Backtransfer:GenotypeA1-	0	0.051	3560	-0.004	0.997
Treatmentsulpiride:Trustee	-0.078	0.102	3560	-0.759	0.448
Backtransfer:Trustee	-0.027	0.076	3560	-0.352	0.725
GenotypeA1-:Trustee	-0.012	0.103	3560	-0.112	0.911
Treatmentsulpiride:Backtransfer:GenotypeA1-	-0.171	0.072	3560	-2.363	0.018
Treatmentsulpiride:Backtransfer:Trustee	-0.084	0.102	3560	-0.819	0.413
Treatmentsulpiride:GenotypeA1-:Trustee	0.145	0.145	3560	1.003	0.316
Backtransfer:GenotypeA1-:Trustee	0.094	0.103	3560	0.917	0.359
Treatmentsulpiride:Backtransfer:GenotypeA1-:Trustee	-0.002	0.145	3560	-0.013	0.99

Supplementary Table 6 A frequentist mixed model predicting changes in investment from one trial to the next including drug treatment, trial, genotype, back-transfer and trustee as predictors. Treatment coded as 1 sulpiride and 0 placebo, trial used as a continuous variable. Variables Trial, Genotype (0.5 A1+, -0.5 A1-), Backtransfer (-1 betray, 1 equalize) and Trustee (0.5 good, -0.5 bad) were centred. T-values all two-sided, p values uncorrected.

ReciprocalTrial ~Treatment*Genotype*Trustee_c + (Trustee |ID), family = binomial	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.634	0.133	0.377	0.894
Treatmentsulpiride	0.399	0.179	0.045	0.755
GenotypeA1-	-0.015	0.179	-0.37	0.335
Trustee	-0.027	0.23	-0.474	0.427
Treatmentsulpiride:GenotypeA1-	-0.223	0.256	-0.727	0.276
Treatmentsulpiride:Trustee	-0.023	0.315	-0.637	0.591
GenotypeA1-:Trustee	0.375	0.307	-0.219	0.975
Treatmentsulpiride:GenotypeA1-:Trustee	-0.103	0.439	-0.965	0.76

Supplementary Table 7. Bayesian mixed logistic model predicting Reciprocal trials from drug treatment, genotype and trustee as predictors.

ReciprocalTrial ~Treatment*Genotype*Trustee_c + (Trustee_c|ID), family = binomial	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.604	0.126	4.804	<10e3
Treatmentsulpiride	0.392	0.171	2.296	0.022
GenotypeA1-	-0.018	0.169	-0.106	0.916
Trustee	-0.049	0.22	-0.222	0.824
Treatmentsulpiride:GenotypeA1-	-0.217	0.24	-0.903	0.366
Treatmentsulpiride:Trustee	-0.009	0.299	-0.03	0.976
GenotypeA1-:Trustee	0.38	0.295	1.289	0.197
Treatmentsulpiride:GenotypeA1-:Trustee	-0.108	0.421	-0.257	0.797

Supplementary Table 8. Frequentist model mixed logistic model predicting Reciprocal trials from drug treatment, genotype and trustee as predictors. P values uncorrected.

ReciprocalTrial ~log(Serum)*Genotype*Trustee_c + (Trustee |ID), family = binomial, data = Sulpiride group only	Estimate	Est.Error	Q2.5	Q97.5
Intercept	1.012	0.116	0.783	1.236
logserum_s	0.186	0.113	-0.04	0.41
GenotypeA1-	-0.235	0.17	-0.568	0.103
Trustee	-0.058	0.175	-0.404	0.282
logserum_s:GenotypeA1-	-0.222	0.175	-0.565	0.126
logserum_s:Trustee	0.325	0.169	-0.003	0.661
GenotypeA1-:Trustee	0.263	0.256	-0.241	0.774
logserum_s:GenotypeA1-:Trustee	-0.429	0.256	-0.935	0.074

Supplementary Table 9. Bayesian mixed logistic model predicting Reciprocal trials Sulpiride group only, from serum levels (log-scaled), genotype and trustee as predictors.

MistakeTrial ~Logserum*Genotype*Trustee_c + (Trustee |ID), family = binomial, data = Sulpiride group only	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-2.068	0.248	-2.589	-1.607
logserum_s	-0.331	0.242	-0.818	0.143
GenotypeA1-	0.427	0.361	-0.314	1.136
Trustee	-0.156	0.233	-0.622	0.300
logserum_s:GenotypeA1-	0.941	0.356	0.25	1.663
logserum_s:Trustee	-0.57	0.231	-1.048	-0.127
GenotypeA1-:Trustee	-0.191	0.314	-0.812	0.428
logserum_s:GenotypeA1-:Trustee	0.595	0.315	-0.01	1.227

Supplementary Table 10. Bayesian mixed logistic model predicting Mistake trials Sulpiride group only, from serum levels (log-scaled), genotype and trustee as predictors.

Punishment ~ Treatment*Genotype + (1|ID)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	6.16	1.177	3.893	8.447
Treatmentsulpride	1.776	1.446	-1.097	4.619
genotypeA1-	-0.764	1.435	-3.585	2.086
Treatmentsulpride:genotypeA1-	-0.476	1.83	-4.091	3.077

Supplementary Table 11. Negative reciprocity task. Bayesian mixed model predicting Punishment from treatment, genotype, and trustee as predictors.

Back-transfer ~ Treatment*Genotype + (1|ID)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	380.565	14.042	349.795	403.707
Treatmentsulpride	-0.016	2.923	-5.911	5.634
genotypeA1-	-0.219	2.978	-5.983	5.698
Treatmentsulpride:genotypeA1-	-0.095	3	-5.999	5.792

Supplementary Table 12. Positive reciprocity task. Bayesian mixed model predicting Back-transfer from treatment, genotype and trustee as predictors.

PrecisionWeights_s ~ logserum_s * Genotype + (1 | ID)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.489	0.137	0.219	0.762
logserum_s	0.286	0.126	0.037	0.539
GenotypeA1-	-1.118	0.2	-1.507	-0.721
logserum_s:GenotypeA1-	-0.218	0.193	-0.595	0.158

Supplementary Table 13. Bayesian mixed model predicting PrecisionWeights from log serum levels and genotype in the Sulpiride group only.

prec_weights_s ~ logserum_s * Genotype + (1 | ID)	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.511	0.13	1862	3.94	<10e3
logserum_s	0.293	0.126	34	2.317	0.027
GenotypeA1-	-1.165	0.194	34	-6.005	<10e3
logserum_s:GenotypeA1-	-0.234	0.196	34	-1.196	0.24

Supplementary Table 14. Frequentist mixed model predicting precision weights from log serum levels and genotype in the Sulpiride group only. T-values all two-sided, p values uncorrected.

Points Earned ~ Treatment*Genotype*Trustee	Estimate	Est.Error	Q2.5	Q97.5
Intercept	324.303	7.176	309.962	338.305
TrusteeBad	-120.505	10.308	-140.018	-100.036
Treatmentsulpiride	7.333	9.678	-12.076	26.262
GenotypeA1-	6.415	9.543	-11.988	25.366
TrusteeBad:Treatmentsulpiride	-1.634	14.08	-29.781	26.353
TrusteeBad:GenotypeA1-	-12.337	13.671	-40.087	13.72
Treatmentsulpiride:GenotypeA1-	-8.555	13.507	-34.12	19.582
TrusteeBad:Treatmentsulpiride:GenotypeA1-	17.192	19.632	-21.076	55.809

Supplementary Table 14. Bayesian linear model predicting overall points earned in the repeated trust game from Genotype, Treatment and Trustee.

Side Effects

side effects	time point	N	Plac.	Sulp.	Sign. (*p*)
Heart rate	base	76	69.2	67.5	0.807
	3 h	76	63.8	64.9	0.596
		76	-5.4	-2.6	0.666
Blood pressure systolic [mm hg]	base	76	132.2	132.8	0.783
	3 h	76	128.1	127.5	0.975
		76	-4.1	-5.4	0.621
Blood pressure diastolic [mm hg]	base	76	76.1	76.9	0.856
	3 h	76	72.0	70.9	0.629
		76	-4.1	-6.0	0.240
VAS: alertness (mean)	base	76	22.6	23.4	0.880
	3 h	75	28.5	28.8	0.945
		75	5.9	5.7	0.719
VAS: contentedness (mean)	base	76	18.7	19.6	0.767
	3 h	75	20.6	22.1	0.660
		75	2.0	3.0	0.304
VAS: calmness (mean)	base	76	22.6	24.9	0.659
	3 h	75	23.0	23.4	0.812
		75	0.4	-0.8	0.890
NVL: any effect	3h	75	-31.5	-38.3	0.743
NVL: bad effects	3h	75	-42.4	-43.1	0.439
NVL: good effects	3h	75	-40.2	-40.7	0.570
NVL: high	3h	75	-43.3	-41.6	0.204
NVL: rush	3h	75	-41.3	-43.6	0.270
NVL: like drug	3h	75	-16.8	-14.6	0.417
NVL: stimulated	3h	75	-39.4	-36.7	0.100
NVL: performance impaired	3h	75	-38.2	-35.6	0.152
NVL: performance improved	3h	75	-38.1	-41.1	0.629
NVL: willing to take again	3h	75	8.8	2.9	0.656
NVL: willing to pay for	3h	75	-39.3	-40.8	0.402
NVL: active-alert-energetic	3h	75	-35.1	-37.9	0.844
NVL: shaky/jittery	3h	75	-42.0	-36.0	0.337
NVL: euphoric	3h	75	-43.3	-38.7	0.158
NVL: irregular or racing heart	3h	75	-45.8	-44.7	0.153
NVL: talkative-friendly	3h	75	-39.1	-31.9	0.049
NVL: nauseated, queasy or sick to stomach	3h	75	-46.4	-46.5	0.393
NVL: nervous or anxious	3h	75	-42.9	-45.1	0.734
NVL: restless	3h	75	-36.2	-30.8	0.085
NVL: sluggish-lazy-fatigued	3h	75	-25.7	-23.6	0.487

Supplementary Table 15. Physiological and self-reported side effects following drug. (Notes. Base = baseline; 3h = 3 hours after drug loading; δ = difference between the value 3 hours after drug loading and the baseline; N = number of observations; Plac. = Placebo group; Sulp. = Sulpiride group; Sign. = Significance of Mann-Whitney tests for differences.)

The effect of sulpiride on blood-pressure and heart-rate, self-reported side-effects and mood were adopted from previous published work with the same cohort(Eisenegger et al., 2014; Naef et al., 2017). Supplementary Table 13 shows all side-effects measures, their changes over time, as well as the

results of a Mann-Whitney test for differences across treatment groups. Significance levels are not above chance level if corrected for multiple testing (Holm-Bonferroni correction).

Supplementary Note 1

Definition of reciprocal and mistake trials

To avoid ambiguity, we note the exact definition of Reciprocal and Mistake Trials used in the main text. Denoting *Backtransfer* as a variable for positive (1) or negative (-1) feedback from the Trustee, and *Change* as a variable for the relative change in investment from the previous trial, we defined Reciprocal and Mistake Trials as follows:

$$Reciprocal\,Trial = ((Backtransfer == 1\,\&\,Change > 0) \,|\, (Backtransfer == 1\,\&\,Investment == 10)) \,|\, ((Backtransfer == -1\,\&\,Change < 0) \,|\, (Backtransfer == -1\,\&\,Investment == 0)),$$

$$Mistake\,Trial = ((Backtransfer == 1\,\&\,Change < 0) \,|\, (Backtransfer == -1\,\&\,Change > 0))$$

Supplementary Note 2

Optimal behaviour in the repeated trust game

Optimal behaviour in the repeated trust game can be defined within various frameworks. Within the utility maximisation framework, a "rational" outcome-maximising agent would at each trial choose the investment that brings the highest expected value. When the trustee betrays the investor after an investment I, the investor ends up with an outcome of $10 - I$. When the trustee equalizes both players receive half of the overall gain. The outcome V in that case is:

$$V = \frac{I - 10 + 10 + 3 * I}{2} = 10 + I$$

If an agent has a subjective belief that the trustee will equalize with probability p, then the expected value of his investment I is:

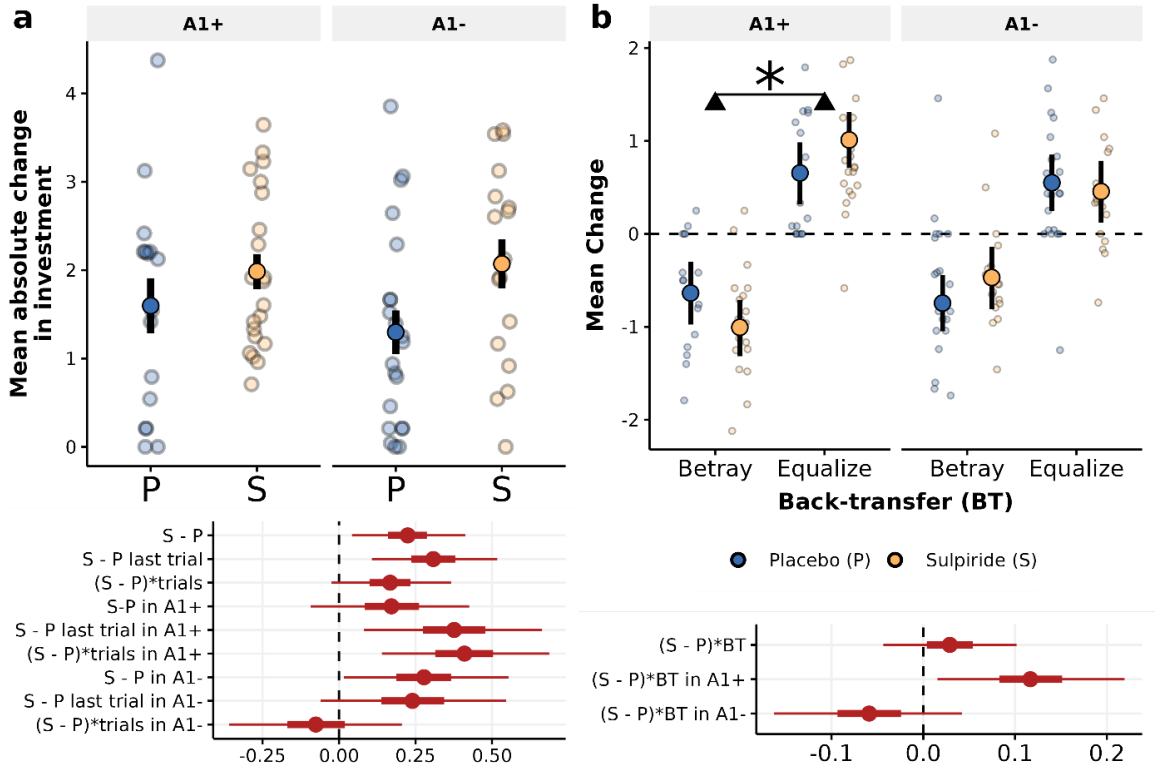
$$EV(I) = p * (10 + I) + (1 - p) * (10 - I) = 10 + I(2p - 1),$$

which is a linear function of p. Meaning that as soon as p is believed to be above 0.5 the outcome maximising agent should give a ten and as soon as p is below 0.5 they should invest 0.

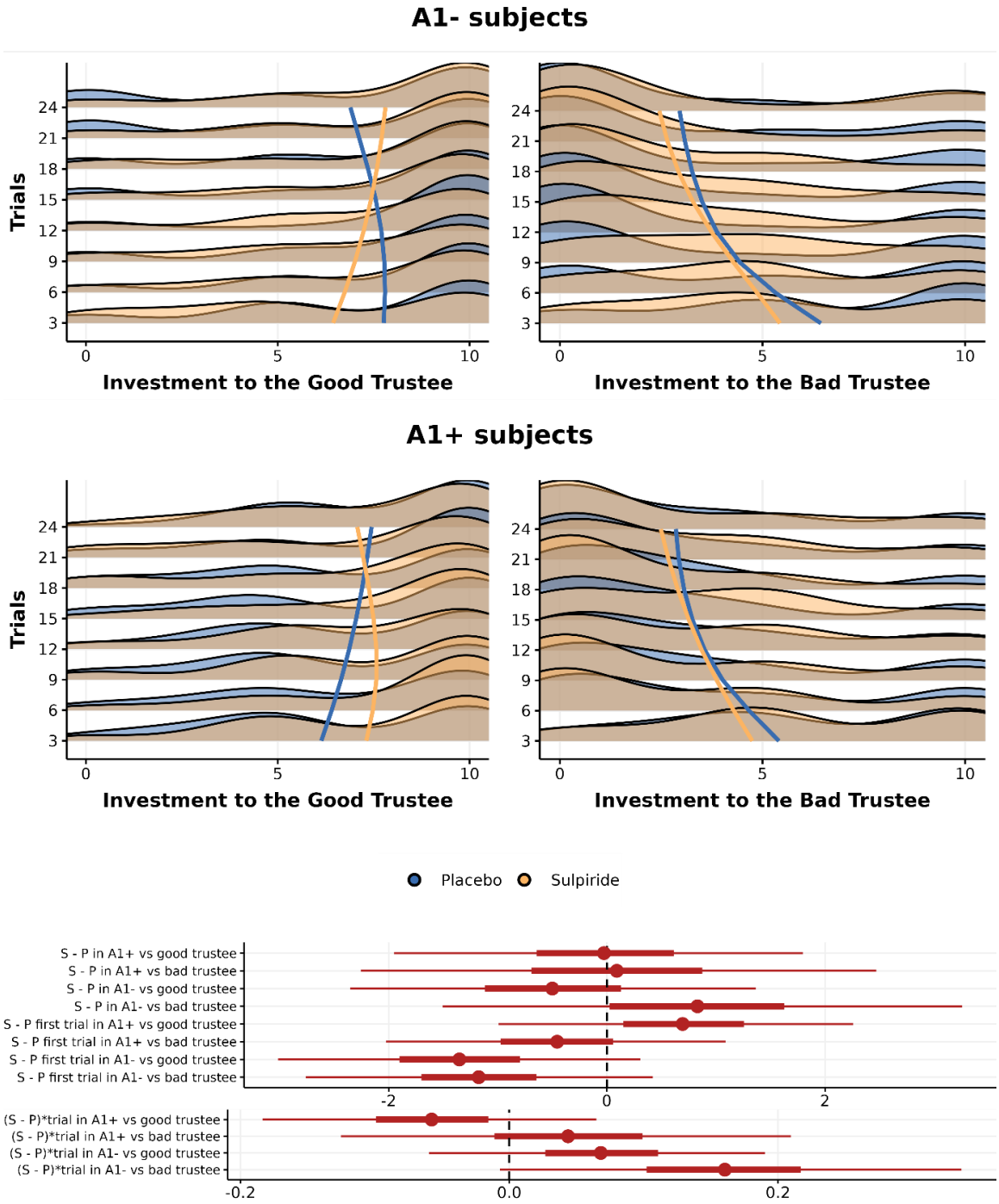
We note that in our model, an outcome-maximizing agent would have extremely high values of γ, which would also correspond to highly exploitative behaviour defined within the reinforcement learning framework.

Importantly, this outcome-maximizing agent does not consider the uncertainty about their estimate. On the other hand, a Bayesian perspective defines behaviour in this task as optimal given the investor's (participant's) subjective prior beliefs(Daunizeau et al., 2010). Update equations of the model are determined deterministically from the free parameters and describe Bayes optimal belief trajectories(Mathys et al., 2011). For example, if one's prior is that other people's behaviour is highly volatile, then it is optimal to adjust your belief after each feedback, whereas if one expects individuals to behave consistently, one should form an opinion quickly. With this, the Bayesian framework can describe subjectively optimal behaviours that are objectively maladaptive(Mathys et al., 2011).

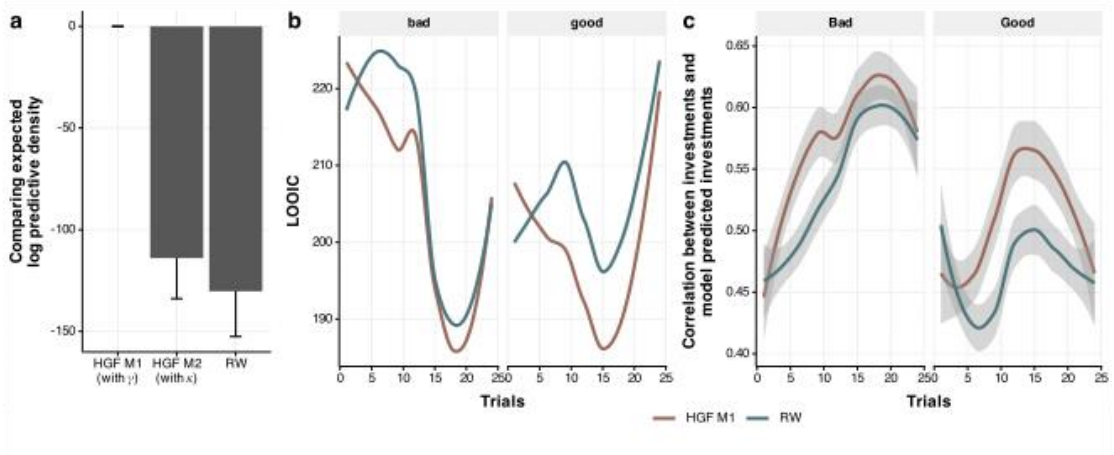
Supplementary Figures



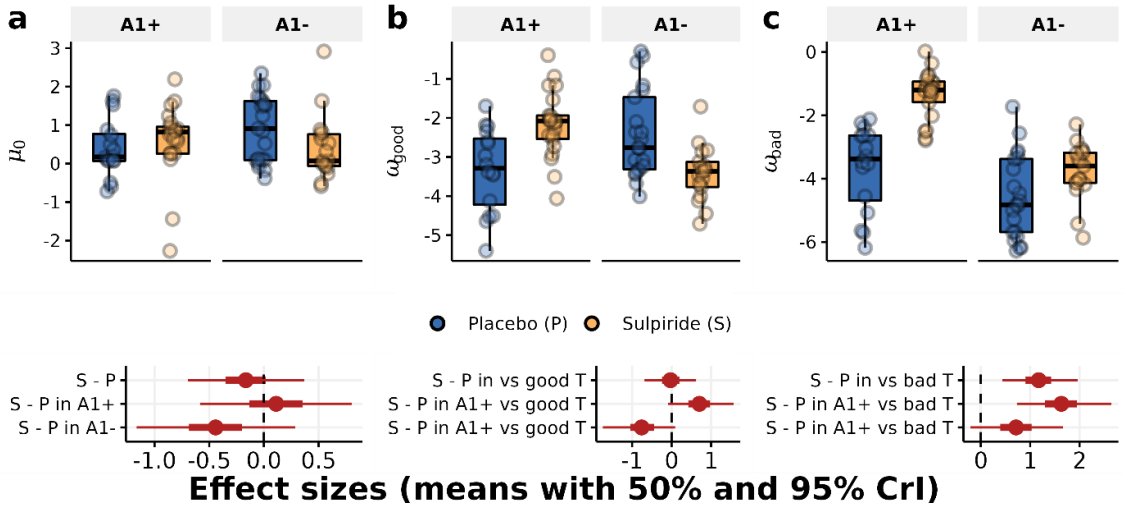
Supplementary Figure 1. Effects of sulpiride on behaviour in the repeated Trust game. a, Mean and standard errors of absolute change of investment from one trial to the next for both treatment and genotype groups. 95% CrI of effect sizes show a main effect of sulpiride, no credible evidence for an interaction effect of treatment and genotype, but a differential effect of sulpiride on the slope of two genotype groups. Sample sizes in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride. **b**, Mean and 95% CrI of relative change in investment following positive and negative feedback, for both genotype groups separately, predicted from a Bayesian multilevel model including Genotype as a predictor. Dots are raw means for each participant. Effect sizes below. The 95% CrI of the interaction effect of Drug and Back-transfer was above zero only in the A1+ group. We found no credible evidence for an interaction effect of sulpiride and Back-transfer (b = 0.089, 95% CrI [-0.135, 0.314], P(b<0) = 0.217, d = 0.029, 95% CrI [-0.044, 0.102]), but observed a three-way interaction effect, with the Genotype variable (b = -0.545, 95% CrI [-0.987, -0.103], P(b>0) = 0.007, d = -0.176, 95% CrI [-0.319, -0.034]). In the A1+ group, participants administered the D2 antagonist tended to increase their investment following positive, and decrease their investment following negative, responses from the trustees (b = 0.361, 95% CrI [0.048, 0.68], P(b<0) = 0.013, d = 0.116, 95% CrI [0.016, 0.219]). We found no credible evidence of a difference between sulpiride and placebo administration in A2 homozygotes (b = -0.179, 95% CrI [-0.502, 0.131] , P(b>0) = 0.127, d = -0.057, 95% CrI [-0.159, 0.042]). Sample sizes as in **a**.



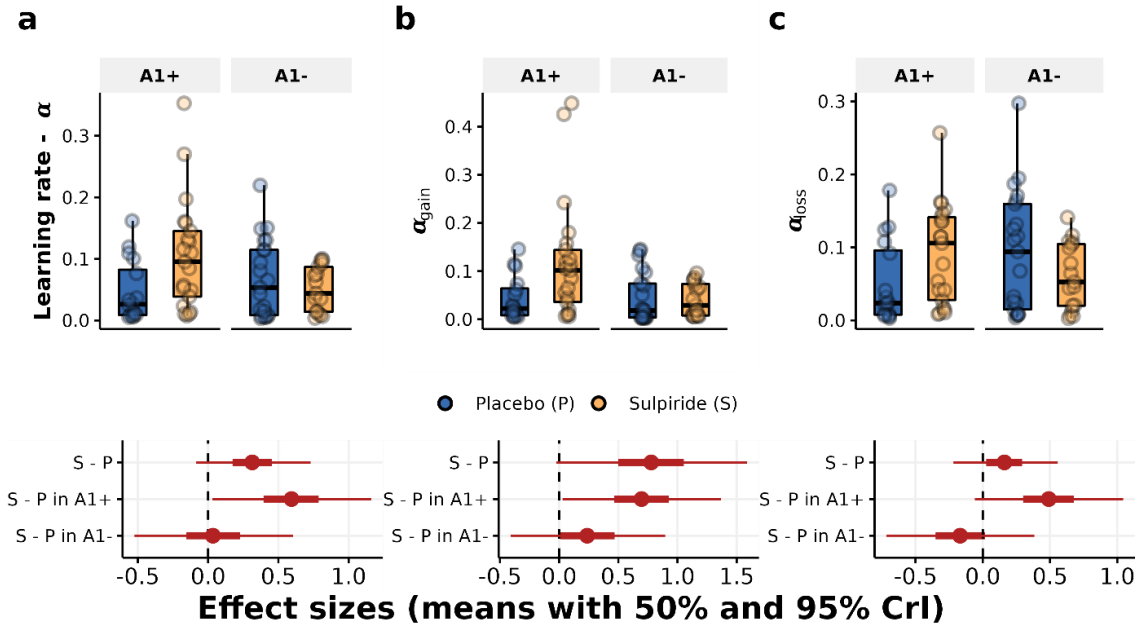
Supplementary Figure 2. Effects of sulpiride on average investments in the repeated Trust game. Density plots of investments for placebo and sulpiride group overlayed and grouped within the trustee and genotype (binned across 3 trials for clarity). Lines are means of investments predicted from a multilevel Bayesian model that included Trial, Genotype, Trustee and Treatment as fixed and Trustee and Trial as participant-level random effects. Effect sizes are shown below for average investments, differences in the first trial (initial trustworthiness) as well as for the slope with means (and 50% and 95% CrI).



Supplementary Figure 3. Model comparison. a, Comparing the HGF model with all four parameters (ω, μ_0, η, γ) to an HGF model that includes a coupling parameter κ, but does not include γ and to an RW model. On the y-axis is the relative difference in the predictive density (on a log scale) together with the standard errors (obtained by the loo_compare function of the loo package in R). Model comparison done for all 76 participants. **b**, The HGF model with γ and the RW model compared with the leave-one-out-information criterion across trials and trustees (25 trials per trustee), lower values imply a better fit. **c**, Same two models compared based on the correlations of actual investments and investments predicted by the model (higher values indicate a better fit). Lines depicted means per model and standard errors across all participants (n = 76).



Supplementary Figure 4. Sulpiride's effects on model parameters. a, Effect of sulpiride on initial trustworthiness belief. **b**, Effects of sulpiride on belief volatility when playing against the good trustee. **c**, Effects of sulpiride on belief volatility when playing against the bad trustee. Below are mean effect sizes with 50% and 95% CrI. Sample sizes in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride. Boxplots with centre line as medians, box bounds 25th and 75th percentiles, and whiskers terminating at maxima/minima (a distance of 1.5 times the IQR away from the 25th and 75th percentiles).



Supplementary Figure 5. Sulpiride effects on the RW model's learning rate. a, Average learning rate across trials. **b**, Learning rate for positive outcomes. **c**, Learning rate for negative outcomes. Below are mean effect sizes with 50% and 95% CrI. Sample sizes in A1+ group, n=17 placebo, n=21 sulpiride, and in A1- group n=21 placebo and n=17 sulpiride. Boxplots with centre line as medians, box bounds 25th and 75th percentiles, and whiskers terminating at maxima/minima (a distance of 1.5 times the IQR away from the 25th and 75th percentiles).

Chapter 4: Effects of dopamine D2 and opioid receptor antagonism on the trade-off between model-based and model-free behaviour in healthy volunteers

Original manuscript as published in eLife

DOI: 10.7554/eLife.79661

Nace Mikus•, [1, 2], Sebastian Korb[1, 3], Claudia Massaccesi[4], Christian Gausterer[5], Irene Graf[6], Matthäus Willeit[6], Christoph Eisenegger†, [1], Claus Lamm[1], Giorgia Silani*, [4], Chris Mathys*, [2, 7, 8]

* Correspondence to:

Nace Mikus, nace.mikus@univie.ac.at,

Affiliations:

[1] Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Austria,

[2] Interacting Minds Centre, Aarhus University, Aarhus, Denmark,

[3] Department of Psychology, University of Essex, Colchester, UK,

[4] Department of Clinical and Health Psychology, Faculty of Psychology, University of Vienna, Austria,

[5] FDZ-Forensisches DNA Zentrallabor GmbH, Medical University of Vienna, Austria,

[6] Department of Psychiatry and Psychotherapy, Medical University of Vienna, Austria,

[7] Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zurich and ETH Zurich, Zurich, Switzerland,

[8] Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy

* Senior authors contributed equally

† Deceased on the 27th of February 2017

Abstract

Human behaviour requires flexible arbitration between actions we do out of habit and actions that are directed towards a specific goal. Drugs that target opioid and dopamine receptors are notorious for inducing maladaptive habitual drug consumption; yet, how the opioidergic and dopaminergic neurotransmitter systems contribute to the arbitration between habitual and goal-directed behaviour is poorly understood. By combining pharmacological challenges with a well-established decision-making task and a novel computational model, we show that the administration of the dopamine D2/3 receptor antagonist amisulpride led to an increase in goal-directed or 'model-based' relative to habitual or 'model-free' behaviour, whereas the non-selective opioid receptor antagonist naltrexone had no appreciable effect. The effect of amisulpride on model-based/model-free behaviour did not scale with drug serum levels in the blood. Furthermore, participants with higher amisulpride serum levels showed higher explorative behaviour. These findings highlight the distinct functional contributions of dopamine and opioid receptors to goal-directed and habitual behaviour and support the notion that even small doses of amisulpride promote flexible application of cognitive control.

Introduction

Several theories of decision making postulate the existence of two distinct systems that drive our behaviour: a habitual system, which is automatic, reflexive and fast; and a goal-directed system, which is deliberative, reflective and effortful (Balleine & O'Doherty, 2010; Dickinson, 1985; Dolan & Dayan, 2013). This dichotomy of systems has a computational analogue in 'model-free' and 'model-based' decision-making models (Daw et al., 2005; Dolan & Dayan, 2013). A model-free agent simply selects actions that have led to rewarding outcomes in the past. This strategy is fast, computationally cheap, but can be inaccurate. A model-based agent uses an internal model of the environment to flexibly plan behavioural responses. This leads to more goal-oriented behaviour but is slower and relies on effortful cognitive control.

Studies investigating how individuals allocate control between the two systems have shown that model-based control is increased when there is more to gain (Kool, Gershman, & Cushman, 2017; Patzelt, Kool, et al., 2018), and decreased when cognitive resources are scarce (Otto, Gershman, et al., 2013; Otto, Raio, Chiang, Phelps, & Daw, 2013). Everyday behaviour can therefore be thought of as constant weighing of costs and benefits of applying model-based over model-free decision-making strategies (Daw et al., 2011; Kool & Botvinick, 2018; Kool et al., 2017). Failure to exert cognitive control over habitual urges in order to avoid negative outcomes has been suggested to be a hallmark of substance addiction (Dolan & Dayan, 2013; Everitt et al., 2008). In support of this, studies show that decreased model-based control is linked to stimulant use disorder (Voon et al., 2015) and seems to constitute a transdiagnostic dimensional trait related to compulsive behaviour across clinical (Voon et al., 2015, 2016) and non-clinical populations (Gillan et al., 2016). In light of increasing deaths from drug overdoses (EMCDDA, 2020; Seth, Scholl, Rudd, & Bacon, 2018), it is important to understand how different neurotransmitter systems affect the competition for cognitive resources when deciding between habitual and goal-directed actions.

Opiates, psychostimulants, and most other drugs of abuse increase the release of dopamine along the mesolimbic pathway (Chiara, 1999; Koob & Bloom, 1988), a circuit that plays a central role in reinforcement learning (Schultz et al., 1997). On top of this, the reinforcing properties of addictive drugs also depend on their ability to activate the μ opioid receptors (Becker, Grecksch, & Kraus, 2002; Benjamin, Grant, & Pohorecky, 1993; Le Merrer, Becker, Befort, & Kieffer, 2009). This suggests that both the dopamine and the opioid systems might be particularly relevant in model-free reinforcement learning processes that drive the formation of habitual behaviour. Studies in rodents show that activating receptors of both systems across the striatum increases cue-triggered wanting of rewards (Peciña & Berridge, 2013; Soares-Cunha, Coimbra, David-Pereira, et al., 2016). Conversely, inhibition of both D1-type and D2-type dopamine receptors (referred to as D1 and D2 from here on) as well as opioid receptors reduces motivation to obtain or consume rewards (Laurent, Leung, Maidment, &

Balleine, 2012; Peciña, 2008; Soares-Cunha, Coimbra, David-Pereira, et al., 2016). This data raises the hypothesis that the drift towards habitual control is enabled by dopamine and opioid receptors via a common neural pathway.

Recent work in humans provides some evidence in this direction, whereby systemic administration of opioid and D2 dopamine receptor antagonists causes a comparable reduction of cue responsivity and reward impulsivity (S. C. Weber et al., 2016) and decreases the effort to obtain immediate primary rewards (Korb et al., 2020). This suggests that when allocating control between the model-based and model-free system, dopamine or opioid receptor antagonists might comparatively disrupt model-free behavioural strategies and increase model-based behaviour. Yet, no study in humans has directly investigated this. Furthermore, disrupting habit formation might not in itself lead to increased model-based control, without either increasing the perceived value of applying cognitive control or decreasing the cost of doing so.

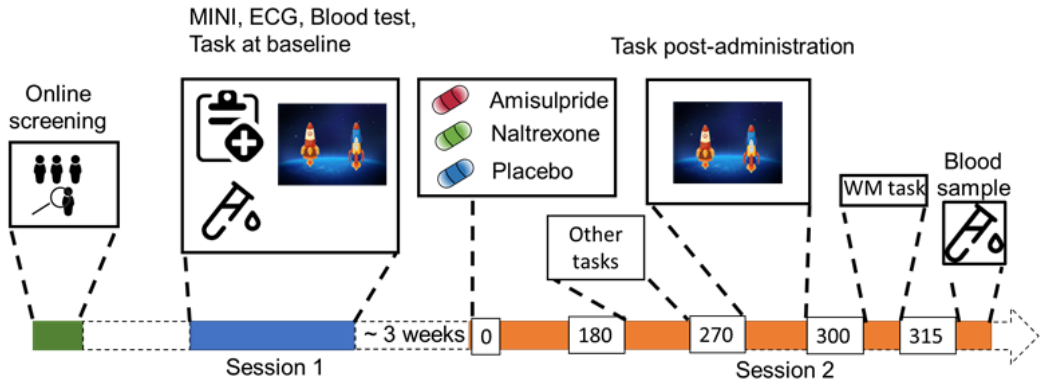
Crucially, there are important differences in how each of the two neurochemical systems relate to the types of cognitive control that are pivotal for model-based behaviour. Across a wide range of studies using various dosing schemes, opioid receptor antagonists did not have an effect on tasks in which exertion of cognitive control plays a key role, such as working memory (Del Campo, McMurray, Besser, & Grossman, 1992; File & Silverstone, 1981; Volavka, Dornbush, Mallya, & Cho, 1979), sustained attention (Zacny, Coalson, Lichtor, Yajnik, & Thapar, 1994), or mathematical problem-solving (Del Campo et al., 1992) (see (van Steenbergen, Eikemo, & Leknes, 2019) for a review). Dopaminergic circuits, on the other hand, are central in a variety of higher cognitive functions and goal-directed behaviour (Brozoski, Brown, Rosvold, & Goldman, 1979). In particular, D1 dopamine receptors in the prefrontal cortex enable maintenance of goal-relevant information and working memory (Goldman-Rakic, 1997; Sawaguchi & Goldman-Rakic, 1991; van Schouwenburg, Aaron, & Cools, 2010; Williams & Goldman-Rakic, 1995), while the D2 dopamine receptor activity disrupts prefrontal representations (Durstewitz & Seamans, 2008). In support of this, decreased working memory performance was observed after blocking prefrontal D1, but not prefrontal D2 receptors (Arnsten, 2011; Sawaguchi & Goldman-Rakic, 1991; Seamans & Yang, 2004). In humans, systemic administration of D2 antagonism increased the ability to maintain and manipulate working memory representations (Dodds et al., 2009; Frank & O'Reilly, 2006) and increased the value of applying cognitive effort (Westbrook et al., 2020). This data suggests that blocking D2 receptors, in contrast to blocking opioid receptors, could further facilitate model-based behaviour through enabling or encouraging flexible use of cognitive control.

This study compared the effects of the highly selective D2/3 dopamine receptor antagonist amisulpride with the opioid receptor antagonist naltrexone on the model-based/model-free trade-off in healthy volunteers. We tested 112 participants with a deterministic version of the two-step task (Daw et al., 2005; Kool et al., 2016), at baseline and after administering amisulpride (N = 38, 400 mg), naltrexone (N = 39, 50 mg), or placebo (N = 35) in a randomized, double-blind, between-subject design (Figure 1). Based on the assumption that the dopamine and opioid systems support the allocation of control between habitual and goal-directed systems through overlapping neural circuits, we expected both antagonists to comparatively increase model-based relative to model-free behaviour. Furthermore, based on the hypothesis that blocking D2 receptors affects the cost/benefit analysis of applying the model-based strategy, we expected a stronger shift towards model-based behaviour following amisulpride administration.

The two-step task used to assess the model-based/model-free trade-off is a well-established incentivized paradigm, where participants need to make choices at the first stage, which influence their outcomes in the second stage. Each trial of the task starts in one of two possible first-stage states, each featuring a pair of spaceships (Figure 2a). Upon choosing a spaceship, participants were taken to one of two planets in the second stage, where they encountered an alien that gave them points, which were converted to money at the end of the experiment. In each pair of spaceships, one spaceship flew to the red planet and one to the green. This deterministic mapping from spaceships to planets, as well as

which spaceships were paired together, stayed constant throughout the task. However, the points received on each planet changed independently with a Gaussian random walk. When playing the game, participants could therefore simply repeat the choice of spaceship that previously led to positive outcomes, regardless of which planet it was associated with (model-free behaviour, Figure 2b). Or they could attempt to remember which spaceship flew to which planet and choose the spaceship based on its associated planet (model-based behaviour, Figure 2b). The task tries to mimic real-life decision-making dilemmas, where we can use our knowledge about the world (e.g., a potentially dangerous virus is circulating) to override a habitual action (e.g., shaking hands when greeting someone) with a more appropriate one (e.g., greeting with an elbow bump or a bow).

Figure 1. Study Procedure. After an initial online screening, participants were invited to the lab for a first visit (Session 1),



where they were subjected to a medical check-up, before playing the two-step task for the first time. If they fulfilled the study criteria, they were invited for another visit (Session 2), where they received either 400 mg of amisulpride, 50 mg of naltrexone, or placebo (mannitol). After 180 minutes of waiting time, participants started with the test battery. Approximately 270 minutes after drug intake, participants performed the two-step task the second time, followed by a Reading Span (working memory) task, and a blood draw to determine amisulpride serum levels.

As an initial straightforward approach to dissociate model-based and model-free behaviour, we first examined how previous reward points affected the probability to stay with the previous choice. We then turned to computational modelling and explicitly modelled the relative weight, ω, of model-based compared to model-free behaviour and disentangled it from other processes that might affect participants' choices, such as exploration and reward devaluation. To assess any effects of the two pharmacological compounds on cognitive control, we also collected data on working memory performance through the Reading Span task (Klaus & Schriefers, 2016). We also aimed to control for changes in mood through a mood questionnaire delivered at drug intake and 3 h later. Whereas the dose we chose for naltrexone provides sufficient occupancy of opioid receptors, the dose for amisulpride is limited by the potentially detrimental side-effects of D2 antagonists (Takano et al., 2006). To explore the variance of drug effects across participants, we examined amisulpride serum levels in the blood. Finally, previous research implicated several genotypes related to baseline dopamine function in model-based/model-free behaviour. To control for these baseline differences and to explore potential genotype drug interactions, we collected data on four genetic markers related to prefrontal and striatal dopamine levels. The data and the analysis scripts are available open-access at https://github.com/nacemikus/mbmf-da-op.git.

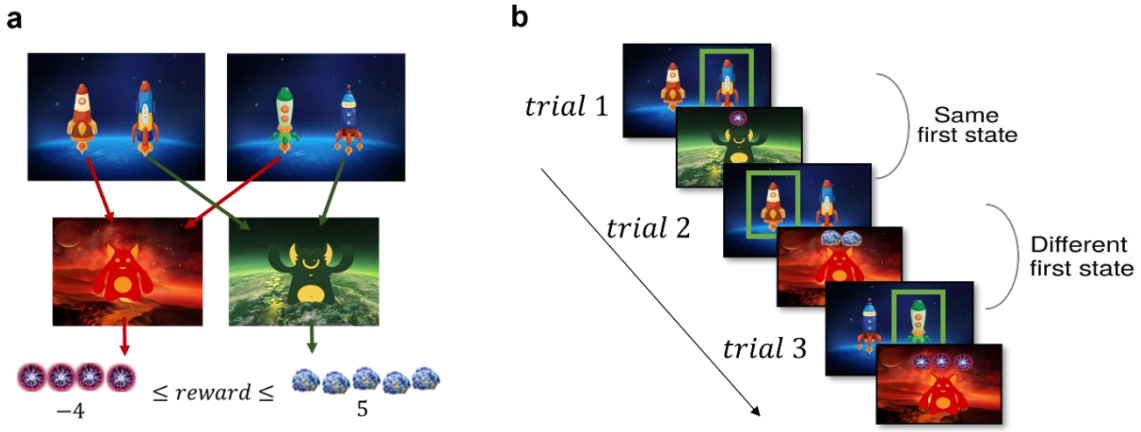


Figure 2. Task Design and Hypotheses. a, In the first stage of the trial, the participants were presented with one of the two pairs of spaceships. Each of the spaceships flew to one of two planets in the second stage, where they encountered an alien that gave them points. The transition from each of the four spaceships to the planets was deterministic and stayed the same throughout the experiment. The points each alien gave changed independently according to a discretized Gaussian random walk with bouncing boundaries at -4 and 5. **b,** Behaviour in the task showcasing trials where the previous first-stage state (spaceship pair) was the same (trial 2) and trials where it was different (trial 3). High number of points should encourage choosing the spaceship that flies to the same planet, what we term as "staying with the previous choice". Note that in the featured example, the participant after receiving -1 point in the trial 1 opted against staying with the previous choice in the next trial, and after receiving 2 points in trial 2 opted for staying with the previous choice, by choosing the spaceship, that flew to the same planet.

Results

Effects of dopaminergic and opioidergic antagonism on staying with previous choices

One way to quantify the trade-off between the two systems is by looking at the interaction effect of previous points on staying behaviour in trials where knowledge about the task structure is irrelevant (same first stage state as in the previous trial) with trials where it is relevant (different first stage states as in the previous trial). With this we control for more general effects of the drugs on decision-making behaviour and isolate the effect on the allocation of control between the two systems. As evident from Figure 3 (and Supplementary File 1a), previous points increased the likelihood to stay with the previous choice (β_{logods} = 0.391, with a 95% credible interval (CI) [0.238, 0.542], and the proportion of the interval below zero P($\beta_{logodds}$<0) < 10e-3), however this was significantly less the case in trials with different compared to the same first stage states (β_{logods} = -0.186 (95% CI [-0.321, -0.051], P($\beta_{logodds}$>0) < 0.005). This indicates that participants often failed to consider the mapping from spaceship to planets when making choices in trials where the first stage state differed from the previous trial.

Compared to placebo, amisulpride significantly increased the difference between the effects of previous points on staying behaviour in different vs. same first state trials, as indicated by a significant four-way interaction ($\beta_{logodds}$ = 0.176, 95% CI [0.036, 0.351], P($\beta_{logodds}$<0) < 0.007, Figure 3). When looking at each type of trial separately, we found that amisulpride also decreased the effect of previous reward points on the probability to stay when the first-stage state was the same ($\beta_{logodds}$ = -0.140, 95% CI = -0.274, -0.009], P($\beta_{logodds}$>0) = 0.02), and slightly (and non-significantly) increased it when the first-stage state was different ($b_{logodds}$ = 0.036 (95% CI -0.091, 0.158], P($\beta_{logodds}$<0) = 0.285). For naltrexone, the log odds estimate of the four-way interaction coefficient was centred around 0.016, (95% CI [-0.127, 0.159], P($\beta_{logodds}$<0) = 0.41), indicating no marked difference between trial types, neither were there significant differences when looking at each trial-type separately (all P > 0.32).

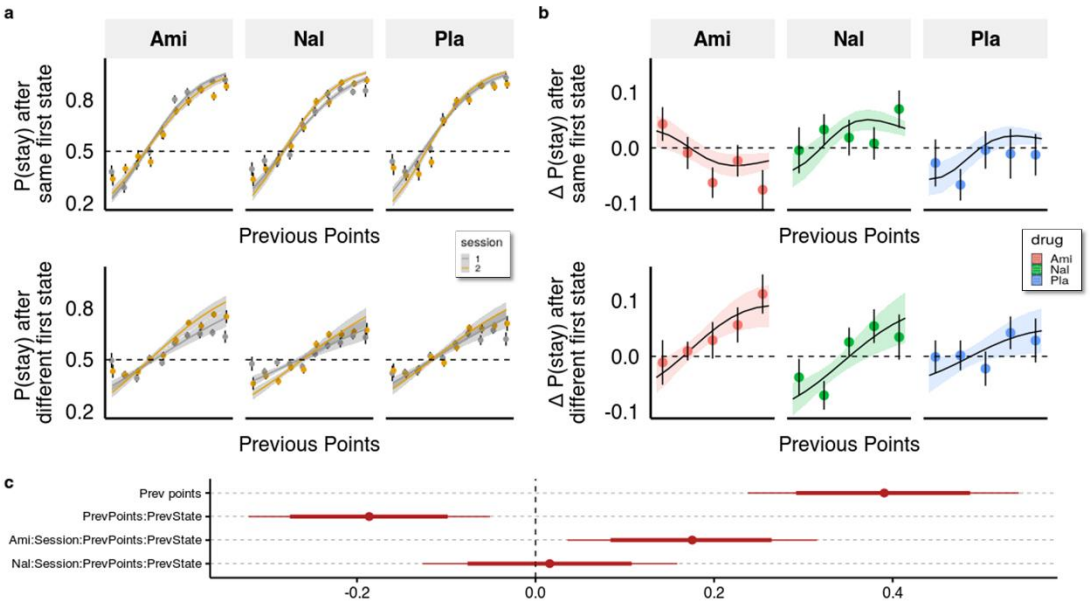


Figure 3. Behavioural analysis. a, We used a hierarchical Bayesian logistic regression model to analyse how the probability to stay with the previous choice depended on previous points, previous first-stage state, session, and drug administration and their interaction. We allowed the intercept and the slopes for previous first-stage state and session to vary by participant (full table of coefficients in Supplementary File 1a). Points and error bars depict mean proportions and standard errors in trials where the participants stayed with their previous choice for each previous point averaged within session and drug group. Lines and ribbons depict the mean estimates and 80% credible intervals. When encountering a trial with the same first-stage state as in its preceding trial, participants were 1.495 times (95% CI [1.286, 1.739], $P(\beta_{logodds} < 0) <$ 10e-3) more likely to stay with their previous choice for each additional previous point ($\beta_{logodds} = 0.402$, 95% CI [0.251, 0.553], $P(\beta_{logodds} < 0) <$ 10e-3). In contrast, in trials where the first-stage state was different from the preceding trial, the odds (on the logarithmic scale) of repeating the previous choice of spaceship were reduced, with each additional point earned in the previous trial, by -0.204 (95% CI [-0.263, -0.146], $P(\beta_{logodds} > 0) <$ 10e-3), and participants were only 1.219 times (95% CI [1.053, 1.410], $P(\beta_{logodds} < 0) <$ 10e-3) more likely to stay with their choice for each additional previous point. This indicates that participants often failed to consider the mapping from spaceship to planets when making choices in trials where the first-stage state differed from the previous trial. The effects of the drugs can be seen by the different slopes of previous points in the two sessions depicted for both trial types. **b,** Differences in staying behaviour between sessions, binned into 5 different reward levels for clarity. Means with standard errors overlayed with means and 80% CI of estimated posterior distributions. **c**, Means with 80% and 95% CIs of effect sizes (in logodds space) of selected regression coefficients of the hierarchical logistic regression model predicting staying behaviour from previous points (PrevPoints), previous first-stage states (PrevState), session, drug administrations, and all the interactions between them. Ami, amisulpride; Nal, naltrexone, Pla, placebo.

These results suggest that naltrexone did not affect the trade-off between goal-directed and habitual behaviour in this task. Amisulpride, on the other hand, increased performance in trials where keeping track of the mapping between spaceships and planets was advantageous, compared to those where it was not. However, somewhat counterintuitively, amisulpride also had a seemingly detrimental effect on decision-making overall, as indicated by the reduced effect of previous points on staying behaviour in trials with the same first-stage state. This pattern of behaviour might be due amisulpride's effect on both model-based/model-free behaviour and other decision-making processes, such as the tendency to explore other options. We note also that this analysis approach ignores one important aspect of the task, namely that subjects were comparing the relative points between two independently changing planets, thereby the points received in each trial should be seen in relation to the points that the participants expected to get were they to choose the other spaceship. To address these issues, we employed computational modelling.

Estimation of drug effects with computational modelling

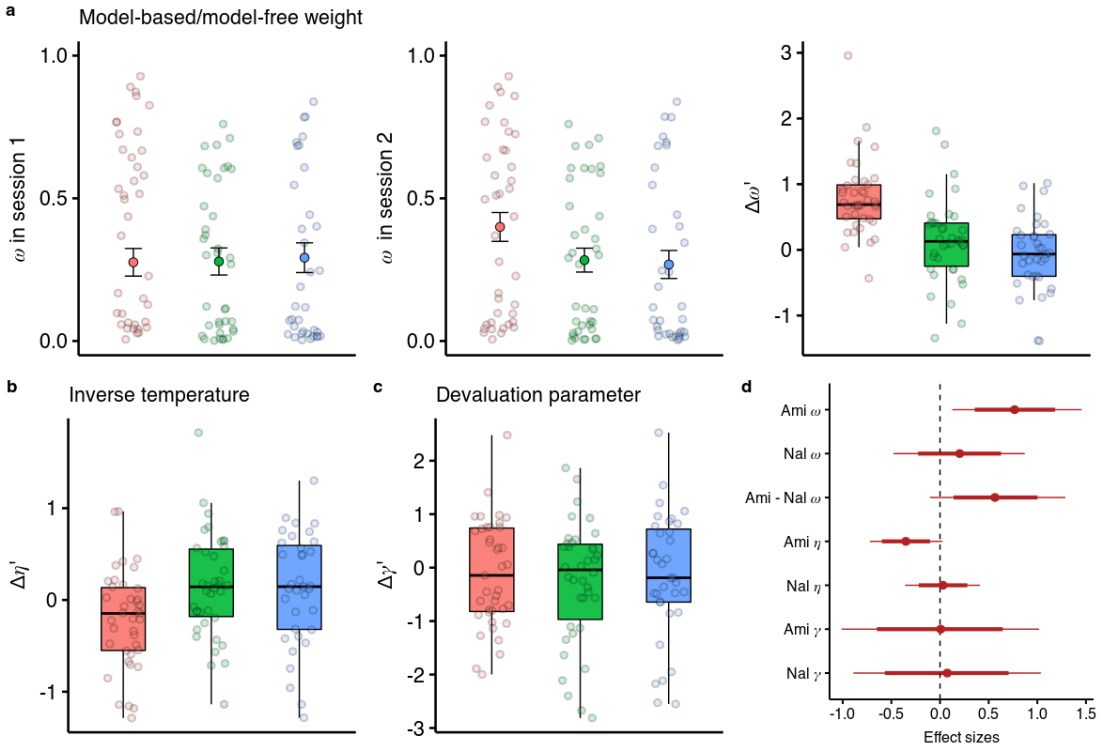


Figure 4. Effects of amisulpride and naltrexone on the model parameters. The best performing model (M1) is described with three free parameters, ω (the degree of model-based vs model-free value contributions to choice), γ (the degree of devaluation of unencountered spaceships), and η (the inverse temperature in the softmax mapping from values to probabilities). The parameters for both sessions and effects of drug treatments were estimated in one hierarchical model. **a**, Going from session 1 to session 2, amisulpride administration led to higher estimations of ω, and therefore increased model-based relative to model-free control. The difference between sessions of the model-based weights is shown in parameter estimation space (hence the prime). **b**, Difference in the parameter γ. **c,** Difference in the inverse temperature parameter η. Lower values mean higher exploration. **d,** Posterior distributions of the effect of drugs, compared to placebo, on group-level mean session differences. Effect sizes with 80% and 95% CI. Ami, amisulpride; Nal, naltrexone, Pla, placebo.

We defined a model (M1) where, similarly to the computational models previously used with this task (Kool et al., 2016, 2017), the trade-off between the goal-directed and habitual components is captured by the weighting parameter ω, which embodies the degree to which the choice of participants on each trial is influenced by model-based ($\omega = 1$), or model-free ($\omega = 0$) subjective values of each of the spaceships. In this model, the subjective values of both the model-free and the model-based model components are defined as the last observed outcome following the choice of that spaceship. The crucial difference between the two components is that, whereas the model-free agent learns only by experiencing direct outcomes of spaceship selection, the model-based component always considers the deterministic mapping from spaceships to planets, and thus learns the subjective values of planets. We also include an inverse temperature parameter η (lower η indicates more explorative or noisy behaviour) and a *discounting parameter* γ that marks the degree of devaluation of non-chosen (and non-encountered) spaceships for the model-free component in each trial.

We compared the above-described model to two Dual-systems Reinforcement Learning (RL) models (Figure S1a). In these two models, the degree to which an outcome in each trial affects the subjective values of actions at each stage is determined by a learning rate parameter. The model-free agent thus learns the subjective action values in each stage from experience, by increasing the values of actions and states that lead to outcomes that were better than expected and decreasing the values when the outcome was worse than expected. We allow the learning rate at both stages to either be different (model M2) or the same (model M3). Note, that the model M1 is version of these RL model, where the learning rates are set to 1 and a devaluation parameter is included on the non-chosen option. This

is motivated by the observation that the rewards change according to a Gaussian random walk (and are not probabilistic), and therefore the last encountered outcome is the best guess the agent can make. When comparing the performance of the three models we found that the model M1 has better out of sample predictive accuracy compared to the other two models (Figure S1b). We verified the winning model with parameter recovery (Figure S1c) and posterior predictive checks (Figure S1d-f).

We first observed that the more model-based choices the participants made, the more money they earned (r = 0.65, 95% CI [0.53, 0.76]). This serves as a validity check of the task, which was designed to make cognitive control pay off (literally) (Kool et al., 2016). We then looked at how the model parameters relate to the random slopes from the behavioural analysis of staying behaviour and found that the participant-level (random effect) slope for the effect of previous points on staying behaviour in different vs. same first state trials was most strongly related to ω (d = 0.493, P < 10e-3) and negatively related to the inverse temperature parameter η (d = -0.328, P < 10e-3), and the slope for trials with same first states was mostly related to η (d = 0.822, P < 10e-3), and less so to ω (d = 0.235, P < 10e-3).

Dopaminergic antagonism increases model-based relative to model free control

We embedded the model parameter estimation within a hierarchical Bayesian inference framework and estimated the drug effects on all three parameters in one model (L. Zhang et al., 2020). We found that under amisulpride the difference in ω between the two sessions is higher than in the placebo group ($\beta_{ami}^{\Delta\omega}$ = 0.787, 95% CI [0.131, 1.510], P($\beta_{ami}^{\Delta\omega}$ <0) = 0.010, Figure 4a), with effect size d = 0.758, (95% CI [0.126, 1.455]). In contrast, there was no session difference in ω between naltrexone and placebo (β_{nal}^{ω} = 0.238, 95% CI [-0.443, 0.856], P(β_{nal}^{ω} <0) = 0.250), with a marginally significant difference with a moderate effect size between the effects of the two compounds ($\beta_{ami-nal}^{\Delta\omega}$ = 0.578, 95% CI [-0.108, 1.337], P($\beta_{ami-nal}^{\Delta\omega}$ <0) = 0.046, d = 0.557, 95% CI [-0.104, 1.289]).

Dopaminergic antagonism increases the exploration parameter in participants with high serum levels

We found no evidence that either the blockade of dopamine D2 or opioid receptors influenced the devaluation parameter γ (Figure 4b, d, β_{ami}^{γ}= 0.01, 95% CI [-1.00, 1.00], β_{nal}^{γ}= 0.07, 95% CI [-0.88, 1.02]). However, we found some evidence for amisulpride effects on the inverse temperature parameter η (Figure 4c, d, β_{ami}^{η} = -0.33, 95% CI [-0.67, 0.02], P(β_{ami}^{η} >0) = 0.034, d = -0.38, 95% CI [-0.78, 0.03]). This would imply that amisulpride increases "explorative" choices or choices that the model would not predict based on estimated action values. To verify if our model performs worse for the amisulpride group we looked at how the pharmacological treatment predicts out-of-sample prediction accuracy and found no differences between groups (Figure 4 – figure supplement 1, 1f, Supplementary File 1b). We also note that the differences between sessions in ω and η are not negatively correlated (r = 0.12, 95% CI [-0.06, 0.30]), suggesting that the two effects are not related to each other.

We next looked at whether the effects of amisulpride were different based on the effective dose. We reran the parameter estimation including a variable for amisulpride blood serum levels as a group-level covariate in the hierarchical model (Figure 5). We used a categorical variable due to the highly skewed distribution of serum levels(DeCoster, Gallucci, & Iselin, 2011) (see Methods for details). We found that amisulpride increased ω in both the low serum group (b = 0.919, 95% CI [0.216, 1.722], P(b<0) = 0.006, d = 1.013, 95% CI [0.238, 1.897]) and the high serum group (b = 0.872, 95% CI [0.008, 1.853], P(b<0) = 0.024, d = 0.961, 95% CI [0.009, 2.041]), with no difference between the effects (P(b>0) = 0.458). However, when looking at the effects of amisulpride on η we found that it was not reduced in the low serum group (b = -0.105, 95% CI [-0.523, 0.348], P(b>0) = 0.323, d = -0.096, 95% CI [-0.477, 0.317]), but was reduced in the high serum group (b = -0.492, 95% CI [-0.96, -0.033], P(b>0) = 0.018, d = -0.45, 95% CI [-0.877, -0.03], with a (non-significant) moderate effect size difference

between the effects (b = -0.393, 95% CI [-0.918, 0.118], P(b>0) = 0.066, d = -0.359, 95% CI [-0.838, 0.108]).

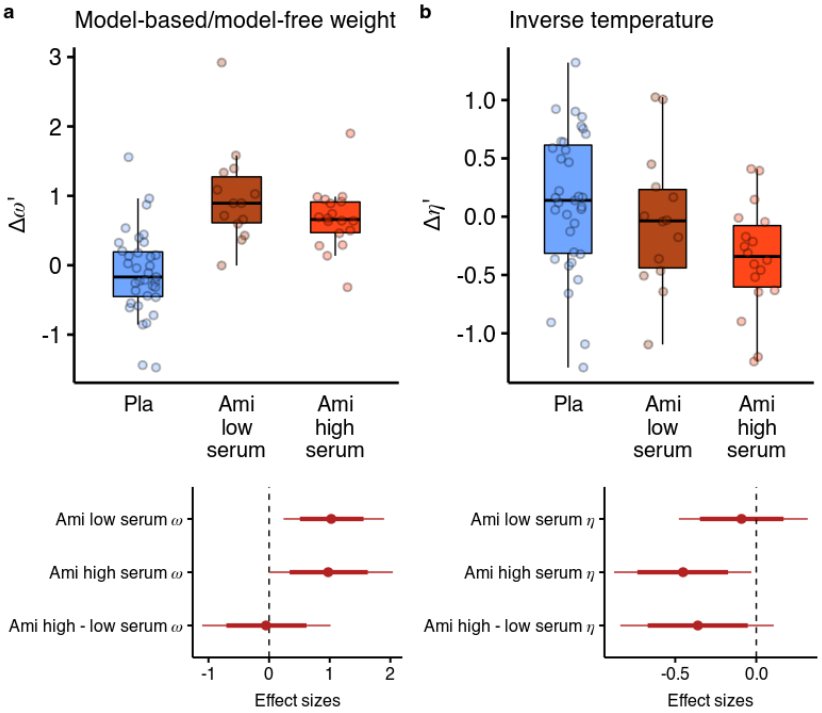


Figure 5. Effects of amisulpride across the high and low blood serum levels. a, Across both effective dose groups of participants, amisulpride increased the model-free/model-based weight parameter. **b**, Amisulpride increased exploration (decreased η) in the group that had high blood serum levels, but not in the group with low serum levels. Effect sizes with 80% and 95% CI. Ami, amisulpride; Pla, placebo.

We also reran the behavioral analysis, predicting the likelihood to stay with the previous choice as before, but including the serum variable. In line with the results from the computational model, we found that amisulpride significantly increased the difference between the effects of previous points on staying behavior in different vs. same first state trials both in the low serum group ($\beta_{logodds}$= 0.235, 95% CI [0.047, 0.43], P($\beta_{logodds}$<0) = 0.006), as well as in the high serum group, although to a lesser extent ($\beta_{logodds}$ = 0.143, 95% CI [-0.03, 0.323], P($\beta_{logodds}$<0) = 0.054). Conversely, in the high serum group, amisulpride decreased the effect of previous points in the same first state trials (b = -0.172, 95% CI [-0.34, -0.002], P($\beta_{logodds}$>0) = 0.024), mirroring the effect of the computational model. In the low serum group amisulpride also decreased the effect of previous points on staying behaviour in the first state trials ($\beta_{logodds}$ = -0.147, 95% CI [-0.34, 0.036], P($\beta_{logodds}$>0) = 0.060). Note that, although the 95% CI contained values above 0, this represents a slight inconsistency to the results from the computational model, where the exploration parameter η was not reliably reduced in the low serum group.

Lower values of η as well as lower effect of previous points on behaviour implies that participants behaviour more stochastically (or noisily) and that the computational model cannot explain that behaviour based on estimated values of the two spaceships. To see whether some of the increased variance can be explained by the tendency of participants to repeat actions or choices regardless of outcomes, we extended the M1 serum model to include a response stickiness (ρ) and stimulus stickiness (π) parameter. We found some evidence that amisulpride in participants with high serum levels increased switching between responses, as indicated by the reduced ρ parameter (b = -0.466, 95% CI [-0.962, 0.043], P(b>0) = 0.039). This effect was not present in the low serum group (b = -0.158, 95% CI [-0.618, 0.311], P(b>0) = 0.251), with a non-significant difference between the effects (b = -0.294, 95% CI [-0.723, 0.13], P(b>0) = 0.082). The inclusion of the two stickiness parameters, did not markedly change the posterior distributions of effects of amisulpride on η, but did reduce the effect on

ω in the high serum group (b = 0.571, 95% CI [-0.361, 1.527], P(b<0) = 0.111, see Figure 4 – figure supplement 2 for other relevant posterior distributions). We note however, that the model including stickiness parameters performed worse than the model without the stickiness parameters, with a lower out-of-sample trial-wise predictive performance ($\Delta elpd$(sd)= -413 (31.5)).

Working memory, mood, and genetics

To examine whether working memory capacity was affected by our drug manipulation, we analysed performance in the Reading Span task and found no evidence for drug effects on proportion of correct recalls (β_{ami}^{wm} = -0.02, 95% CI [-0.07, 0.02], β_{nal}^{wm} = -0.03, 95% CI [-0.08, 0.02]). We also found no difference in mood changes following amisulpride administration from drug intake and 3 hours in either serum group(Supplementary Files 1c, 1d, 1e), but found that naltrexone flattened both positive (b = -0.411, 95% CI [-0.855, 0.029], P(b>0) = 0.035) and negative affect (b = -0.524, 95% CI [-1.001, -0.061], P(b>0) = 0.013). To explore what other group level variables might influence the drug effects, we ran two linear models predicting differences in ω and η between sessions including working memory and mood ratings as covariates in the analysis and Sex, Weight and Age as moderator variables. We found no significant effects and neither did conditioning on the covariates affect inference of drug effects on ω or η (Supplementary Files 1f, 1g).

We also found no robust effects of genotype variables on either baseline measures or on the drug effects on either ω or η (Supplementary Files h, i). In line with previous studies using similar paradigms, we used the *COMT* single nucleotide polymorphism (SNP) as a marker for prefrontal dopamine function, and *DARPP-32* SNP as a marker of striatal D1 receptor function (Doll, Bath, Daw, & Frank, 2016; Mannisto & Kaakkola, 1999; Trantham-Davidson, Neely, Lavin, & Seamans, 2004). We also investigated two genetic markers for striatal dopamine levels, namely the *Taq1A* SNP and the *DAT1* polymorphism, a 40 base-pair variable number tandem repeat polymorphism (VNTR) of the dopamine transporter (Eisenegger et al., 2014, 2013; Laakso et al., 2005) (see Supplementary Note 1 for details on genotypes and how they relate to striatal and prefrontal dopamine function).

Discussion

Drugs that stimulate opioid or dopamine neurotransmitter systems can lead to compulsive drug-taking, which can result either from increased reliance on habits or from failures to exert cognitive control over habitual urges. Experimental studies strongly implicate both the dopamine and opioid neurotransmitter systems in processes related to habit formation. Here, we asked whether pharmacologically blocking opioid and dopamine receptors promotes the use of goal-directed over habitual control in healthy volunteers. Using the two-step task and a novel computational model, we found that blocking D2-like dopamine receptors increases the weight on model-based relative to model-free control, whereas blocking opioid receptors has no appreciable effect, with a marginally significant difference between the effects (with a moderate average effect size). Additionally, we found that amisulpride increases choice exploration, particularly in the group of participants with high amisulpride serum levels in the blood. The results from the computational model mirrored those obtained when analyzing the effects of previous points on staying with prior choices. The degree to which reward in a previous trial affected choice repetition in trials with the same first stage states was mostly related to the exploitation/exploration parameter in the model and was reduced under amisulpride. Conversely, the relative increase in the effect of previous points on choice repetition between different vs same first stage state trials was more related to the model-based/model-free weight and was increased under amisulpride. Our findings provide initial evidence for a divergent involvement of the dopamine and opioid neurotransmitter systems in the shift between habitual and goal-directed behaviour. The lack of effects of naltrexone on the model-based/model-free trade-off also provides some support for the notion that simply disrupting neurobiological systems that subserve habitual behaviour might not be enough to increase goal-directed behaviour in this task. An increase in the model-based/model-free weight following amisulpride administration advocates for dopamine playing

a decisive role in flexibly applying cognitive control to facilitate model-based behaviour and highlights the specific functional contribution of the D2 receptor subtype.

Prior research on dopamine's involvement in the arbitration between the two systems has produced inconsistent results. Elevated dopamine levels, either at baseline or following pharmacological treatments such as L-dopa, are related to increased model-based behaviour in some studies (Deserno, Huys, et al., 2015; M. E. Sharp, Foerde, Daw, & Shohamy, 2016; Wunderlich, Smittenaar, & Dolan, 2012), but show no correlation or even opposite effects in others (Kroemer et al., 2019; Voon et al., 2020). This suggests that the effect of dopamine promoting agents on model-based/model-free trade-off might depend on the relative contributions of the various region-specific dopamine receptor subtypes. Studies have repeatedly shown that prefrontal dopamine circuits enable goal stability and distractor avoidance, primarily through D1 dopamine receptors (Sawaguchi & Goldman-Rakic, 1991; van Schouwenburg et al., 2010; Williams & Goldman-Rakic, 1995). According to the dual-state theory of prefrontal dopamine function (Durstewitz & Seamans, 2008), a state dominated by D1 receptor activity is characterized by a high-energy barrier and supports stabilization of prefrontal representations. Conversely, a D2 receptor dominated state has a low-energy barrier and facilitates flexible switching between representational states. In line with this, D2 antagonists impair behaviour in tasks that require constant attentional set shifting (Mehta, Manes, Magnolfi, Sahakian, & Robbins, 2004), but might improve performance in tasks where prefrontal goal representations are required (Kahnt, Weber, Haker, Robbins, & Tobler, 2015). For example, D2 antagonism improves the ability of participants in certain cognitive tasks that require manipulation of task-relevant information in the working memory (Dodds et al., 2009; Frank & O'Reilly, 2006), but does not improve working memory capacity, or memory retrieval (Dodds et al., 2009; Mehta et al., 1999; Naef et al., 2017). The effects of D2 antagonism on model-based/model-free behaviour in our study can be interpreted within this framework to result from increased ability to maintain prefrontal representation of the mapping between the spaceships and the planets online. However, this is difficult to reconcile with the fact that model-based behaviour in dynamic learning paradigms, such as the one used here, also requires flexible updating of action values.

Theories of adaptive control emphasize that the prefrontal regions subserving cognitive control interact with striatal dopaminergic circuits (Frank, 2005; McNab & Klingberg, 2008). In the striatum, phasic dopaminergic bursts subserve prediction error propagation and action initiation(Frank et al., 2004) mainly, but not exclusively, through D1 dopamine receptors (Frank et al., 2004; Montague et al., 1996; Soares-Cunha, Coimbra, David-Pereira, et al., 2016). Slower baseline dopamine changes in tonic activity are believed to invigorate physical action (Niv, Daw, & Joel, 2006) based on cost-benefit tradeoffs (Salamone et al., 2016). More recent work extended this framework and proposed that the same circuits subserve cost-benefit analysis of applying cognitive effort, framing the role of striatal dopamine as mediating the decision to apply cognitive control based on perceived value vs costs of doing so (Cools, 2016, 2019; Westbrook, Frank, & Cools, 2021). In a recent study that combined neurochemical imaging with pharmacological administration of 400 mg of sulpiride (another D2 dopamine receptor antagonist) and methylphenidate (a drug that boosts striatal dopamine availability) showed that both drugs comparatively increased the willingness to exert cognitive effort, particularly in subjects with low striatal dopamine availability (Westbrook et al., 2020). An important aspect of drugs that target D2 receptors is that they are known to mainly act on presynaptic D2 receptors at low doses, while binding to postsynaptic receptors prevailingly accounts for their effects at higher doses (Schoemaker et al., 1997). Presynaptic D2 receptors are believed to play an autoregulatory control of dopamine signalling through inhibition of synthesis and dopamine release (Ford, 2014). In our data, we find that the increase in model-based control did not scale with higher effective dose and was present in participants with either high or low serum levels in the blood. The effect in the high serum group was slightly reduced (and not robustly above zero, $p = 0.111$) in the computational model that included stickiness parameters, although we note that including stickiness parameters led to a poorer model fit. Taken together, this provides some evidence for the notion that the effects of amisulpride on model-based behaviour are not driven by postsynaptic effects but might rather be due to amisulpride

increasing dopamine levels through presynaptic D2 receptor blockade and thereby boosting the willingness to apply cognitive effort.

Interestingly, amisulpride also increased choice stochasticity parametrised by the softmax inverse temperature parameter. In a paradigm with two choice options, it cannot be definitively determined whether this indicates higher decision-noise or increased exploration of alternative choices. We can however speculate that increased decision noise would lead to overall detrimental effects on learning in both trial types with same and different consecutive first stage states, which we do not observe in our data. The effect on the choice stochasticity parameter was only present in participants with a higher effective dose (Rosenzweig et al., 2002), suggesting that the effect was more likely to be post-synaptic. Similarly, in the same effective dose group, we found some evidence that amisulpride reduces response stickiness indicating increased switching between actions. This is well in line with a prominent model of the cortico-striatal circuitry implicating post-synaptic D2 receptors in exploration/exploitation (Frank, 2005) and supported by empirical data. In animal studies, activation of D2 receptors was shown to lead to choice perseverance and more deterministic behaviour, whereas D2 receptor inhibition increases the probability of performing competing actions and increases randomness in action selection (Sridharan et al., 2006). In humans, a recent neurochemical imaging study showed that D2 receptor availability in the striatum correlated with choice uncertainty parameters across both reinforcement learning and active inference computational modelling frameworks(Adams et al., 2020). Increased choice uncertainty was also observed in a social and non-social learning tasks in a study using 800 mg of sulpiride, a dose that is known to exert post-synaptic effects (Eisenegger et al., 2014; Mikus, Eisenegger, et al., 2022). We note, however, that the evidence for the difference in exploration between the low and high serum groups was not robust (p=0.066). Furthermore, it has been suggested that increased striatal dopamine is also related to tendency for stochastic, undirected exploration (Frank, Doll, Oas-Terpstra, & Moreno, 2009; Gershman & Uchida, 2019), arising due to overall uncertainty across available options (Gershman & Uchida, 2019) or through increasing the opportunity cost of choosing the wrong option (Cools, 2016; Niv et al., 2006). This suggests that the same biological signature that leads to increased cognitive effort expenditure also promotes choice exploration. In line with this, both prior studies that investigated the effect of increasing dopamine availability with L-DOPA on model-based/model-free behaviour observed increase choice exploration as well as increased model-based behaviour (although in one it was only present in individuals with a higher working memory capacity) (Kroemer et al., 2019; Wunderlich et al., 2012).

The lack of a pronounced effect of naltrexone on model-based/model-free behaviour is unlikely to be due to issues of dosage or timing, since 50 mg of naltrexone lead to a ~90% of μ receptor occupancy even 48 hours post intake (M. C. Lee et al., 1988; Trøstheim, Eikemo, Haaker, Frost, & Leknes, 2022), with μ receptors being the primary (but not only) opioid receptor type that the drug binds to. One possibility is that opioid receptors are not crucial for model-free learning required in this task. This is in opposition to previous studies showing that acute administration of naltrexone, comparably to amisulpride, causes a reduction of cue responsivity and reward impulsivity (S. C. Weber et al., 2016), decreases effort to obtain immediate primary rewards (Korb et al., 2020), and decreases the wanting of rewards (Soutschek, Weber, Kahnt, Quednow, & Tobler, 2021). Similarly, preclinical studies in rats show that naltrexone and naloxone, another opioid antagonist, decreased sucrose reinforced place preference (Ågmo, Galvan, & Talamantes, 1995; Delamater, Sclafani, & Bodnar, 2000). Furthermore, both naltrexone and naloxone are used to reduce craving in patients with substance use disorder (Quednow & Herdener, 2016). However, there is also evidence that opioid receptors are important for goal-directed behaviour. In particular, a study in rats showed that opioid receptor blockade leads to decreased sensitivity to reward value and accelerated habitual control of actions (Wassum, Cely, Maidment, & Balleine, 2009), and in a recent neurochemical imaging study in humans μ opioid receptor availability correlated with goal-directed behaviour in a loss-only version of the two-step task (Voon et al., 2020). In fact, opioid agonists (but not antagonists) can in some cases lead to increased performance in cognitive control tasks (van Steenbergen et al., 2019). One explanation of our results is that naltrexone simply reduced the intrinsic value of reward and therefore decreased the

motivation to exert cognitive effort, which would be in line with the observed flattening of both positive and negative affect after naltrexone administration.

An important limitation of the experimental approach used here is that it rests on the assumption that there are only two ways of learning, which moreover can be distinguished clearly. This assumption has been questioned in the reinforcement learning literature (Daw, 2018; Feher da Silva & Hare, 2020), and mirrors the scepticism of the two-system division of decision making in cognitive psychology (Melnikoff & Bargh, 2018). Although we made sure that participants understood what the task rules were and how they could maximize their gains, we cannot exclude the possibility that participants searched for alternative models to obtain rewards, such as different pressing patterns or simply favouring one stimulus over the other (Feher da Silva & Hare, 2020). In fact, the increased exploration parameter in the amisulpride group could be due to participants' increased exploration of this model space. Importantly, it should also be acknowledged that the behavioural setup in our study does not allow us to draw definite conclusions about the mechanisms that mediate amisulpride's effects on model-based or model-free behaviour. For example, it is not clear whether amisulpride increases the perceived benefit of applying cognitive control, or whether it increases the participant's ability to do so through various possible complementary processes, such as goal maintenance or planning abilities. Future studies should further elucidate the mechanistic contributions of dopamine receptors to the distinct coding and utilisation of task relevant representations (Langdon, Sharpe, Schoenbaum, & Niv, 2018; Stalnaker et al., 2019).

One of the strengths of our design is a baseline measure, and the fact that the participants were all introduced to the task under no administration, thus avoiding potential effects of the treatment on task training. Although this design allowed to reduce between-subjects variability, we cannot completely exclude order effects. Although unlikely, it is possible that the effects of the treatment that we observe come indirectly from the effects of the two drugs on either skill transfer from the previous session, or simply on the effect of the drugs on the part of the experiment that preceded the task. For instance, participants under amisulpride could be less tired from other tasks and therefore more willing to exert effort in the task presented here. Speaking against this is the observation that we found no differences in mood between amisulpride and placebo regardless of low or high serum levels.

We note that opioid and dopamine systems are not the only neurotransmitter systems that have been implicated in the model-based/model-free tread-off. Most notably, depletion of the serotonin precursor tryptophane biases behaviour towards habitual and away from goal-directed (Worbe, Savulich, de Wit, Fernandez-Egea, & Robbins, 2015), suggesting that serotonin supports goal-directed behaviour. This claim is also in line with a recent neurochemical imaging study that showed that serotonin terminal density in the striatum correlated with higher goal-directed behaviour in a version of the two-step task (Voon et al., 2020) and decreasing prefrontal serotonin levels has been shown to promote drug seeking behaviour (Pelloux, Dilleen, Economidou, Theobald, & Everitt, 2012). In light of this, we also note that amisulpride also blocks serotonin receptors, albeit with lower affinity, which should be kept in mind when considering the findings of our study. In conclusion, we provide a first comparison of the contributions of the dopamine D2 and opioid receptors to arbitration between the model-based and model-free systems. Our results suggest that D2 dopamine antagonists might promote goal-directed behaviour when alternative habitual choices are available, already in low doses, while opioid blockade does not have such an effect on model-based behaviour. These findings are a step forward in understanding how neuromodulators control the arbitration between habitual and goal-directed decision-making systems, which can on the long run be crucial for developing targeted pharmacological treatments for addiction and other disorders of compulsivity.

Methods

Procedure

The study took place in the Department of Psychiatry and Psychotherapy at the Medical University of Vienna. After initial online screening, the participants were invited for a first visit, where

they underwent a physical examination (electrocardiogram, hemogram) and a psychiatric screening before being subjected to the two-step task that consisted of a 25 - minute automated training period followed by 200 trials of the task (see Figure 1 for the study outline). Participants eligible for the study were invited back to the clinic approximately 3 weeks after their first visit.

In their second visit participants received either placebo, 400 mg of amisulpride, or 50 mg of naltrexone. After a waiting period of 3 h, participants solved two additional tasks (explained elsewhere (Korb et al., 2020)) before performing the two-step task (on average 4 h 36 min, SD = 22 min, after the pill intake) as well as a working memory task. Waiting times and the doses of drugs were chosen based on previous pharmacological studies with the same compounds (S. C. Weber et al., 2016). To ensure comparable absorption, participants were asked to come with an empty stomach and received a standardized meal before pill intake. Participants filled out a questionnaire assessing mood and side effects right after pill intake and 3 hours later. There were no profound differences in side effects across the three drug groups, apart from a trend level effect of amisulpride on tiredness (Figure 1 – figure supplement 1). Blood plasma levels confirmed amisulpride levels above 212.6 μg/l (mean (sd) = 548.8 (96.9)), in all participants of the amisulpride group. For 7 subjects serum levels were not possible to calculate due to technical issues with blood samples. For 18 participants the value of amisulpride was at maximum (>604 μg/l), leading to a highly skewed distribution, with no variance in the upper 50% quantile. Because of this we dichotomized the data into a categorical variable, whereby the high serum group consisted of participants with serum levels above 604 μg/l (n = 18) and the low serum group had levels between 212.6 μg/l and 602.5 μg/l (n = 14).

Participants

Data was collected from 120 volunteers and constituted a subset of volunteers involved in a larger study (Korb et al., 2020). All participants were assigned an initial screening code and a subject ID number. For six participants the task data from the first session were lost due to failures in assigning screening codes from the first session to subject ID numbers in the second. For two participants the task data from the second session were lost owing to technical issues, resulting in 112 participants included in the main analysis (see Supplementary File 1l for exact subject numbers for each analysis). All participants had no history of drug abuse or other psychiatric disorders and were matched in age, sex, and BMI (Supplementary File 1j). The study was approved by the Ethical Committee of the Medical University of Vienna (EK N. 1393/2017) and in line with the Declaration of Helsinki ("World Medical Association Declaration of Helsinki," 2013). Participants received a monetary compensation of 90€ plus the extra money earned in the task.

Genotypic analysis

Peripheral blood was collected by lancet and stored on Whatman FTA micro cards (Sigma-Aldrich). DNA was extracted using the QIAamp DNA Mini kit (Qiagen, Hilden, Germany). The VNTR polymorphism in the *DAT1* gene was investigated by PCR with 5'-fluorescent-dye-labeled forward primer and automated detection of PCR products by capillary electrophoresis (details of the procedure provided in the supplement). The single base primer extension (SBE) method also known as SNaPshot minisequencing was applied for the typing of single nucleotide polymorphism (SNP) variants (details provided in supplement). Accordingly, five informative SNPs [*DRD2/ANKK1 Taq1A* (*rs1800497*), *BDNF* (*rs6265*), *CDH13* (*rs3784943*), *OPRM1* (*rs1799971*) and PPP1R1B/DARPP-32 (*rs907094*)] were analysed simultaneously applying a multiplex strategy for PCR and SNaPshot minisequencing of purified PCR products. Typing of Val158Met variants (*rs4680*) in the *COMT* gene was carried out separately, applying a singleplex approach for PCR and SNaPshot. The *OPRM1* and *BDNF* SNPs were not included in the analysis in this task. Genetic data from three participants were lost. This led to the group distributions depicted in Supplementary File 1k. For details on the analysis see Supplementary Note 2).

Task design

In the task, participants made an initial choice at stage one which took them to one of the two possible states ('planets') in stage two. There were two possible states in the first stage, each featuring a pair of spaceships. Each of the spaceships flew deterministically to one of the two planets where they encountered an alien who gave them either positive or negative points. The points in each planet changed according to a Gaussian random walk in the interval [-4, 5], rounded to whole numbers. Importantly, since both first-stage states led to the same two possible second-stage states, participants could transfer knowledge from the pair of spaceships in one first-stage state to the pair in the other. According to the conventional definition adopted here, a completely model-free agent relies entirely on its direct experience and will choose a spaceship based only on its reward history in previous trials featuring the same first-stage state, regardless of their experience with the other pair of spaceships. A completely model-based agent however will use the causal structure of the task to update the value of the spaceships based on which planet they fly to. This means that knowing that spaceship A and C fly to the same planet (but appear in different first step states) enables the model-based agent to learn about spaceship C by experiencing the outcome of its choice of spaceship A.

Participants had 2000 ms to choose in the first step and then again 2000 ms to press the space bar once they had encountered the alien. Each acquired point was translated to 4 Eurocents and added to the overall compensation of the participant at the end of the experiment. The story and a thorough training session were employed to increase the comprehension of the task, as done before (Decker, Otto, Daw, & Hartley, 2016). To avoid all participants behaving in a model-based way (Feher da Silva & Hare, 2020) we increased the difficulty of the task by dynamically changing the drift of the reward-determining Gaussian walks. The Gaussian random walks therefore had various drift rates (0.5, 1, and 2). We generated one trajectory and shuffled it around to create two different testing sessions of similar difficulty. Each participant first went through an automated rigorous explanation of the task followed by 25 practice trials before completing 200 trials of the task.

Behavioural analysis

Table 3. Prior distributions for the behavioural analysis

Standard Deviations	$\boldsymbol{\sigma \sim HalfCauchy(0,2)}$
Regression Coefficients	$\beta \sim N(0,3)$
Intercept	$\beta_0 \sim Student(3,0,10)$
Prior for the correlation matrix	$R \sim LKJcorr(2)$

To regularize our inference, we used a hierarchical Bayesian approach in all our analyses. We report estimates in parameter estimation space and indicate the precision of our estimates with credible intervals (Kruschke, 2014; L. Zhang et al., 2020). We also report the proportion of the credible interval that is above or below zero. Model-agnostic analysis of behaviour focused on predicting the probability of staying with the previous choice (choose the spaceship that flies to the same planet as in the previous trial). Behaviour was analysed using the *brms* package in R (Bürkner, 2017) which employs the probabilistic programming language Stan(Carpenter et al., 2017). We fit a binomial model that predicted staying with the previous choice from reward obtained in the previous trial, modulated by the previous state (same or different), drug treatment, and session. The effects of session, previous state, previous points and all interactions between them were drawn from a multivariate normal distribution (were considered as correlated random effects). We report 95% credible intervals of estimates on the log-odds scale. Parameters were estimated with 4 chains, each 3000 iterations (1000 warmup), with priors listed in Table 1.

The number of participants used in the analysis depend on whether both sessions were included and whether genetic data were used (see Supplementary File 1l for an overview). Analysis with amisulpride blood serum levels included a categorial variable for the serum. A categorical variable was

chosen due to the skewed distribution of serum levels. To relate behaviour to the parameters of the computational model we extracted the random slopes for each participant for the second session for both the effects of previous points on staying behaviour in trials with the same first stage state as in the previous trial, as well as the random slopes for the different vs. same first state trials. We then standardized all slopes and the computational parameters and used two separate linear models to predict both slopes from the three computational parameters.

Computational Models

We defined three computational models that define the subjective values Q of participants' action a, in trial t, with a first stage state s, as the weighted average of model-based (Q_{MB}) and model-free (Q_{MF}) subjective values.

$$Q(a,s,t) = \omega Q_{MB}(\mathrm{a},\mathrm{s},t)) + (1-\omega) Q_{MF}(\mathrm{a},\mathrm{s},t)$$

where ω is the weighting parameter (larger influence on choice of model-based values is indicated by ω being close to 1 and model-free control by ω being close to 0). The subjective values are then mapped on to probability of choosing a, and not a', with the soft-max transformation:

$$P(a,s,t) = \frac{e^{\eta * Q(\mathrm{a},\mathrm{s},t))}}{e^{\eta * Q(\mathrm{a},\mathrm{s},t)} + e^{\eta * (\mathrm{Q}(\mathrm{a}',\mathrm{s},t))}}$$

where η is the *inverse temperature* parameter, that determines the stochasticity of choices and the exploration-exploitation trade-off. In the models that include stickiness parameters, the value function for action a was extended as follows:

$$Q(a,s,t) = \omega Q_{MB}(\mathrm{a},\mathrm{s},t)) + (1-\omega) Q_{MF}(\mathrm{a},\mathrm{s},t) + \rho * resp(a) + \pi * stim(a)$$

where the indicator functions were defined as being 1 for the action that was the same as in the previous trial ($resp(a)$) or the stimulus that was chose in the previous trial ($stim(a)$. The ρ and π parameters therefore determine the degree to which previous actions (or stimuli) tended to be repeated (Kool et al., 2016).

Deterministic learning model (M1)

Because the outcomes at the second stage are determined by Gaussian random walks, the objectively best prediction by the agent is the last encountered outcome. We defined a model (M1), where the model-free subjective value of each of the spaceships is the last encountered outcome following the choice of that spaceship, and the model-based subjective value of each spaceship is the subjective value of the last encountered outcome following the planet that this spaceship flies to. This means that the update equation of the model-free agent after receiving outcome $r(s_2,t)$ following an action a, in trial t, with first state s_1, is defined as

$$Q_{MF}(a,s_1,t) = r(s_2,t).$$

In the same trial, the value of the for the unchosen actions across both first level states are shrunk towards 0 with forgetting parameter $\gamma \in [0,1]$:

$$Q_{MF}(a',s_1,t) = (1-\gamma)\, Q_{MF}(a',s_1,t-1),$$
$$Q_{MF}(a,s_1{}',t) = (1-\gamma)\, Q_{MF}(a,s_1',t-1),$$
$$Q_{MF}(a',s_1',t) = (1-\gamma)\, Q_{MF}(a',s_1',t-1),$$

where a' is the unchosen action and $s_1{}'$ is the unencountered state. The model-based agent, on the other hand, after receiving outcome $r(t)$ following an action a, in trial t, with first state s_1, updates not only the experienced first stage state, but also the unencountered first stage which would have led to the same outcome:

$$Q_{MB}(a,s_1,t) = Q_2(s_2,t) = r(s_2,t)\,,$$

$$Q_{MB}(a, s_1', t) = Q_2(s_2, t) = r(s_2, t)\,,$$

where, $Q_2(s_2, t)$ is the subjective value of the second stage state that action a deterministically leads to.

Dual-system reinforcement learning models (M2 and M3)

We compared our model M1 to two versions of a dual-system reinforcement learning model inspired by Kool et al (2016, 2017). In these models, the model-free agent learns the subjective values of spaceships and planets through a temporal difference-learning algorithm(Daw et al., 2011; Sutton & Barto, 2017). The model-free agent was defined by 3 free parameters: the learning rate at the first stage (α_1), and the second stage (α_2), where the eligibility trace (λ) determines the degree to which the outcome at the second stage retrospectively transfers to the first stage. In simple terms, the model increases (or decreases) the subjective value of an action at stage 1 proportionally to how positively (or negatively) surprising the outcome was, but discounted by the learning rate that describes the contributions of previous outcomes of that specific action. Conversely, the model-based agent is aware of the structure of the task. Specifically, in trial t, with first stage state s_1 we define the model-free subjective value of action a as

$$Q_{MF}(a, s_1, t) = Q_{MF}(a, s_1, t) + \alpha_1 \delta_1$$

$$\delta_1 = Q_2(s_2, t) - Q_{MF}(s_1, a, t)$$

and the model-based subjective values of action a as

$$Q_{MB}(s_1, a, t) = T(s_1, a) * Q_2(s_2, t)$$

where $Q_2(s_2, t)$ is the subjective value of the second stage s_2 at time t and $T(s_1, a)$ is the transition matrix from first to second stage states. The agents learn the values of the planets, Q_2, by

$$Q_2(s_2) = Q_2(s_2) + \alpha_2 \delta_2$$

$$\delta_2 = r - Q_2(s_2)$$

$$Q_{MF}(s_1, a) = Q_{MF}(s_1, a) + \lambda \alpha_1 \delta_2$$

The learning rates at the two stages were either different (M2), or the same (M3).

Model Estimation

We estimated the model parameters for both sessions in one hierarchical (multilevel) Bayesian model. This approach pools information across different levels (drug groups, and participants) and thus leads to more stable individual parameter estimates (Ahn et al., 2017), reduces overfitting (McElreath, 2018), and enables us to estimate in one model both individual and group level parameters as well as differences between sessions (Lengersdorff, Wagner, & Lamm, 2020). Models were implemented in Stan (Carpenter et al., 2017) using R as the interface. Stan uses a Markov Chain Monte Carlo sampling method to describe posterior distributions of model parameters. We ran each candidate model with four independent chains and 3000 iterations (1000 warm-up). Convergence of sampling chains was estimated through the Gelman-Rubin $\hat{R}$ statistic (Gelman & Rubin, 1992), whereby we considered $\hat{R}$ values smaller than 1.01 as acceptable.

For all subject level parameters (e.g., ω) we drew both the baseline (ω_0) and the session difference ($\Delta\omega$), from a multivariate Gaussian prior. Specifically, for model M1:

$$\begin{pmatrix} \omega_0' \\ \gamma_0' \\ \eta_0' \\ \Delta\omega' \\ \Delta\gamma' \\ \Delta\eta' \end{pmatrix} \sim MVNormal \begin{pmatrix} \mu_{\omega_0'} \\ \mu_{\gamma_0'} \\ \mu_{\eta_0'} \\ \mu_{\Delta\omega'} \\ \mu_{\Delta\gamma'} \\ \mu_{\Delta\eta'} \end{pmatrix}, \; S \Bigg)$$

where S is the covariance matrix, that was factored into a diagonal matrix with standard deviations and the correlation matrix R (Bürkner, 2017; McElreath, 2018). The prime denotes the parameters in estimation space. The parameters that were fully constrained (i.e., ω, α_1, α_2, λ, γ) were estimated in the inverse probit space and the parameters that only had a lower bound (i.e., η) were estimated in log space. The hyper-priors for all group-level means were weakly informative, $\mu_{\omega'} \sim N(0,1)$, the prior for group-level standard deviations were $\sigma_{\omega'} \sim HalfNormal(0,1)$, and the prior for the correlation matrix was $R \sim LKJcorr(2)$. The parameters in the second session were therefore defined by their estimated baseline and difference between the sessions. For instance, the model-based weight ω in the second session for participant i was defined as:

$$\omega(i) = \phi(\omega_0'(i) + \Delta\omega'(i))$$

where ϕ is the cumulative distribution function of the standard normal distribution (inverse of probit). Effect sizes are calculated by normalizing the relevant regression coefficients by the pooled standard deviation (square root of the sum of all relevant variance components (Hedges, 2007; Nalborczyk et al., 2019)). In models where there are no random effects, this reduces to a Cohen's d. For example, the effect sizes for drug effects on ω were calculated by dividing the estimated difference between group means by the square root of the sum of the variance of both the baseline ($\sigma^2_{\omega_0'}$) and the session difference ($\sigma^2_{\Delta\omega'}$). We used the same procedure in all computational models.

Model Comparison and Validation

We used the trial-based Leave-One-Out Information Criterion (LOOIC) to compare the three models using the loo package in R (Vehtari et al., 2017). The LOOIC estimates out-of-sample predictive accuracy of each trial and is more informative than simpler point-estimate information criterions used commonly (such as the Akaike information criterion). Lower LOOIC scores indicate better prediction accuracy out of sample. Additionally, we compared models with bootstrapped pseudo Bayesian Model Average relative weights of the models that reflect the posterior probability of each model given the data (Y. Yao, Vehtari, Simpson, & Gelman, 2018). To validate the novel model M1, we first used the posterior means of the estimated parameters to simulate behaviour in both sessions. To see how well we can retrieve the model parameters we reran the parameter estimation on the simulated behaviour with the same model (Figure 4 – figure supplement 1, 1d). We also ran the same analysis of staying behaviour on this synthetic dataset and reproduced the behavioural plots from Figure 3 (compare to Figure 4 – figure supplement 1, 1e). We used the same logistic hierarchical Bayesian model to statistically evaluate the crucial aspects of the behavioural analysis (Figure 4 – figure supplement 1, 1e, compare to Figure 3c). To get the posterior predictive accuracy of the model we predicted the choice on each trial for each participant for 8000 samples drawn from the posterior distribution and then calculated the average accuracy for each participant.

Estimation of the effects of the pharmacological treatment

To statistically evaluate the effect of our treatment on all the parameters in the model, we included two regression terms when defining the group-level means of the difference between sessions:

$$\begin{pmatrix} \omega_0' \\ \gamma_0' \\ \eta_0' \\ \Delta\omega' \\ \Delta\gamma' \\ \Delta\eta' \end{pmatrix} \sim MVNormal \left(\begin{matrix} \mu_{\omega_0'} \\ \mu_{\gamma_0'} \\ \mu_{\eta_0'} \\ \mu_{\Delta\omega'} + \beta^{\omega}_{Ami} * X_{Ami} + \beta^{\omega}_{Nal} * X_{Nal'} \\ \mu_{\Delta\gamma'} + \beta^{\gamma}_{Ami} * X_{Ami} + \beta^{\gamma}_{Nal} * X_{Nal} \\ \mu_{\Delta\eta'} + \beta^{\eta}_{Ami} * X_{Ami} + \beta^{\eta}_{Nal} * X_{Nal} \end{matrix}, \; S \right)$$

where β^*_{Ami} and β^*_{Nal} are coefficients drawn from prior distribution $N(0,1.5)$ and X_{Ami} and X_{Nal} are dummy variables for the two drugs. To see which of the parameters are affected by the drug treatment we included group-level effects on all parameters. This model was extended in the second

step to include a group-level categorical variable for high/low serum levels of amisulpride (coded as 1 for high serum level and 0 otherwise). In the third step, the model also included the subject-level stickiness parameters and all corresponding group level effects.

To estimate the effects of mood, working memory performance, sex, weight and age, we ran another model predicting estimated differences in ω and η across sessions from those variables. Similarly, in another model, the session differences were predicted from drug variables and their interactions with the four genotype variables. Full model definitions and outputs is provided in supplementary material.

Reading Span Task

We used an automated version of the Reading Span Task, where in each block participants saw 2 to 6 words serially presented that they had to recall by the end of the trial in any order. The words were interlaced with sentences that participants were instructed to judge as either making sense or not (Klaus & Schriefers, 2016). Participants played 15 blocks. Correctly recalled items were calculated as a proportion within the block and then averaged across blocks. The effect of the drug treatment was calculated with a Bayesian linear model where the mean score was predicted by drug treatment.

Acknowledgements

The study was supported by the Vienna Science and Technology Fund (WWTF) with a grant (CS15- 003) awarded to Giorgia Silani and Christoph Eisenegger and a grant (VRG13-007) awarded to Christoph Eisenegger and Claus Lamm. We thank Prof. Boris Quednow for his advice on the study design. This work would not be possible were it not for the students who have carried out the data collection: Mani Erfanian Abdoust, Anne Franziska Braun, Raimund Bühler, Lena Drost, Manuel Czornik, Lisa Hollerith, Berit Hansen, Luise Huybrechts, Merit Pruin, Vera Ritter, Frederic Schwetz, Conrad Seewald, Carolin Waleew, Luca Wiltgen, Stephan Zillmer. We thank Catherine Hartley for sending us her version of the task implemented in matlab, and for allowing us to use the stimuli. We are grateful to Wouter Kool for making his version of the task implemented in JsPsych freely available on github. We also thank Lei Zhang and Lukas Lengersdorff for helpful discussions on the statistical procedures used in the manuscript. The manuscript has significantly improved from the constructive and insightful remarks of the three reviewers.

Supplementary Material

Figure supplements

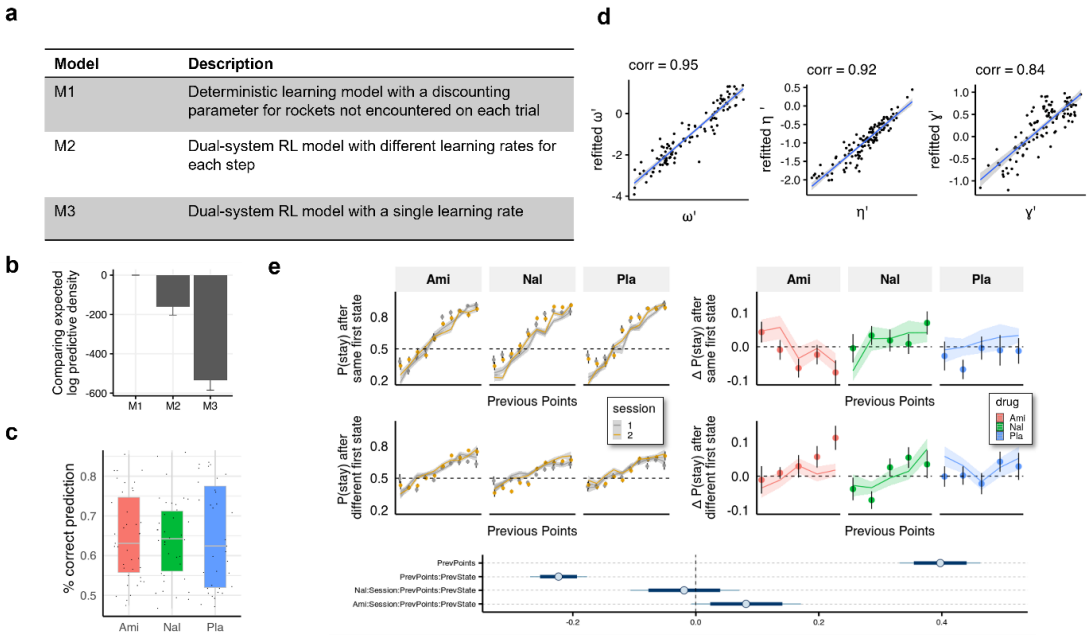


Figure 4 – figure supplement 1. Computational modelling. a, We compared our model to two dual-system reinforcement learning (RL) models used previously with this task(Kool et al., 2016, 2017). In model M1 both model-free and model-based agents remember only the last outcome of a choice of spaceships, but only the model-based agent is aware of the relation between first-stage states. Models M2 and M3 are reinforcement learning-based models adapted from Kool et al (2016) with either separate (M2) or equal (M3) first and second stage learning rates. **b**, Models were compared using the leave-one-out information criterion (looic), where lower values indicate better out of sample trial-by-trial predictive performance. Refer to the Methods for weights from Bayesian Model Averaging of all models, indicating the posterior probability of each model given the data. Model M1 performs best in both metrics. **c**, For the winning model (M1) we calculated the posterior predictive accuracy averaged within participants and found that the mean (SD) posterior prediction accuracy was 65% (11%). **d,** Parameter recovery. **e**, Simulated data plotted to visually and statistically investigate whether the model captures the crucial aspects of behaviour and group differences (compare to Figure 3).

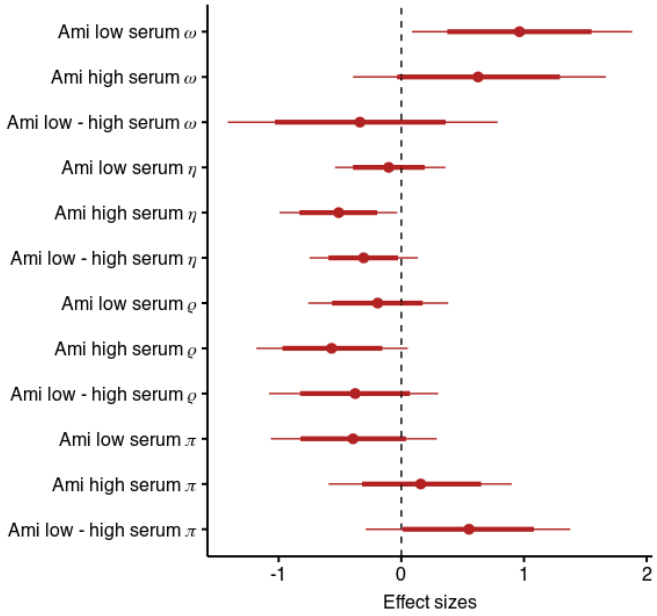


Figure 4 - figure supplement 2. Results of the model including stickiness parameters. Posterior distribution of effect sizes of group level effects on model parameters in a model including stickiness parameters with means, 95% and 80% intervals.

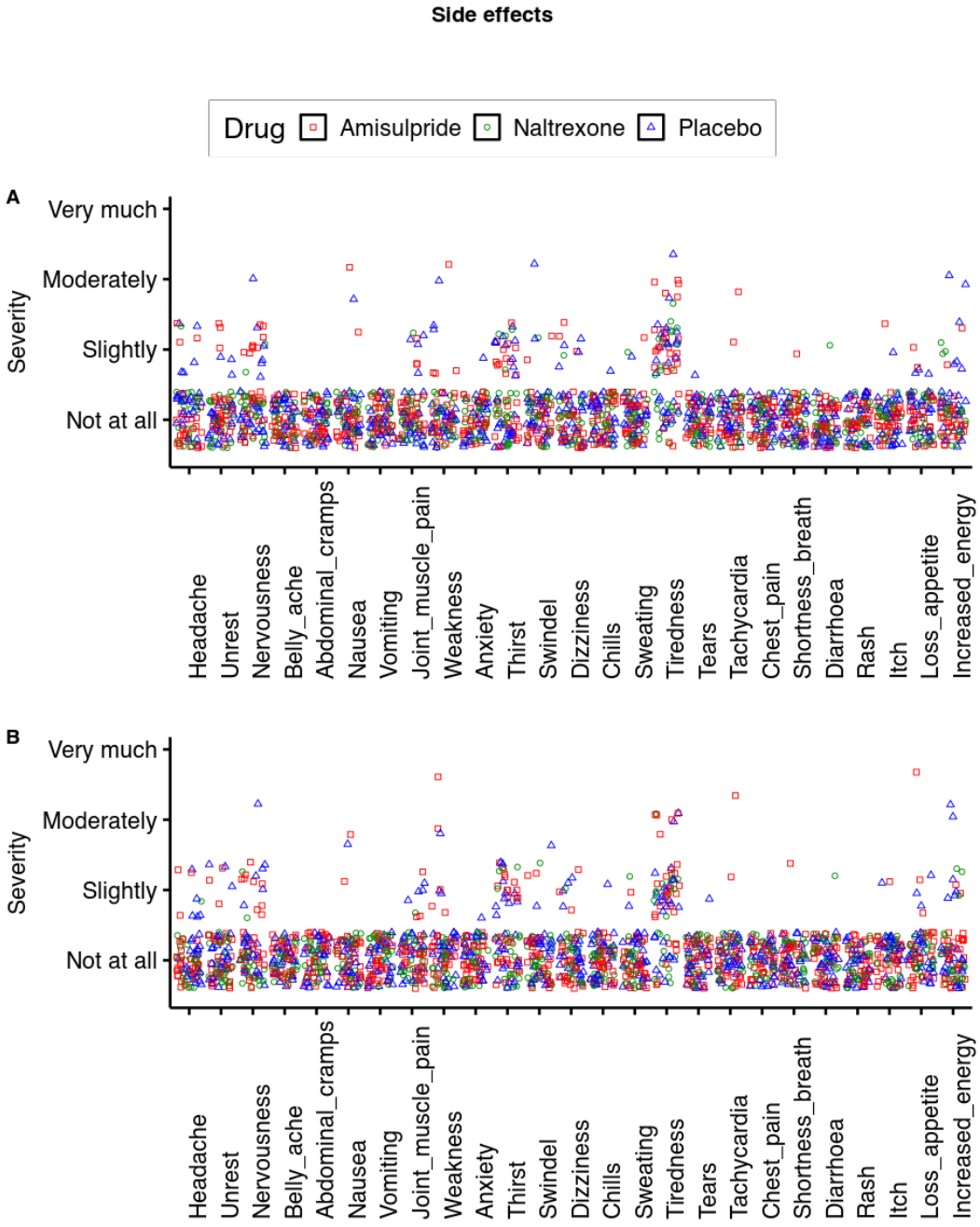


Figure 1 – figure supplement 1. Side effects. On most measures most participants rated the severity of the side effect as low as possible. The only measures where more than 50% of data points were above 1 was tiredness. Here, the effect of amisulpride was at trend level (b = 0.20, p = 0.06) and not significant for naltrexone (b = -0.15, p = 0.17).

Supplementary Files

Stay ~ session * prev_points * prev_state_diff * Ami + session * prev_points * prev_state_diff * Nal +(session * prev_state_diff * prev_points | ID)	Estimate	SE	Q2.5	Q97.5
Intercept	0.7	0.0	0.57	0.84
session	-0.01	0.1	-0.21	0.19
prev_points	0.43	0.0	0.34	0.52
prev_state	-0.53	0.1	-0.73	-0.34
Ami	-0.08	0.0	-0.27	0.1
Nal	-0.12	0.1	-0.32	0.07
session:prev_points	0.08	0.0	-0.02	0.18
session:prev_state	0.04	0.1	-0.21	0.31
prev_points:prev_state	-0.25	0.0	-0.34	-0.16
session:Ami	-0.08	0.1	-0.35	0.2
prev_points:Ami	0.04	0.0	-0.09	0.16
prev_state:Ami	0.04	0.1	-0.22	0.3
session:Nal	0.17	0.1	-0.11	0.44
prev_points:Nal	-0.05	0.0	-0.18	0.07
prev_state:Nal	0.01	0.1	-0.26	0.27
session:prev_points:prev_state	-0.03	0.0	-0.13	0.07
session:prev_points:Ami	-0.14	0.0	-0.27	-0.01
session:prev_state:Ami	0.2	0.1	-0.17	0.55
prev_points:prev_state:Ami	-0.03	0.0	-0.16	0.1
session:prev_points:Nal	0.02	0.0	-0.12	0.15
session:prev_state:Nal	-0.24	0.1	-0.6	0.13
prev_points:prev_state:Nal	0.01	0.0	-0.11	0.14
session:prev_points:prev_state:Ami	0.18	0.0	0.04	0.32
session:prev_points:prev_state:Nal	0.02	0.0	-0.13	0.16

Supplementary File 1a. Estimates and CIs of fixed effects of the Bayesian logistic regression predicting staying behaviour. Q2.5 and Q97.5 are the 2.5% and 97.5% quantiles of the posterior parameter distribution. For details on how the posterior distributions were calculated refer to the code online (Ami = Amisulpride, Nal = Naltrexone).

PosteriorPredictiveAccuracy ~ Ami + Nal	Estimate	Q2.5	Q97.5
Intercept	0.69	0.51	0.87
Ami	-0.05	-0.3	0.2
Nal	-0.07	-0.32	0.17

Supplementary File 1b.Results of a Bayesian logistic linear model predicting percentage correct from drug groups

	PANAS pos T1	PANAS neg T1	PANAS pos T2	PANAS neg T2
Placebo	29.5 (6.4)	11.3 (1.6)	27 (7.0)	10.5 (0.9)
Amisulpride	30.5 (5.6)	12.2 (3.3)	26.7 (6.2)	10.9 (2.9)
Naltrexone	30.1 (6.6)	13.9 (7.0)	25.4 (7.5)	11.4 (4.4)

Supplementary File 1c.Mood in mean and standard deviation at time of pill intake (T1) and 3 hours later (T2).

$\Delta PANAS_{pos} \sim$ ami + serum_ami + nal	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.2	0.16	-0.12	0.52
Ami	-0.2	0.29	-0.76	0.36
serum_ami_high	0.07	0.33	-0.59	0.71
Nal	-0.41	0.22	-0.85	0.03

Supplementary File 1d. Drug effects on differences in positive PANAS scales (centralized) between sessions. Drug variables coded as follows: *nal* is as dummy variable for naltrexone (1 for naltrexone, 0 otherwise), *ami* is a dummy variable for amisulpride, and *serum_ami* is a dummy variable for serum (1 only in the high serum group, and 0 otherwise).

$\Delta PANAS_{neg} \sim$ ami + serum_ami + nal	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.21	0.17	-0.13	0.54
Ami	-0.1	0.3	-0.69	0.49
serum_ami_high	-0.02	0.35	-0.73	0.65
Nal	-0.53	0.24	-1	-0.06

Supplementary File 1e. Drug effects on differences in negative PANAS scales (centralized) between sessions. Drug variables coded as before.

$\Delta\omega \sim$ (ami + serum_ami + nal)*(Weight + Sex + Age) + $PANAS_{neg}^{t0}$ +$\Delta PANAS_{neg}$ + $PANAS_{pos}^{t0}$ + $\Delta PANAS_{pos}$ + WM + Age	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-0.12	0.13	-0.39	0.14
Ami	1.00	0.22	0.56	1.42
serum_ami_high	-0.21	0.31	-0.8	0.4
Nal	0.23	0.19	-0.15	0.61
Sex1	-0.31	0.28	-0.84	0.24
Weight_s	0.15	0.14	-0.14	0.43
Age_s	0	0.11	-0.21	0.22
wm_s	0.05	0.07	-0.09	0.19
pos_mood_diff_s	-0.14	0.07	-0.29	0
neg_mood_diff_s	-0.14	0.18	-0.5	0.22
PANAS_1_negative_s	-0.09	0.12	-0.33	0.15
PANAS_1_positive_s	-0.11	0.09	-0.28	0.06
Ami:Sex1	-0.36	0.44	-1.23	0.49
Ami:Weight_s	0.11	0.27	-0.43	0.64
Ami:Age_s	-0.17	0.19	-0.55	0.22
serum_ami_high:Sex1	0.41	0.57	-0.71	1.51
serum_ami_high:Weight_s	-0.22	0.3	-0.81	0.39
serum_ami_high:Age_s	0.21	0.31	-0.39	0.8
Nal:Sex1	0.06	0.4	-0.74	0.82
Nal:Weight_s	-0.22	0.19	-0.6	0.17
Nal:Age_s	-0.11	0.18	-0.47	0.25

Supplementary File 1f. Drug effects on session differences in ω, including mood at baseline, difference in mood from baseline and working memory performance as covariates, as well as sex, age and weight as moderators of effects. Drug variables coded as before, all other dependent variables scaled and centralized.

$\Delta\eta \sim$ (ami + ami_serum + nal)*(Weight + Sex+ Age) + $PANAS_{neg}^{t0}$ +$\Delta PANAS_{neg}$ + $PANAS_{pos}^{t0}$ + $\Delta PANAS_{pos}$ + WM	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.19	0.12	-0.05	0.43
Ami	-0.32	0.2	-0.71	0.07
serum_ami_high	0.06	0.28	-0.5	0.6
Nal	-0.09	0.18	-0.44	0.27
Sex1	0.01	0.26	-0.5	0.52
Weight_s	0.04	0.14	-0.23	0.31
Age_s	-0.02	0.1	-0.22	0.18
wm_s	-0.06	0.06	-0.18	0.07
pos_mood_diff_s	-0.02	0.07	-0.15	0.11
neg_mood_diff_s	0.02	0.17	-0.31	0.35
PANAS_1_negative_s	-0.01	0.11	-0.23	0.21
PANAS_1_positive_s	-0.04	0.08	-0.2	0.11
Ami:Sex1	0.04	0.42	-0.78	0.85
Ami:Weight_s	0.17	0.25	-0.32	0.66
Ami:Age_s	-0.06	0.18	-0.42	0.28
serum_ami_high:Sex1	0.41	0.55	-0.66	1.48
serum_ami_high:Weight_s	-0.15	0.28	-0.69	0.42
serum_ami_high:Age_s	-0.09	0.29	-0.66	0.47
Nal:Sex1	-0.16	0.36	-0.88	0.57
Nal:Weight_s	-0.13	0.18	-0.48	0.24
Nal:Age_s	0.17	0.17	-0.16	0.5

Supplementary File 1g. Drug effects on session differences in η, including mood at baseline, difference in mood from baseline and working memory performance as covariates, as well as sex, age and weight as moderators of effects. Drug variables coded as before, all other dependent variables scaled and centralized.

Δω ~ (ami + ami_serum + nal)*(comt+ dat + darpp + ankk)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-0.15	0.11	-0.38	0.07
Ami	1.13	0.21	0.71	1.54
serum_ami_high	-0.24	0.25	-0.73	0.26
Nal	0.34	0.16	0.02	0.65
ankk_c1	0.19	0.2	-0.2	0.58
dat1_c1	-0.34	0.2	-0.74	0.06
comt_s	-0.01	0.1	-0.21	0.19
darpp_c1	0.09	0.2	-0.32	0.49
Ami:ankk_c1	0.43	0.41	-0.36	1.21
Ami:dat1_c1	-0.03	0.37	-0.76	0.69
Ami:comt_s	0.25	0.19	-0.1	0.61
Ami:darpp_c1	0.39	0.4	-0.39	1.17
serum_ami_high:ankk_c1	-0.26	0.46	-1.17	0.63
serum_ami_high:dat1_c1	0.52	0.41	-0.31	1.32
serum_ami_high:comt_s	-0.07	0.22	-0.48	0.36
serum_ami_high:darpp_c1	-0.38	0.44	-1.24	0.49
Nal:ankk_c1	-0.12	0.29	-0.69	0.44
Nal:dat1_c1	0.51	0.3	-0.08	1.11
Nal:comt_s	-0.12	0.16	-0.43	0.19
Nal:darpp_c1	-0.25	0.29	-0.83	0.34

Supplementary File 1h. Drug effects on session differences in ω, from genetic variables. Drug variables coded as before, all other dependent variables scaled and centralized.

Δη ~ (ami + ami_serum + nal)*(comt+ dat + darpp + ankk)	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.17	0.1	-0.03	0.38
Ami	-0.09	0.19	-0.48	0.29
serum_ami_high	-0.48	0.23	-0.95	-0.02
Nal	-0.02	0.15	-0.31	0.27
ankk_c1	0.21	0.19	-0.16	0.56
dat1_c1	0.16	0.2	-0.23	0.55
comt_s	-0.02	0.09	-0.21	0.16
darpp_c1	-0.09	0.19	-0.46	0.29
Ami:ankk_c1	0.38	0.38	-0.39	1.1
Ami:dat1_c1	0.06	0.35	-0.63	0.75
Ami:comt_s	-0.06	0.17	-0.4	0.26
Ami:darpp_c1	0.26	0.38	-0.48	0.99
serum_ami_high:ankk_c1	-0.9	0.43	-1.75	-0.03
serum_ami_high:dat1_c1	-0.44	0.39	-1.2	0.32
serum_ami_high:comt_s	0.15	0.2	-0.24	0.53
serum_ami_high:darpp_c1	-0.35	0.42	-1.14	0.49
Nal:ankk_c1	-0.19	0.27	-0.74	0.33
Nal:dat1_c1	-0.14	0.28	-0.68	0.42
Nal:comt_s	0.07	0.15	-0.23	0.37
Nal:darpp_c1	-0.09	0.27	-0.62	0.44

Supplementary File 1i. Drug effects on session differences in η, from genetic variables. Drug variables coded as before, all other dependent variables scaled and centralized.

	N	BMI, m (sd)	Age, m (sd)	Sex(m,f)
Placebo	39	21.7 (2.2)	23.2 (3.6)	12,27
Amisulpride	39	22.7 (2.6)	22.8 (2.8)	14,26
Naltrexone	40	23.0 (2.4)	23.2 (4.0)	13,26

Supplementary File 1j. Description of participants in terms of body mass index (BMI), age and sex with mean (m) and standard deviation (sd).

	COMT Val/Val, Val/Met, Met/Met	DAT1 10/10, other	ANKK, a1-, a1+	DARPP C/C+C/T, T/T
Placebo	9,14,12	23,12	20,15	13,22
Amisulpride	10,15,12	20,17	27,10	13,24
Naltrexone	6,20,11	20,17	24,13	13,24

Supplementary File 1k. Distribution of genotypes.

	BEHAVIOURAL ANALYSIS WITH BOTH SESSIONS	COMPUTATIONAL MODELLING WITH BOTH SESSIONS	COMPUTATIONAL MODELLING, SECOND SESSION ONLY	COMPUTATIONAL MODELLING WITH BOTH SESSIONS AND GENETIC DATA
PLACEBO	35	35	39	35
AMISULPRIDE	38	38	39	37
NALTREXONE	39	39	40	37

Supplementary File 1l. Number of participants per drug group used in analysis.

Supplementary Notes

Supplementary Note 1: Description of genotypes and how they relate to baseline dopamine function

Taq1a is a D2 dopamine receptor polymorphism with two allelic variants, A1 and A2. The A1 allele carriers (a1+) have reportedly higher dopamine synthesis rate(Laakso et al., 2005) that could potentially result in lower density of D2 receptors in the striatum(H. Ito et al., 2011; Jönsson et al., 1999). The *Dat1* VNTR is a polymorphism of the dopamine transporter (DAT) that is involved in reuptaking dopamine from the synaptic cleft back into the cell and thus a regulator of tonic dopamine levels and occurs with highest density in the striatum(Hiroshi Ito, Takahashi, Arakawa, Takano, & Suhara, 2008). The 9-repeat variant of the polymorphisms is associated with lower levels of DAT, and thus potentially higher extra-cellular dopamine levels(Fuke, 2001; Heinz et al., 2000). Based on the above both an effect on baseline tendency for model-based behaviour as well as an augmentation of the drug effects could be expected in the group carrying at least one A1 allele (a1+ group) and at least one the 9-repeat variant of the DAT1 polymorphism (9 repeats group). As in previous studies, we used the valine-to-methionine (Val/Met) substitution in the Val^{158}Met polymorphism of the *COMT* gene as a marker for prefrontal dopamine function(Doll, Hutchison, & Frank, 2011; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). COMT (or Catechol-O-methyl transferase) is an enzyme that

has an effect on dopamine extracellular levels through metabolizing dopamine especially in cortical areas of the brain(Männistö & Kaakkola, 1999). Participants with Met mutations have lower enzymatic activity of the *COMT* and have been associated with higher dopamine levels the prefrontal cortex(Slifstein et al., 2008). And lastly, we used the *DARPP-32* as a marker of striatal D1 receptor function. T homozygotes (T/T allele carriers) have putatively higher efficiency of striatal D1 receptors(Svenningsson et al., 2004) and have been related to model-free learning (Doll et al., 2016; Frank et al., 2009).

Supplementary Note 2: Details of Genotyping

I. Typing of the variable number tandem repeat (VNTR) polymorphism in the DAT1 gene

Peripheral blood was collected by lancet and stored on Whatman FTA micro cards (Sigma-Aldrich). DNA was extracted using the QIAamp DNA Mini kit (Qiagen, Hilden, Germany) and eluted in a final volume of 50 µL of buffer AE (Qiagen, Hilden, Germany). Human DNA concentration was determined with the Quantifiler HP Quantification Kit (AB) on the Applied Biosystems (AB) 7500 real-time PCR instrument (Thermo Fisher Scientific, Waltham, MA). Template DNA (10 ng per sample) was subjected to PCR in a total reaction volume of 25 µL consisting of 1 × GeneAmp PCR Buffer (AB), 0.25 mM each dNTP, 2.5 U AmpliTaq Gold Polymerase (AB) and target specific primers (details are provided in Table S1). The following thermal protocol was applied using a Veriti 96-Well Thermal Cycler (AB): 35 amplification cycles at 95 °C for 30 seconds, 55 °C for 1 minute, and 72 °C for 1 minute. Before the first cycle, an initial "hot start" denaturation (5 minutes at 95°C) was included, and the last cycle was followed by a final extension step at 72 °C for 45 minutes. Aliquots of PCR products were diluted with Hi-Di formamide (AB) mixed with internal lane standard LIZ 600 v.2 (AB) and separated on the ABI 3500 Genetic Analyzer applying standard conditions. The number of repeats predicted by the GeneMapper ID-X software (AB) was in full agreement to the number of repeats determined by direct sequencing of PCR products using the BigDye Terminator Sequencing Kit v3.1 (AB) in selected DNA samples.

Table S1. Primer set used for the typing of DAT1 VNTR repeat length polymorphisms

Marker	Location[a]	Primer sequence 5'-3' [b]	Dye [c]	Orientation	Conc.(nM)[d]
DAT1 VNTR	chr5:1393559-1394008(-)	TGTGGTGTAGGGAACGGCCTGAGA	6-FAM	forward	400
		TGTTGGTCTGCAGGCTGCCTGCAT		reverse	

[a] Chromosome number and genomic location of targeted sequence (orientation provided in brackets) according to UCSC version hg38 (http://genome.ucsc.edu/) is provided.

[b] The non-specific primer tail is underlined in Italics.

[c] 5' Fluorescein (6-FAM)-labeled forward primer was used.

[d] The final primer concentrations in the reaction mix is shown.

II. Typing of single nucleotide polymorphisms (SNPs) by SNaPshot minisequencing

Five informative SNPs [ANKK1 (rs1800497), BDNF (rs6265), CDH13 (rs3784943), OPRM1 (rs1799971) and PPP1R1B (rs907094)] were typed simultaneously applying a multiplex strategy for PCR and SNaPshot minisequencing of purified PCR products. Typing of Val158Met variants (rs4680) in the COMT gene was carried out separately, applying a singleplex approach for PCR and SNaPshot.

Step 1: PCR and purification

1.1 Singleplex PCR (COMT)

First, a 177 bp genomic fragment of the COMT gene harbouring the causative single nucleotide polymorphism (SNP rs4680) in its centre was targeted by PCR. The reaction mix contained 5 ng template DNA, 1 × GeneAmp PCR buffer (AB), 0.25 mM each dNTP, 2.5 units AmpliTaq Gold polymerase (AB) and specific primers (details provided in Table 2) in a total reaction volume of 25 µL. Thermal cycling was conducted using the Veriti cycler (AB) and the following conditions: 95 °C for 5 min; 35 cycles of 95 °C for 15 seconds, 59 °C for 30 seconds and 72 °C for 1 minute; final extension at 72 °C for 5 minutes. Excess of primers and unincorporated dNTPs were removed by adding 2 µL of ExoSAP-IT (Thermo Fisher Scientific) to each 5 µL PCR product. Reactions were incubated at 37 °C for 15 min followed by 80 °C for 15 min for enzyme deactivation.

1.2 Multiplex PCR (ANKK1, BDNF, CDH13, OPRM1, PPP1R1B)

Amplification by multiplex PCR was carried out in a total volume of 20 µL, containing 1x Phire Hot Start II PCR Master Mix (Thermo Fisher Scientific), PCR primers concentrations as specified in table 2 and template DNA (5 ng per sample). Thermal cycling was carried out on the Veriti thermal cycler (AB), and the following conditions: 98 °C pre-incubation step for 30 sec; 40 cycles of 98 °C denaturation for 10 sec, annealing at 62 °C for 30 sec and extension at 72 °C during 30 sec; followed by 1 min for final extension at 72 °C. PCR products were purified by ExoSAP-IT treatment.

Step 2: SNaPshot Minisequencing

2.1 Singleplex SNaPshot minisequencing (COMT)

Singleplex SNaPshot minisequencing of COMT(rs4680) was performed in a total volume of 10 µL containing 3 µL of purified PCR product, 5 µL SNaPshot Multiplex Ready Reaction mix (Thermo Fisher) and 2 µL of diluted minisequencing primer (pCOMT 2 µM; details see Table 3). The cycling conditions (25 cycles) using the Veriti thermal cycler (AB) were as follows: denaturation at 96 °C for 10 seconds, annealing at 50 °C for 5 seconds and extension at 60 °C for 30 seconds.

ExoSAP-IT treatment was again applied for the clean-up of the minisequencing reaction. 5 µl of purified reaction product was then mixed with 9.3 µL Hi-Di formamide (AB) and 0.2 µL of GeneScan-LIZ 120 internal size standard (AB). After a denaturing step for 5 min at 98 °C followed by cooling to 4 °C the fragments were separated on an ABI PRISM 310 Genetic Analyzer (AB) with POP4 polymer and analysed with GeneMapper v3.2 software. Calling of SNP variants based on minisequencing was in full agreement to results from direct sequencing of PCR products in selected DNA samples.

2.2 Multiplex SNaPshot (ANKK1, BDNF, CDH13, OPRM1, PPP1R1B)

Multiplex SNaPshot minisequencing was performed in a total volume of 10 µL containing 3 µL of purified PCR product, 5 µL SNaPshot Multiplex Ready Reaction mix (Thermo Fisher) and 2 µL multiplex primers mix (1 µM of pBDN; 2 µM of pANKK1, pCDH13, pOPRM1 and PPP1R1B; details provided in Table 3). Thermal cycling conditions (28 cycles) were as follows: denaturation at 96 °C for 10 seconds, annealing at 58 °C for 5 seconds and extension at 60 °C for 30 seconds. ExoSAP-IT clean-up and automated detection of minisequencing products by capillary electrophoresis was conducted

using the same conditions as for the singleplex SNaPshot (details provided in 2.1). Calling of SNP variants based on minisequencing was in full agreement to results from direct sequencing of PCR products in selected DNA samples.

Table S2. Panel of loci and primer sets used for PCR

Marker[a]	Location[b]	Primer sequence 5'-3'	Orientation	Conc.(nM)[c]
ANKK1	chr11:113400010-113400194(-)	GACATGATGCCCTGCTTTCG	forward	500
		CATCACGCAAATGTCCACGC	reverse	
BDNF	chr11:27658331-27658494(-)	CAAACATCCGAGGACAAGGT	forward	250
		CCGAACTTTCTGGTCCTCAT	reverse	
CDH13	chr16:83678696-83678883(+)	CAGTGCTGTGTTCCCAAATG	forward	500
		CTGCGAATCCACAATACCTGT	reverse	
COMT[a]	chr22:19963623-19963799(+)	GGGCCTACTGTGGCTACTCA	forward	400
		GCCCTTTTTCCAGGTCTGA	reverse	
OPRM1	chr6:154039591+154039751(+)	CCTTGGCGTACTCAAGTTGC	forward	500
		CGTGATCATGGAGGGACTG	reverse	
PPP1R1B	chr17:39634064-39634252(-)	CTCAGGTAGGGCTGAGTTCG	forward	500
		CCTGAAGGTCATCAGGCAGT	reverse	

[a] COMT primers only used in singleplex PCR; all other primers combined in a multiplex PCR.

[b] Chromosome number and genomic location of targeted sequence (orientation provided in brackets)

according to UCSC version hg38 (http://genome.ucsc.edu/).

[c] The final primer concentrations in the reaction mix is listed.

Table S3 Minisequencing primer information

Primer name[a]	SNP (Alleles) Location[b]	Primer sequence 5'-3'[c]	Primer binding site[d]
pCOMT	rs4680 (G/A) chr22:19963748	*(GATC)*$_4$GGATGGTGGATTTCGCTGGC	chr22:19963728-19963747(+)
pANKK1	rs1800497 (C/T) chr11:113400106	CCATCCTCAAAGTGCTGGTC	chr11:113400086-113400105(+)
pBDNF	rs6265 (G/A) chr11:27658369	*(GATC)*$_8$TCATTGGCTGACACTTTCGAACAC	chr11:27658370-27658393(-)
pCDH13	rs3784943 (G/A) chr16:83678830	*(GATC)*$_{10}$CCTACTTTGTCATCAGCACTGCTTT	chr16:83678805-83678829(+)
pOPRM1	rs1799971 (G/A) chr6:154039662	CGCATGGGTCGGACAGGT	chr6:154039663-154039680(-)
pPPP1R1B	rs907094 (C/T) chr17:39634118	*(GATC)*$_6$GTATACTCAAGGAGGACCCACAG	chr17:39634119-39634141(-)

[a] COMT primer was only used in singleplex SNaPshot; all other primers were combined in a multiplex SNaPshot.

[b] Chromosome number and genomic location of target single nucleotide polymorphism (SNP) according to UCSC version hg38 (http://genome.ucsc.edu/).

[c] The non-specific primer tail is underlined in Italics

[d] Chromosome number and genomic location of primer binding site (orientation provided in brackets) according to UCSC version hg38.

Chapter 5: General Discussion

The aim of this thesis was to advance our knowledge of the neurochemical and neurocomputational mechanisms of antipsychotic drugs. I presented three papers in which my co-authors and I outlined how some variance in psychotic experiences across the clinical severity spectrum could be explained by aberrant belief updating processes, which are in turn affected by D2 receptor antagonists in healthy volunteers.

Psychotic symptoms, such as holding on to bizarre delusional beliefs or experiencing anomalous sensations, are thought to result from misattributing relevance to random chance events, but it is not clear whether the core computational mechanisms that drive this are due to individuals 1) failing to adjust their beliefs about how reliable observations are, or 2) overestimating the likelihood of environmental changes. In Chapter 2, we present a computational model that describes how individuals learn in environments where both observational noise and the volatility of environmental events can change. Three parameters of the model representing noise uncertainty, volatility uncertainty, ,and tonic belief uncertainty describe inter-individual differences in how they update their beliefs about outcome variance, task volatility and the underlying mean. We first showed with simulations that these three parameters can reflect stable trait-like computational patterns given adequate task length. This reliability is significantly more difficult to achieve when the inputs in the task and participants' responses are binary, as is the case in the classical reversal learning tasks predominantly used in the literature. We then show that these parameters have moderate to high test-retest scores. By estimating the parameters through the hierarchical (multilevel) Bayesian framework, where all participant levels and group levels are estimated in one inferential step, we were able to show that participants' averaged beliefs about volatility or outcome variance, as well as precision weights on prediction errors, had excellent test-retest reliability. The excellent reliability was also reflected in three behavioral patterns: implicit learning rate, learning rate adaptation across variance blocks, and confidence adaptation to prediction error.

The parameter reflecting higher uncertainty about volatility correlated with higher subclinical schizotypal traits across two studies with healthy volunteers, and with positive symptoms in a sample of chronic schizophrenia patients, when playing to gain points. At the behavioral level, this parameter was associated with higher confidence even when predictions made were not very good (such as in trial immediately following mean changes). These findings suggest that psychotic symptoms in stable patients and schizotypal traits in healthy individuals are primarily related to an impaired use of higher-order representations, rather than a failure to attenuate sensations. This trait could have an adaptive value, as it allows for a rapid uncertainty reduction when faced with unpredictable situations. On the other hand, we found that traits related to anxiety and depression in a non-clinical sample were related to decreased volatility of beliefs about outcome variance and decreased tonic belief volatility, suggesting reduced learning rates for more rigid updating of beliefs about the mean as well as outcome variance. The same pattern was observed in relation to negative symptoms in the clinical sample of patients with schizophrenia, when playing to avoid losses.

In subsequent chapters, we investigated how D2 dopamine receptor antagonists affect the computational mechanisms identified in Chapter 2 as relevant to schizophrenia symptoms. In Chapter 3 we examined how sulpiride, a selective D2 receptor antagonist, affects belief rigidity when learning about prosocial attitudes through a repeated trust game. In a study with 76 male participants, using a HGF to model belief updating, we found that D2 antagonism increased the volatility of beliefs about others' trustworthiness, leading to higher precision weights on prediction errors. This effect was particularly pronounced in participants with higher dopamine availability (indirectly through a genetic polymorphism). The results from the model were reflected in behavior in two ways: participants under sulpiride changed their investment more from trial to trial and were more likely to reciprocate the other player's actions.

In Chapter 3, we examined the effects of another selective D2 receptor antagonist, amisulpride, on the ability to use structural knowledge about the world for planning and decision making. We

embedded this problem within the model-based/model-free reinforcement learning framework. Given the role of dopamine D2 receptors in both cognitive control and habit formation, a potential increase in model-based behavior can also be due to a reduction in habitual (model-free) processes. We therefore compared the effect of amisulpride with that of an antagonist of the opioid system (naltrexone), a neuromodulatory system that is known to be involved in the formation of habitual behavior but has no known direct effects on cognitive control. We found that amisulpride, but not naltrexone increased the weight on model-based compared to model-free behavior, suggesting that D2 receptor antagonism boosted processes that support model-based control, while leaving model-free or habitual tendencies intact.

In the remaining sections of this chapter, I will discuss the involvement of dopamine D2 receptors in hierarchical inference and learning before discussing the implications of these findings for our understanding of the mechanisms underlying the therapeutic effects of D2 antagonists. Finally, I will mention some potential applications of these findings for precision psychiatry.

The role of D2 receptors in learning and inference

Taken together, the results of Chapters 3 and 4 contribute to elucidating the rather complex role of D2 antagonists in learning. A consistent finding across both studies is that higher doses of D2 antagonists are likely going to increase the parameter that encodes stochasticity and randomness of choice. This was evident across both genetic groups when participants were administered 800 mg of sulpiride and was present in the group of individuals with higher serum levels when given 400 mg of amisulpride. As discussed in both chapters, this is consistent with a widely accepted model of cortico-striatal loops that defines the role of postsynaptic D2 receptors as inhibiting alternative actions (Frank, 2005; Maia & Frank, 2017), and is supported by numerous animal and human studies. Unfortunately, in the context of the experimental paradigms used in this thesis, it is not clear how this choice randomness can be interpreted at the psychological or cognitive level. Several plausible explanations have been suggested, such as increased exploratory behavior, sloppiness in decision making, or increased uncertainty (lower confidence) in action selection. While increased exploration makes sense in the Two-step task due to the lack of feedback on unchosen options, it is not a meaningful concept in the context of the repeated Trust game, where participants were told that responses were prerecorded. Nevertheless, we cannot completely exclude the option that participants explored different investment strategies (see the discussion on multiple models below). The conclusion that D2 antagonism leads to increased sloppiness is plausible because both tasks require at least some cognitive control. Indeed, sulpiride impaired working memory performance in the cohort reported in Chapter 3 (Naef et al., 2017), a finding consistent with some previous work (Mehta et al., 2008), but not others (Dodds et al., 2009; Westbrook et al., 2020). The second explanation, that blocking post-synaptic D2 receptors increases choice uncertainty, is elegant and parsimonious. Given the active inference formulation of action as inference, one might conclude that post-synaptic D2 receptors encode precision more generally, thereby affecting both beliefs and action selection strategies. For a less speculative conclusion as to which of these explanations are correct, future studies will have to delineate the various behavioral patterns more precisely with a richer set of variables within a single task design. Such tasks would for instance differentiate between directed and random exploration (Gershman & Tzovaras, 2018), better measurement of attentional focus, and different degrees of complexities of the "cognitive maps" needed to navigate the task (van Opheusden et al., 2023).

The main finding of Chapter 4, that D2 antagonism increases the capacity for model-based behavior, seems somewhat contradictory to the finding in Chapter 3 that it makes beliefs more flexible. How can the same mechanism of action make beliefs looser and increase the ability to maintain an online representation of the world while making decisions? However, flexible belief updating plays an important role in both experimental paradigms. In the Two-step task, the points on each planet are constantly changing, therefore both flexible belief updating and keeping the mapping between levels online are prerequisites for good performance. We did not explicitly measure the rate of belief updating in the Two-step task, which precludes a direct comparison of belief rigidity across the two tasks. Of

note, the two processes (flexible belief updating and model-based control) are thought to be supported by distinct brain circuits. Seminal work by Goldman-Rakic has shown that prefrontal D1 receptors (and not D2 receptors) are pivotal for maintaining working memory representations (Sawaguchi & Goldman-Rakic, 1991), whereas D2 dopaminergic receptors in the midbrain-basal ganglia circuits regulate the flow of information to the cortex (Frank, 2005). It is also possible that the effect of D2 antagonism in the Two-step task is due to improved flexibility in updating beliefs. Across the two studies, our findings support the notion that D2 receptors play an important role in both processes. Blocking D2 receptors throughout the brain may leave the D1-supported working memory capacity intact, while making the striatal circuits more D1-based and therefore more flexible. Increased belief flexibility would however benefit both the "model-based" and the "model-free" components of the dual-state reinforcement learning model applied to task behavior in Chapter 4.

It is worth noting that the behavior in the Two-step task described by the single dimension along the model-based/model-free dichotomy has recently been scrutinized (Collins & Cockburn, 2020). Participants can appear to be less model-based and more model-free by misunderstanding the task instructions or by using strategies not described by either system, such as perceiving one stimulus as unlucky, perceiving a row of losses as indicative of a future win ("gambler's fallacy"), or finding spurious patterns in button-presses or outcomes (e.g. twice left and twice right) (Feher da Silva & Hare, 2020). This line of thinking extends to the behavior in the repeated Trust game, where participants might form different strategies that are not accounted for in the model. To move forward, what is needed are more precise considerations of the computational building blocks that underpin various behaviors, complemented by behavioral paradigms that facilitate the estimation of these mechanisms (Collins & Cockburn, 2020; Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017). In hindsight, this could have been addressed by investigating the effects of a D2 antagonist on behavior in a task that measures belief rigidity and the capacity for finding more abstract statistical patterns, such as the predictive inference task used in Chapter 2.

An important difference between the two studies in Chapters 3 and 4 is the dose and drug used: 800 mg of sulpiride in the Trust game and 400 mg of amisulpride in the Two-step task. Amisulpride and sulpiride have a similar chemical structure, both are highly selective for D2 and D3 dopamine receptor subtypes and both are believed to target predominantly presynaptic D2 auto-receptors at low doses. Evidence from neurochemical imaging studies suggest that 400 mg of sulpiride leads to ~30% receptor occupancy (Mehta et al., 2008), whereas 800 mg is believed to more than double that (at ~60%) and is therefore more likely to target postsynaptic D2 receptors (Takano et al., 2006). There are however several limitations that make the comparison of drug effects across the two studies difficult. The conclusions on acute dose occupancy are based on relatively low sample numbers (only two individuals were given 800 mg sulpiride in the study by Takano et al.), and should be assumed with caution, especially given the large interindividual variability in the dose-occupancy relationship reported in clinical studies (Howes, Egerton, et al., 2009). Furthermore, there are slight differences in the chemical structure and therefore in the pharmacological profile of the two compounds. Amisulpride is considered to be more potent and shows greater selectivity presynaptic dopamine receptors (Pani, Villagrán, Kontaxakis, & Alptekin, 2008). With these differences in mind, a strength in both studies was that serum levels in the blood were collected. The effect of amisulpride on model-based behavior was only present in the group of participants with low serum levels, and the effect of sulpiride on belief volatility scaled positively with serum levels. This tentatively suggests that the effects of amisulpride are more likely to have been caused by presynaptic D2 receptor antagonism, while the effects of sulpiride are likely due to postsynaptic D2 receptor antagonism.

The mechanistic role of D2 antagonism in therapeutic outcome

An overarching goal of this thesis was to better understand how antipsychotic medication affects behavioral and computational processes relevant for patients with schizophrenia: belief rigidity and the ability to use higher-order representations (models) for prediction and planning. In contrast to studies in patients treated with antipsychotics, we conducted psychopharmacological experiments in healthy

volunteers. This allowed for better control of variables that could confound the effects of D2 receptor occupancy but limits the implications we can make for outlining the mechanistic role of D2 antagonists on clinical outcome. I will now discuss these implications and the limitations.

Delusions are particularly difficult to treat due to their imperviousness to rational argument. D2 hyperstimulation can be one neural mechanism through which the rigidity of beliefs is manifested. The results in Chapter 3 suggest that D2 antagonism may increase the volatility of patients' beliefs and therefore help patients in letting go of their rigid delusional convictions. Furthermore, delusional beliefs are often interpersonal in nature and revolve around the idea of being harmed or threatened by others. Patients have often had a traumatic interpersonal experience (Howes, McCutcheon, Owen, & Murray, 2017), and many tend to isolate themselves socially (Meyer-Lindenberg & Tost, 2012). Beliefs of persecution are likely to stem from a lack of genuine and trusting relationships with others. Treatment of these symptoms, particularly in severely ill patients, often requires a multifaceted therapeutic approach in which antipsychotic medication is accompanied by psychosocial therapy and psychotherapy. he findings in Chapter 3 tentatively suggest that D2 antagonism may lead to greater openness to others, thereby reducing paranoid tendencies. This may be particularly relevant to the formation of a therapeutic alliance between the patient and the therapist, which could increase the effectiveness of adjunct psychotherapy treatment. Indeed, one study showed that patients with higher levels of belief instability (a concept similar to belief volatility) were more likely to benefit from cognitive behavioral therapy (Hauke et al., 2022).

Similarly, results from Chapter 4 suggest that D2 antagonism improves model-based planning. Findings from Chapter 2 support the notion that positive symptoms are related to the inability to recognize more abstract "higher order" statistical patterns, a computational parameter that was negatively correlated with the capacity for model-based behavior as measured by the Two-step task. This is in accordance with previous modelling, empirical and theoretical work showing that patients with schizophrenia might have impairments in model-based learning (Culbreth et al., 2016), goal-directed behavior (Morris, Quail, Griffiths, Green, & Balleine, 2015), and cognitive control more generally (Barch & Ceaser, 2012). These impairments are centered around the dorsolateral prefrontal cortex and its interactions with the thalamus and striatum (Barch & Ceaser, 2012). Importantly, the same connectivity patterns are diminished in patients with psychotic symptoms (Anticevic et al., 2014) and increase the risk of a full-blown episode in individuals at risk of developing psychosis (Anticevic et al., 2015). Electrophysiological work has demonstrated that it is the role of postsynaptic D2 receptors (but not D1 receptors) in the striatum in facilitating this thalamic-cortical communication and enabling goal-directed behavior and cognitive control (Goto & Grace, 2005). The implications of the findings in Chapter 4 are that D2 antagonists may improve the capacity for model-based behavior. We might speculate that the underlying clinical effect might be that patients find it easier to avoid habitual thought processes and build more appropriate models of the world. The implication of this is that the capacity of model-based behavior could be an important clinical target, but this has never been tested in practice, and evidence from studies in other disorders speaks against it. Impaired model-based behavior has been recognized as a transdiagnostic measure of compulsivity in clinical and non-clinical populations (Gillan et al., 2016; Voon et al., 2015). A recent study found that behavioral measures including a Two-step task administered on a smartphone did not predict clinical improvement in patients with depression and anxiety following internet-based psychotherapy (C. T. Lee et al., 2023).

In the data presented in Chapter 4, amisulpride increased model-based behavior only in participants with low serum levels in the blood. At low doses, both amisulpride and sulpiride are believed to primarily target D2 presynaptic auto-receptors, resulting in increased levels of dopamine in the synaptic cleft. This is thought to be one of the reasons why these two compounds are sometimes effective in treating depression and dysthymia, a chronic form of depression (Riederer, Laux, Nagatsu, Le, & Riederer, 2022). Although, as I reviewed above, 400 mg of amisulpride could already reach 60% occupancy of postsynaptic receptors, it is possible that the enhancement of model-based behavior

occurred through increased dopamine availability, reducing the degree of generalization for therapeutic pathways of antipsychotics.

To assess the relevance of these findings for understanding the therapeutic effects of D2 antagonists in clinical practice, it is important to acknowledge the similarities and differences between the acute effects of D2 antagonists in healthy participants and the daily intake of antipsychotics in treatment of patients with schizophrenia. The clinical efficiency of antipsychotic medication depends largely on the ability of the drug to occupy D2 receptors, with a therapeutic window between 65% and 80% receptor occupancy (Kaar et al., 2020), which is comparable to the receptor occupancy that could be expected from acute dosing in our studies. However, therapeutic effects of antipsychotics are known to only manifest with a delay of weeks or at least a few of days (Agid, Kapur, Arenovich, & Zipursky, 2003), which would imply that the therapeutic effects cannot be deducted from acute effects in either healthy volunteers or patients. One reason for this delayed clinical response may be that the instruments we use to assess clinical improvement are simply not designed for repeated testing. Clinical assessments based on questionnaires often ask patients to rate their mental state over the past couple of days, making them unsuitable for frequent testing. The second reason may be that antipsychotics alter a particular mechanism that is more downstream of the core biological abnormality, which means that therapeutic effects need to compound to produce a clinically observable difference. For example, if D2 antagonisms helps patients with model-based reasoning, it will take time for this change to manifest itself in reduced psychopathology as measured by established clinical scales. It may also be that there is a biological adaptation to antipsychotic treatment that takes days. One such model of the therapeutic effect is the depolarization blockade, according to which repeated administration of antipsychotics induces a state of dopamine neuron inactivation (Grace et al., 1997). However, there is evidence that the antipsychotic-induced depolarization block might occur much earlier (Valenti et al., 2011) consistent with the finding that high D2 receptor occupancy within the first two days of treatment predicts the clinical response (Catafau et al., 2006).

The Valenti et al. (2011) study also alludes to another important caveat for generalizing our findings to treatment in patients. The acute response to an antipsychotic may be very different in patients with increased presynaptic dopamine availability, compared to healthy volunteers (Grace, 2016). Although we attempted to account for these changes with genetic phenotyping in the study in Chapter 3, the effect sizes across genetic subgroups are not comparable to the difference in dopamine synthesis between patients and controls.

Utilizing computational phenotypes for precision psychiatry

In this thesis I showed that symptoms relevant of schizophrenia are related to computational processes that are in turn affected by D2 antagonism in healthy volunteers. Despite major differences in the responses of healthy volunteers to acute effects of antipsychotics and repeated intake of antipsychotics in patients, these findings open several interesting questions that can be addressed in future research.

A major problem in the treatment of psychotic symptoms in schizophrenia is that many patients do not respond to dopamine-based antipsychotics and have so-called treatment-resistant schizophrenia. Patients who do not improve after two dopamine-based antipsychotic treatments (given at an adequate dose for at least 6 weeks) are put on clozapine, the only treatment licensed for treatment-resistant schizophrenia. Treatment with clozapine is known to be most effective when given as early as possible (Correll et al., 2022). One reliable predictor of a treatment response is the level of presynaptic dopamine availability as measured by PET – a measure that is expensive and not available in routine clinical use. As mentioned above, another predictor might be an acute response to a dopaminergic antipsychotic, however, clinical scales used in clinical practice are not suitable for repeated assessment.

The notion of aberrant salience and its computational implementation within hierarchical Bayesian inference suggests that increased dopamine synthesis leads to weak predictions relative to the precision attributed to sensory signals. Work presented in this thesis and elsewhere has demonstrated

that positive symptoms are correlated (albeit with a small effect size) with poor higher-order representations and an over-reliance on sensory input. A testable hypothesis is that patients whose computational processes improve following acute antipsychotic administration are more likely to benefit from dopamine-based antipsychotic treatment. If this is the case, computational models coupled with behavioral tasks may deliver clinical markers that can be used for patient stratification.

Another important challenge in psychiatry is comorbidity. Patients with schizophrenia, schizoaffective disorder, and bipolar disorder can all have psychotic symptoms. On the other hand, both schizoaffective disorder and bipolar disorder have mood (depressive) symptoms as a central feature, whereas in schizophrenia the focus is predominantly on psychotic symptoms. However, the boundaries between these disorders are far from clear. The prevalence of depressive mood disorder in schizophrenia is reported to be around 40% (Conley, Ascher-Svanum, Zhu, Faries, & Kinon, 2007) and depends on the stage of the illness (early vs chronic), with up to 80% of patients with schizophrenia experiencing a depressive episode (Upthegrove, Marwaha, & Birchwood, 2017). Similarly, negative symptoms in schizophrenia are differentiated from depressive symptoms purely on the basis of severity. Despite the similarities between diagnoses, misdiagnosis is known to lead to ineffective or even harmful treatment interventions. Accurate diagnosis becomes paramount not only for treatment efficacy but also for predicting prognosis, managing potential side effects, and improving overall patient outcomes. As we have shown in Chapter 2, we can use computational approaches to explicitly model the interdependence of various computational mechanisms relevant for a range of psychopathologies. These latent computational phenotypes are independent of patient' diagnostic categories but have so far been benchmarked against the existing diagnostic labels. Future work will need to explore the potential of these computational dimensions for clinical outcomes, patient stratification or treatment prediction.

An important distinguishing feature between disorders is the causal structure of different symptoms. For example, schizoaffective and bipolar disorders are thought to be separable on the basis of the timing of psychotic symptoms in relation to mood disturbances. Indeed, it has been proposed that mental disorders emerge from complex systems of interacting processes (Fried, 2022). To accommodate the temporal dynamics of mental health symptoms, computational frameworks are being developed that allow for simultaneous modelling of within- and between-person dynamics (Borsboom et al., 2021).

Finally, it is crucial to acknowledge the important limitations of the computational frameworks we have adopted in this thesis. The core underlying assumption of this approach is that computational mechanisms can sufficiently "explain" the wide range of phenomenological and psychological experiences of patients with mental illness. This approach is fundamentally reductionist, assuming that complex mental health disorders can be fully understood in terms of lower-level computational processes. This perspective may overlook the multifactorial nature of psychopathology, which may involve a combination of genetic, environmental, social, and psychological factors.

While it is theoretically possible to construct more complex models and enhance experimental paradigms with domain-specific (e.g. social) stimuli, in practice it is incredibly difficult to experimentally capture the highly complex contexts in which patients' symptoms manifest, including the extremely rich social and cultural lives of people with mental illness. There is no doubt that more precise methods and personalised treatments will improve psychiatric care, but real improvements in mental health outcomes will only be achieved if these advances are coupled with more integrated and efficient health and social support systems.

References

Abbott, A. (2011). Novartis to shut brain research facility. *Nature*, *480*(7376), 161–162.

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

Adams, R. A., Moutoussis, M., Nour, M. M., Dahoun, T., Lewis, D., Illingworth, B., … Roiser, J. P. (2020). Variability in Action Selection Relates to Striatal Dopamine 2/3 Receptor Availability in Humans: A PET Neuroimaging Study Using Reinforcement Learning and Active Inference Models. *Cerebral Cortex*, *30*(6), 3573–3589. https://doi.org/10.1093/cercor/bhz327

Adams, R. A., Napier, G., Roiser, J. P., Mathys, C. D., & Gilleen, J. (2018). Attractor-like dynamics in belief updating in schizophrenia. *Journal of Neuroscience*, *38*(44), 9471–9485. https://doi.org/10.1523/JNEUROSCI.3163-17.2018

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The Computational Anatomy of Psychosis. *Frontiers in Psychiatry*, *4*(May), 1–26. https://doi.org/10.3389/fpsyt.2013.00047

Adams, R. A., Vincent, P., Benrimoh, D., Friston, K. J., & Parr, T. (2021). Everything is connected: Inference and attractors in delusions. *Schizophrenia Research*, (March). https://doi.org/10.1016/j.schres.2021.07.032

Agid, O., Kapur, S., Arenovich, T., & Zipursky, R. B. (2003). Delayed-Onset Hypothesis of Antipsychotic Action: A Hypothesis Tested and Rejected. *Archives of General Psychiatry*, *60*(12), 1228–1235. https://doi.org/10.1001/archpsyc.60.12.1228

Ågmo, A., Galvan, A., & Talamantes, B. (1995). Reward and reinforcement produced by drinking sucrose: Two processes that may depend on different neurotransmitters. *Pharmacology Biochemistry and Behavior*, *52*(2), 403–414. https://doi.org/https://doi.org/10.1016/0091-3057(95)00128-J

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*. https://doi.org/10.1162/cpsy_a_00002

Amat, J., Baratta, M. V., Paul, E., Bland, S. T., Watkins, L. R., & Maier, S. F. (2005). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nature Neuroscience*, *8*(3), 365–371. https://doi.org/10.1038/nn1399

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

Anticevic, A., Cole, M. W., Repovs, G., Murray, J. D., Brumbaugh, M. S., Winkler, A. M., … Glahn, D. C. (2014). Characterizing thalamo-cortical disturbances in schizophrenia and bipolar illness. *Cerebral Cortex*, *24*(12), 3116–3130.

Anticevic, A., Haut, K., Murray, J. D., Repovs, G., Yang, G. J., Diehl, C., … Cannon, T. D. (2015). Association of thalamic dysconnectivity and conversion to psychosis in youth and young adults at elevated clinical risk. *JAMA Psychiatry*, *72*(9), 882–891. https://doi.org/10.1001/jamapsychiatry.2015.0566

Arnsten, A. F. T. (2011). Catecholamine influences on dorsolateral prefrontal cortical networks. *Biological Psychiatry*, *69*(12), e89–e99. https://doi.org/10.1016/j.biopsych.2011.01.027

Babayan, B. M., Uchida, N., & Gershman, S. J. (2018). Belief state representation in the dopamine system. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04397-0

Baker, S. C., Konova, A. B., Daw, N. D., & Horga, G. (2019). A distinct inferential mechanism for delusions in schizophrenia. *Brain*, *142*(6), 1797–1812. https://doi.org/10.1093/brain/awz051

Balleine, B. W., & O'Doherty, J. P. (2010). Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacology*,

35(1), 48–69. https://doi.org/10.1038/npp.2009.131

Barch, D. M., & Ceaser, A. (2012). Cognition in schizophrenia: core psychological and neural mechanisms. *Trends in Cognitive Sciences*, *16*(1), 27–34. https://doi.org/10.1016/j.tics.2011.11.015

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*.

Becker, A., Grecksch, G., & Kraus, J. (2002). Rewarding effects of ethanol and cocaine in μ opioid receptor-deficient mice, 296–302. https://doi.org/10.1007/s00210-002-0533-2

Benjamin, D., Grant, E. R., & Pohorecky, L. A. (1993). Naltrexone reverses ethanol-induced dopamine release in the nucleus accumbens in awake, freely moving rats. *Brain Research*. https://doi.org/10.1016/0006-8993(93)90309-B

Bentall, R. P. (2004). *Madness explained: Psychosis and human nature*. Penguin UK.

Bentall, R. P., Baker, G. A., & Havers, S. (1991). Reality monitoring and psychotic hallucinations. *British Journal of Clinical Psychology*, *30*(3), 213–222.

Bentall, R. P., Kinderman, P., & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: Towards a model of persecutory delusions. *Behaviour Research and Therapy*, *32*(3), 331–341. https://doi.org/10.1016/0005-7967(94)90131-7

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*. https://doi.org/10.1006/game.1995.1027

Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology*, *191*(3), 391–431. https://doi.org/10.1007/s00213-006-0578-x

Bland, A. R., Roiser, J. P., Mehta, M. A., Schei, T., Boland, H., Campbell-Meiklejohn, D. K., … Elliott, R. (2016). EMOTICOM: A Neuropsychological Test Battery to Evaluate Emotion, Motivation, Impulsivity, and Social Cognition. *Frontiers in Behavioral Neuroscience*, *10*(February). https://doi.org/10.3389/fnbeh.2016.00025

Boehme, R., Deserno, L., Gleich, T., Katthagen, T., Pankow, A., Behr, J., … Schlagenhauf, F. (2015). Aberrant Salience Is Related to Reduced Reinforcement Learning Signals and Elevated Dopamine Synthesis Capacity in Healthy Adults. *Journal of Neuroscience*, *35*(28), 10103–10111. https://doi.org/10.1523/JNEUROSCI.0805-15.2015

Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *American Economic Review*, *98*(1), 294–310.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., … Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1). https://doi.org/10.1038/s43586-021-00055-w

Bressan, R. A., Erlandsson, K., Jones, H. M., Mulligan, R., Flanagan, R. J., Ell, P. J., & Pilowsky, L. S. (2003). Is regionally selective D2/D3 dopamine occupancy sufficient for atypical antipsychotic effect? A in vivo quantitative [123I]epidepride SPET study of amisulpride-treated patients. *American Journal of Psychiatry*, *160*(8), 1413–1420. https://doi.org/10.1176/appi.ajp.160.8.1413

Bressolle, F., Bres, J., & Fauré-Jeantis, A. (1992). Absolute bioavailability, rate of absorption, and dose proportionality of sulpiride in humans. *Journal of Pharmaceutical Sciences*, *81*(1), 26–32. https://doi.org/10.1002/jps.2600810106

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

Browning, M., Carter, C. S., Chatham, C., Den Ouden, H., Gillan, C. M., Baker, J. T., … Paulus, M. (2020). Realizing the Clinical Potential of Computational Psychiatry: Report From the Banbury Center Meeting, February 2019. *Biological Psychiatry*, *88*(2), e5–e10. https://doi.org/10.1016/j.biopsych.2019.12.026

Brozoski, T. J., Brown, R. M., Rosvold, H. E., & Goldman, P. S. (1979). Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science*, *205*(4409), 929–932.

Burke, D. A., Rotstein, H. G., & Alvarez, V. A. (2017). Striatal Local Circuitry: A New Framework for Lateral Inhibition. *Neuron*, *96*(2), 267–284. https://doi.org/10.1016/j.neuron.2017.09.019

Bürkner, P.-C. (2017). brms : An R Package for Bayesian Multilevel Models Using Stan . *Journal of Statistical Software*. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal Regression Models in Psychology: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *2*(1), 77–101. https://doi.org/10.1177/2515245918823199

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. *pup.princeton.edu*. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=10299404497160394093related:bVVOU9LV7o4J%5Cnhttp://pup.princeton.edu/chapters/i7517.html

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v076.i01

Catafau, A. M., Corripio, I., Pérez, V., Martin, J. C., Schotte, A., Carrió, I., & Alvarez, E. (2006). Dopamine D2 receptor occupancy by risperidone: implications for the timing and magnitude of clinical response. *Psychiatry Research: Neuroimaging*, *148*(2–3), 175–183.

Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., … Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: A machine learning approach. *The Lancet Psychiatry*, *3*(3), 243–250. https://doi.org/10.1016/S2215-0366(15)00471-X

Chiara, G. Di. (1999). Drug addiction as dopamine-dependent associative learning disorder. *European Journal of Pharmacology*, *375*, 13–30.

Cohen, J. D. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *351*(1346), 1515–1527. https://doi.org/10.1098/rstb.1996.0138

Cohen, M. X., Krohn-Grimberghe, A., Elger, C. E., & Weber, B. (2007). Dopamine gene predicts the brain's response to dopaminergic drug. *European Journal of Neuroscience*, *26*(12), 3652–3660. https://doi.org/10.1111/j.1460-9568.2007.05947.x

Collins, A. G. E., Albrecht, M. A., Waltz, J. A., Gold, J. M., & Frank, M. J. (2017). Interactions Among Working Memory, Reinforcement Learning, and Effort in Value-Based Choice: A New Paradigm and Selective Deficits in Schizophrenia. *Biological Psychiatry*, *82*(6), 431–439. https://doi.org/10.1016/j.biopsych.2017.05.017

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

Collins, A. G. E., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working Memory Load Strengthens Reward Prediction Errors. *The Journal of Neuroscience*, *37*(16), 4332–4342. https://doi.org/10.1523/JNEUROSCI.2700-16.2017

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

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

Conant, R. C., & Ashby, R. W. (1970). Every good regulator of a system must be a model of that

system. *International Journal of Systems Science*, *1*(2), 89–97.

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

Cools, R. (2016). The costs and benefits of brain dopamine for cognitive control. *Wiley Interdisciplinary Reviews. Cognitive Science*. https://doi.org/10.1002/wcs.1401

Cools, R. (2019). Chemistry of the Adaptive Mind: Lessons from Dopamine. *Neuron*. Elsevier Inc. https://doi.org/10.1016/j.neuron.2019.09.035

Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2011.03.028

Corlett, P. R., & Fletcher, P. C. (2012). The neurobiology of schizotypy: Fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. *Neuropsychologia*, *50*(14), 3612–3620. https://doi.org/10.1016/j.neuropsychologia.2012.09.045

Corlett, P. R., Frith, C. D., & Fletcher, P. C. (2009). From drugs to deprivation: A Bayesian framework for understanding models of psychosis. *Psychopharmacology*, *206*(4), 515–530. https://doi.org/10.1007/s00213-009-1561-0

Corlett, P. R., Honey, G. D., Aitken, M. R. F. F., Dickinson, A., Shanks, D. R., Absalom, A. R., … Fletcher, P. C. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: Linking cognition, brain activity, and psychosis. *Archives of General Psychiatry*, *63*(6), 611–621. https://doi.org/10.1001/archpsyc.63.6.611

Corlett, P. R., Krystal, J. H., Taylor, J. R., & Fletcher, P. C. (2009). Why do delusions persist? *Frontiers in Human Neuroscience*, *3*(JUL), 1–9. https://doi.org/10.3389/neuro.09.012.2009

Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R. F., Shanks, D. R., Robbins, T. W., … Fletcher, P. C. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*, *130*(9), 2387–2400.

Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in Neurobiology*, *92*(3), 345–369. https://doi.org/10.1016/j.pneurobio.2010.06.007

Correll, C. U., Agid, O., Crespo-Facorro, B., de Bartolomeis, A., Fagiolini, A., Seppälä, N., & Howes, O. D. (2022). A Guideline and Checklist for Initiating and Managing Clozapine Treatment in Patients with Treatment-Resistant Schizophrenia. *CNS Drugs*, *36*(7), 659–679. https://doi.org/10.1007/s40263-022-00932-2

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

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*(1). https://doi.org/10.1186/1741-7015-11-126

Daunizeau, J., Den Ouden, H., Pessiglione, M., Kiebel, S. J., Stephan, K. E., & Friston, K. J. (2010). Observing the observer (I): Meta-bayesian models of learning and decision-making. *PLoS ONE*, *5*(12). https://doi.org/10.1371/journal.pone.0015554

Daw, N. D. (2018). Are we of two minds? *Nature Neuroscience*, *21*(11), 1497–1499. https://doi.org/10.1038/s41593-018-0258-2

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

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

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The helmholtz machine. *Neural Computation*, *7*(5), 889–904.

De Berker, A. O., Rutledge, R. B., Mathys, C. D., Marshall, L., Cross, G. F., Dolan, J. R., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, *7*(1), 10996. https://doi.org/10.1038/ncomms10996

De Lafuente, V., & Romo, R. (2011). Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(49), 19767–19771. https://doi.org/10.1073/pnas.1117636108

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

DeCoster, J., Gallucci, M., & Iselin, A.-M. R. (2011). Best Practices for Using Median Splits, Artificial Categorization, and their Continuous Alternatives. *Journal of Experimental Psychopathology*, *2*(2), 197–209. https://doi.org/10.5127/jep.008310

Del Campo, A. F. M., McMurray, R. G., Besser, G. M., & Grossman, A. (1992). Effect of 12-hour infusion of naloxone on mood and cognition in normal male volunteers. *Biological Psychiatry*, *32*(4), 344–353.

Delamater, A. R., Sclafani, A., & Bodnar, R. J. (2000). Pharmacology of Sucrose-Reinforced Place-Preference Conditioning: Effects of Naltrexone. *Pharmacology Biochemistry and Behavior*, *65*(4), 697–704. https://doi.org/https://doi.org/10.1016/S0091-3057(99)00251-8

Deserno, L., Boehme, R., Heinz, A., & Schlagenhauf, F. (2013). Reinforcement learning and dopamine in schizophrenia: Dimensions of symptoms or specific features of a disease group? *Frontiers in Psychiatry*, *4*(DEC), 1–16. https://doi.org/10.3389/fpsyt.2013.00172

Deserno, L., Boehme, R., Mathys, C. D., Katthagen, T., Stephan, K. E., Heinz, A., & Schlagenhauf, F. (2019). Volatility estimates increase choice switching and relate to prefrontal activity in schizophrenia. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2019.10.007

Deserno, L., Heinz, A., & Schlagenhauf, F. (2015). Computational approaches to schizophrenia: A perspective on negative symptoms. *Schizophrenia Research*, *186*, 46–54. https://doi.org/10.1016/j.schres.2016.10.004

Deserno, L., Huys, Q. J. M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A. A., … Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences*, *112*(5), 1595–1600. https://doi.org/10.1073/pnas.1417219112

Diaconescu, A. O., Wellstein, K. V., Kasper, L., Mathys, C. D., & Stephan, K. E. (2020). Hierarchical bayesian models of social inference for probing persecutory delusional ideation. *Journal of Abnormal Psychology*, *129*(6), 556–569. https://doi.org/10.1037/abn0000500

Dickinson, A. (1985). The Development of Behavioural Autonomy. *Philosophical Transactions of the Royal Society of London . Series B.*

Diederen, K. M. J., Spencer, T., Vestergaard, M. D., Fletcher, P. C., & Schultz, W. (2016). Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron*, *90*(5), 1127–1138. https://doi.org/10.1016/j.neuron.2016.04.019

Diederen, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *Journal of Neuroscience*, *37*(7), 1708–1720. https://doi.org/10.1523/JNEUROSCI.1979-16.2016

Dobbs, L. K., Kaplan, A. R., Lemos, J. C., Matsui, A., Rubinstein, M., Alvarez, V. A., … Alvarez, V.

A. (2016). Dopamine Regulation of Lateral Inhibition between Striatal Neurons Gates the Stimulant Actions of Article Dopamine Regulation of Lateral Inhibition between Striatal Neurons Gates the Stimulant Actions of Cocaine. *Neuron*, *90*(5), 1100–1113. https://doi.org/10.1016/j.neuron.2016.04.031

Dodds, C. M., Clark, L., Dove, A., Regenthal, R., Baumann, F., Bullmore, E. T., … Müller, U. (2009). The dopamine D2 receptor antagonist sulpiride modulates striatal BOLD signal during the manipulation of information in working memory. *Psychopharmacology*, *207*(1), 35–45. https://doi.org/10.1007/s00213-009-1634-0

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

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *Journal of Neuroscience*, *36*(4), 1211–1222. https://doi.org/10.1523/JNEUROSCI.1901-15.2016

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic Genes Predict Individual Differences in Susceptibility to Confirmation Bias. *Journal of Neuroscience*, *31*(16), 6188–6198. https://doi.org/10.1523/JNEUROSCI.6486-10.2011

Doll, B. B., Waltz, J. A., Cockburn, J., Brown, J. K., Frank, M. J., & Gold, J. M. (2014). Reduced susceptibility to confirmation bias in schizophrenia. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 715–728. https://doi.org/10.3758/s13415-014-0250-6

Dreyer, J. K., Herrik, K. F., Berg, R. W., & Hounsgaard, J. D. (2010). Influence of Phasic and Tonic Dopamine Release on Receptor Activation. *Journal of Neuroscience*, *30*(42), 14273–14283. https://doi.org/10.1523/JNEUROSCI.1894-10.2010

Dudley, R., Taylor, P., Wickham, S., & Hutton, P. (2016). Psychosis, delusions and the "Jumping to Conclusions" reasoning bias: A systematic review and meta-analysis. *Schizophrenia Bulletin*, *42*(3), 652–665. https://doi.org/10.1093/schbul/sbv150

Durstewitz, D., & Seamans, J. K. (2008). The Dual-State Theory of Prefrontal Cortex Dopamine Function with Relevance to Catechol-O-Methyltransferase Genotypes and Schizophrenia. *Biological Psychiatry*, *64*(9), 739–749. https://doi.org/10.1016/j.biopsych.2008.05.015

Eisenegger, C., Knoch, D., Ebstein, R. P., Gianotti, L. R. R., Sándor, P. S., & Fehr, E. (2010). Dopamine receptor D4 polymorphism predicts the effect of L-DOPA on gambling behavior. *Biological Psychiatry*, *67*(8), 702–706.

Eisenegger, C., Naef, M., Linssen, A., Clark, L., Gandamaneni, P. K., Müller, U., & Robbins, T. W. (2014). Role of Dopamine D2 Receptors in Human Reinforcement Learning. *Neuropsychopharmacology*, *39*(10), 2366–2375. https://doi.org/10.1038/npp.2014.84

Eisenegger, C., Pedroni, A., Rieskamp, J., Zehnder, C., Ebstein, R., Fehr, E., & Knoch, D. (2013). DAT1 Polymorphism Determines L-DOPA Effects on Learning about Others' Prosociality. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0067820

Elkis, H., & Buckley, P. F. (2016). Treatment-Resistant Schizophrenia. *Psychiatric Clinics of North America*. Elsevier Inc. https://doi.org/10.1016/j.psc.2016.01.006

EMCDDA. (2020). European Drug Report 2020: Trends and Developments. Publications Office of the European Union Luxembourg.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., Mackinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test – retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences of the United States of America*, 1–6. https://doi.org/10.1073/pnas.1818430116

Erdmann, T., & Mathys, C. D. (2022). A generative framework for the study of delusions. *Schizophrenia Research*, *245*, 42–49. https://doi.org/10.1016/j.schres.2020.11.048

Everitt, B. J., Belin, D., Economidou, D., Pelloux, Y., Dalley, J. W., & Robbins, T. W. (2008). Neural

mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1507), 3125–3135. https://doi.org/10.1098/rstb.2008.0089

Eyny, Y. S., & Horvitz, J. C. (2003). Opposing roles of D1 and D2 receptors in appetitive conditioning. *Journal of Neuroscience*, *23*(5), 1584–1587. https://doi.org/10.1523/jneurosci.23-05-01584.2003

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

Fehr, E. (2009). On The Economics and Biology of Trust. *Journal of the European Economic Association*, *7*(2–3), 235–266. https://doi.org/10.1162/JEEA.2009.7.2-3.235

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, *13*(1), 1–25. https://doi.org/10.1007/s12110-002-1012-7

FeldmanHall, O., Raio, C. M., Kubota, J. T., Seiler, M. G., & Phelps, E. A. (2015). The Effects of Social Context and Acute Stress on Decision Making Under Uncertainty. *Psychological Science*, *26*(12), 1918–1926. https://doi.org/10.1177/0956797615605807

FeldmanHall, O., & Shenhav, A. (2019a). Resolving uncertainty in a social world. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0590-x

FeldmanHall, O., & Shenhav, A. (2019b). Resolving uncertainty in a social world. *Nature Human Behaviour*. Springer US. https://doi.org/10.1038/s41562-019-0590-x

Fett, A. K. J., Shergill, S. S., Joyce, D. W., Riedl, A., Strobel, M., Gromann, P. M., & Krabbendam, L. (2012). To trust or not to trust: The dynamics of social interaction in psychosis. *Brain*, *135*(3), 976–984. https://doi.org/10.1093/brain/awr359

Feyaerts, J., Henriksen, M. G., Vanheule, S., Myin-Germeys, I., & Sass, L. A. (2021). Delusions beyond beliefs: a critical overview of diagnostic, aetiological, and therapeutic schizophrenia research from a clinical-phenomenological perspective. *The Lancet Psychiatry*. Elsevier Ltd. https://doi.org/10.1016/S2215-0366(20)30460-0

File, S. E., & Silverstone, T. (1981). Naloxone changes self-ratings but not performance in normal subjects. *Psychopharmacology*, *74*(4), 353–354.

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., & Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature Neuroscience*, *22*(12), 2066–2077. https://doi.org/10.1038/s41593-019-0518-9

Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*, *11*(8), 966–973. https://doi.org/10.1038/nn.2159

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete Coding of Reward Dopamine Neurons. *Science*, *299*(March), 1898–1902. Retrieved from http://www.sciencemag.org/content/299/5614/1898.short

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2005). Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behavioral and Brain Functions*, *1*, 1–5. https://doi.org/10.1186/1744-9081-1-7

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*(1), 48–58. https://doi.org/10.1038/nrn2536

Ford, C. P. (2014). The role of D2-autoreceptors in regulating dopamine neuron activity and transmission. *Neuroscience*. https://doi.org/10.1016/j.neuroscience.2014.01.025

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, *17*(1), 51–72. https://doi.org/10.1162/0898929052880093

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, *12*(8), 1062–1068. https://doi.org/10.1038/nn.2342

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences*, *104*(41), 16311–16316. https://doi.org/10.1073/pnas.0706111104

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

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Freedman, R., Lewis, D. A., Michels, R., Pine, D. S., Schultz, S. K., Tamminga, C. A., … Yager, J. (2013). The initial field trials of DSM-5: new blooms and old thorns. *The American Journal of Psychiatry*, *170*(1), 1–5. https://doi.org/10.1176/appi.ajp.2012.12091189

Freeman, D. (2016). Persecutory delusions: a cognitive perspective on understanding and treatment. *The Lancet Psychiatry*, *3*(7), 685–692. https://doi.org/10.1016/S2215-0366(16)00066-3

Fried, E. I. (2022). Studying Mental Health Problems as Systems, Not Syndromes. *Current Directions in Psychological Science*, 1–20. https://doi.org/10.1177/09637214221114089

Friston, K. J. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11), 1000211. https://doi.org/10.1371/journal.pcbi.1000211

Friston, K. J. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. J., Rigoli, F., Ognibene, D., Mathys, C. D., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214. https://doi.org/10.1080/17588928.2015.1020053

Friston, K. J., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T. E. J., & Dolan, J. R. (2013). The anatomy of choice: active inference and agency. *Frontiers in Human Neuroscience*, *7*(September), 1–18. https://doi.org/10.3389/fnhum.2013.00598

Friston, K. J., Shiner, T., FitzGerald, T., Galea, J. M., Adams, R. A., Brown, H., … Bestmann, S. (2012). Dopamine, affordance and active inference. *PLoS Computational Biology*, *8*(1). https://doi.org/10.1371/journal.pcbi.1002327

Friston, K. J., Stephan, K. E., Montague, P. R., & Dolan, J. R. (2014). Computational psychiatry: The brain as a phantastic organ. *The Lancet Psychiatry*. https://doi.org/10.1016/S2215-0366(14)70275-5

Frith, C., Rees, G., & Friston, K. (1998). Psychosis and the experience of self: brain systems underlying self-monitoring. *Annals of the New York Academy of Sciences*, *843*(1), 170–178.

Fromm, S., Katthagen, T., Deserno, L., Heinz, A., Kaminski, J., & Schlagenhauf, F. (2023). Belief Updating in Subclinical and Clinical Delusions. *Schizophrenia Bulletin Open*, *4*(1), 1–11. https://doi.org/10.1093/schizbullopen/sgac074

Fuchs, T. (2015). The intersubjectivity of delusions. *World Psychiatry*, *14*(2), 178–179. https://doi.org/10.1002/wps.20209

Fuke, S. (2001). The VNTR polymorphism of the human dopamine transporter (DAT!) gene affects gene expression. *Pharmacogenomics Journal*. https://doi.org/10.1038/sj.tpj.6500026

Fusar-Poli, P., Howes, O. D., Allen, P., Broome, M., Valli, I., Asselin, M. C., … McGuire, P. (2011).

Abnormal prefrontal activation directly related to pre-synaptic striatal dopamine dysfunction in people at clinical high risk for psychosis. *Molecular Psychiatry*, *16*(1), 67–75. https://doi.org/10.1038/mp.2009.108

Fusar-Poli, P., & Meyer-Lindenberg, A. (2013). Striatal presynaptic dopamine in schizophrenia, part II: Meta-analysis of [18F/11C]-DOPA PET studies. *Schizophrenia Bulletin*, *39*(1), 33–42. https://doi.org/10.1093/schbul/sbr180

Garety, P. A., Hemsley, D. R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease*, *179*(4), 194–201. https://doi.org/10.1097/00005053-199104000-00003

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*. https://doi.org/10.1214/ss/1177011136

Gershman, S. J. (2015). A Unifying Probabilistic View of Associative Learning. *PLoS Computational Biology*, *11*(11), 1–20. https://doi.org/10.1371/journal.pcbi.1004567

Gershman, S. J. (2018a). Dopamine, Inference, and Uncertainty. *Neural Computation*, *3326*(12), 3311–3326. https://doi.org/10.1162/NECO

Gershman, S. J. (2018b). How to never be wrong. *Psychonomic Bulletin and Review*, 1. Retrieved from https://goo.gl/RRen3T

Gershman, S. J., & Tzovaras, B. G. (2018). Dopaminergic genes are associated with both directed and random exploration. *Neuropsychologia*, *120*(September), 97–104. https://doi.org/10.1016/j.neuropsychologia.2018.10.009

Gershman, S. J., & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience*, *20*(November), 703–714. https://doi.org/10.1038/s41583-019-0220-7

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

Gillan, C. M., & Whelan, R. (2017). What big data can do for treatment in psychiatry. *Current Opinion in Behavioral Sciences*, *18*, 34–42. https://doi.org/10.1016/j.cobeha.2017.07.003

Gluskin, B. S., & Mickey, B. J. (2016). Genetic variation and dopamine D2 receptor availability: A systematic review and meta-analysis of human in vivo molecular imaging studies. *Translational Psychiatry*, *6*(January). https://doi.org/10.1038/tp.2016.22

Gold, J. M., Strauss, G. P., Waltz, J. A., Robinson, B. M., Brown, J. K., & Frank, M. J. (2013). Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biological Psychiatry*, *74*(2), 130–136. https://doi.org/10.1016/j.biopsych.2012.12.022

Gold, J. M., Waltz, J. A., & Frank, M. J. (2015). Effort Cost Computation in Schizophrenia: A Commentary on the Recent Literature. *Biological Psychiatry*, *78*(11), 747–753. https://doi.org/10.1016/j.biopsych.2015.05.005

Goldman-Rakic, P. S. (1997). The Cortical Dopamine System: Role in Memory and Cognition. *Advances in Pharmacology*. https://doi.org/10.1016/S1054-3589(08)60846-7

Goto, Y., & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, *8*(6), 805–812. https://doi.org/10.1038/nn1471

Grace, A. A. (2016). Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nature Reviews Neuroscience*. Nature Publishing Group. https://doi.org/10.1038/nrn.2016.57

Grace, A. A., Bunney, B. S., Moore, H., & Todd, C. L. (1997). Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. *Trends in Neurosciences*, *20*(1), 31–37. https://doi.org/10.1016/S0166-2236(96)10064-3

Gromann, P. M., Heslenfeld, D. J., Fett, A. K. J., Joyce, D. W., Shergill, S. S., & Krabbendam, L. (2013). Trust versus paranoia: Abnormal response to social reward in psychotic illness. *Brain*, *136*(6), 1968–1975. https://doi.org/10.1093/brain/awt076

Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *ELife*, *5*(MARCH2016), 1–19. https://doi.org/10.7554/eLife.13388

Haarsma, J., Fletcher, P. C., Griffin, J. D., Taverne, H. J., Ziauddeen, H., Spencer, T. J., … Murray, G. K. (2020). Precision weighting of cortical unsigned prediction error signals benefits learning, is mediated by dopamine, and is impaired in psychosis. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-020-0803-8

Hahn, S., Moritz, S., Elmers, J., & Scheunemann, J. (2021). Do you like cliff-hangers? Objective versus subjective need for closure in the schizophrenia spectrum. *Schizophrenia Research*, *238*, 20–26. https://doi.org/https://doi.org/10.1016/j.schres.2021.09.013

Harrington, A. (2019). *Mind fixers: Psychiatry's troubled search for the biology of mental illness*. WW Norton & Company.

Hartmann, M. N., Kluge, A., Kalis, A., Mojzisch, A., Tobler, P. N., & Kaiser, S. (2015). Apathy in schizophrenia as a deficit in the generation of options for action. *Journal of Abnormal Psychology*, *124*(2), 309–318. https://doi.org/10.1037/abn0000048

Hauke, D. J., Roth, V., Karvelis, P., Adams, R. A., Moritz, S., Borgwardt, S., … Andreou, C. (2022). Increased Belief Instability in Psychotic Disorders Predicts Treatment Response to Metacognitive Training. *Schizophrenia Bulletin*, *48*(4), 826–838. https://doi.org/10.1093/schbul/sbac029

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs.

Heinz, A. (2002). Dopaminergic dysfunction in alcoholism and schizophrenia - Psychopathological and behavioral correlates. *European Psychiatry*. https://doi.org/10.1016/S0924-9338(02)00628-4

Heinz, A., Goldman, D., Jones, D. W., Palmour, R., Hommer, D., Gorey, J. G., … Weinberger, D. R. (2000). Genotype influences in vivo dopamine transporter availability in human striatum. *Neuropsychopharmacology*. https://doi.org/10.1016/S0893-133X(99)00099-8

Heinz, A., Murray, G. K., Schlagenhauf, F., Sterzer, P., Grace, A. A., & Waltz, J. A. (2019). Towards a Unifying Cognitive, Neurophysiological, and Computational Neuroscience Account of Schizophrenia. *Schizophrenia Bulletin*, *45*(5), 1092–1100. https://doi.org/10.1093/schbul/sby154

Heinz, A., & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: Salience attribution revisited. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbq031

Henco, L., Diaconescu, A. O., Lahnakoski, J. M., Brandi, M. L., Hörmann, S., Hennings, J., … Mathys, C. D. (2020). Aberrant computational mechanisms of social learning and decision-making in schizophrenia and borderline personality disorder. *PLoS Computational Biology*, *16*(9), 1–22. https://doi.org/10.1371/journal.pcbi.1008162

Hoekstra, S., Bartz-Johannessen, C., Sinkeviciute, I., Reitan, S. K., Kroken, R. A., Løberg, E. M., … Sommer, I. E. (2021). Sex differences in antipsychotic efficacy and side effects in schizophrenia spectrum disorder: results from the BeSt InTro study. *Npj Schizophrenia*, *7*(1). https://doi.org/10.1038/s41537-021-00170-3

Hoven, M., Lebreton, M., Engelmann, J. B., Denys, D., Luigjes, J., & van Holst, R. J. (2019). Abnormalities of confidence in psychiatry: an overview and future perspectives. *Translational Psychiatry*, *9*(1). https://doi.org/10.1038/s41398-019-0602-7

Howes, O. D., Egerton, A., Allan, V., McGuire, P., Stokes, P., & Kapur, S. (2009). Mechanisms Underlying Psychosis and Antipsychotic Treatment Response in Schizophrenia: Insights from PET and SPECT Imaging. *Current Pharmaceutical Design*, *15*(22), 2550–2559. https://doi.org/10.2174/138161209788957528

Howes, O. D., & Kaar, S. J. (2018). Antipsychotic drugs: challenges and future directions. *World*

Psychiatry, *17*(2), 170–171. https://doi.org/10.1002/wps.20522

Howes, O. D., Kambeitz, J., Kim, E., Stahl, D., Slifstein, M., Abi-Dargham, A., & Kapur, S. (2012). The Nature of Dopamine Dysfunction in Schizophrenia and What This Means for Treatment. *Archives of General Psychiatry*, *69*(8), 776–786. https://doi.org/10.1001/archgenpsychiatry.2012.169

Howes, O. D., McCutcheon, R., Owen, M. J., & Murray, R. M. (2017). The Role of Genes, Stress, and Dopamine in the Development of Schizophrenia. *Biological Psychiatry*, *81*(1), 9–20. https://doi.org/10.1016/j.biopsych.2016.07.014

Howes, O. D., McCutcheon, R., & Stone, J. (2015). Glutamate and dopamine in schizophrenia: An update for the 21st century. *Journal of Psychopharmacology*, *29*(2), 97–115. https://doi.org/10.1177/0269881114563634

Howes, O. D., Montgomery, A. J., Asselin, M. C., Murray, R. M., Valli, I., Tabraham, P., … Grasby, P. M. (2009). Elevated striatal dopamine function linked to prodromal signs of schizophenia. *Archives of General Psychiatry*, *66*(1), 13–20. https://doi.org/10.1001/archgenpsychiatry.2008.514

Howes, O. D., Thase, M. E., & Pillinger, T. (2022). Treatment resistance in psychiatry: state of the art and new directions. *Molecular Psychiatry*, *27*(1), 58–72. https://doi.org/10.1038/s41380-021-01200-3

Huys, Q. J. M., & Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition*, *113*(3), 314–328. https://doi.org/10.1016/j.cognition.2009.01.008

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

Hyman, S. E. (2012, October 10). Revolution stalled. *Science Translational Medicine*. https://doi.org/10.1126/scitranslmed.3003142

Iglesias, S., Kasper, L., Harrison, S. J., Manka, R., Mathys, C. D., & Stephan, K. E. (2021). Cholinergic and dopaminergic effects on prediction error and uncertainty responses during sensory associative learning. *NeuroImage*, *226*(November 2020). https://doi.org/10.1016/j.neuroimage.2020.117590

Iglesias, S., Tomiello, S., Schneebeli, M., & Stephan, K. E. (2017). Models of neuromodulation for computational psychiatry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(3), 1–22. https://doi.org/10.1002/wcs.1420

Ito, H., Kodaka, F., Takahashi, H., Takano, H., Arakawa, R., Shimada, H., & Suhara, T. (2011). Relation between Presynaptic and Postsynaptic Dopaminergic Functions Measured by Positron Emission Tomography: Implication of Dopaminergic Tone. *Journal of Neuroscience*, *31*(21), 7886–7890. https://doi.org/10.1523/JNEUROSCI.6024-10.2011

Ito, Hiroshi, Takahashi, H., Arakawa, R., Takano, H., & Suhara, T. (2008). Normal database of dopaminergic neurotransmission system in human brain measured by positron emission tomography. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2007.09.011

Jaspers, K. (1997). *General Psychopathology*. JHU Press.

Jauhar, S., Nour, M. M., Veronese, M., Rogdaki, M., Bonoldi, I., Azis, M., … Howes, O. D. (2017). A test of the transdiagnostic dopamine hypothesis of psychosis using positron emission tomographic imaging in bipolar affective disorder and schizophrenia. *JAMA Psychiatry*, *74*(12), 1206–1213. https://doi.org/10.1001/jamapsychiatry.2017.2943

Jauhar, S., Veronese, M., Nour, M. M., Rogdaki, M., Hathway, P., Turkheimer, F. E., … Howes, O. D. (2019). Determinants of treatment response in first-episode psychosis: an 18F-DOPA PET study. *Molecular Psychiatry*, *24*(10), 1502–1512. https://doi.org/10.1038/s41380-018-0042-4

Jensen, J., Willeit, M., Zipursky, R. B., Savina, I., Smith, A. J., Menon, M., … Kapur, S. (2008). The formation of abnormal associations in schizophrenia: Neural and behavioral evidence. *Neuropsychopharmacology*, *33*(3), 473–479. https://doi.org/10.1038/sj.npp.1301437

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-Mediated Reinforcement Learning Signals in the Striatum and Ventromedial Prefrontal Cortex Underlie Value-Based Choices. *Journal of Neuroscience*, *31*(5), 1606–1613. https://doi.org/10.1523/JNEUROSCI.3904-10.2011

Jocham, G., Klein, T. A., & Ullsperger, M. (2014). Differential Modulation of Reinforcement Learning by D2 Dopamine and NMDA Glutamate Receptor Antagonism. *Journal of Neuroscience*, *34*(39), 13151–13162. https://doi.org/10.1523/JNEUROSCI.0757-14.2014

Johnson, D. D. P., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, *477*(7364), 317–320.

Jönsson, E. G., Nöthen, M. M., Grünhage, F., Farde, L., Nakashima, Y., Propping, P., & Sedvall, G. C. (1999). Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. *Molecular Psychiatry*, *4*(3), 290–296. https://doi.org/10.1038/sj.mp.4000532

Kaar, S. J., Natesan, S., McCutcheon, R., & Howes, O. D. (2020). Antipsychotics: Mechanisms underlying clinical response and side-effects and novel treatment approaches based on pathophysiology. *Neuropharmacology*, *172*(July 2019), 107704. https://doi.org/10.1016/j.neuropharm.2019.107704

Kahnt, T., Weber, S. C., Haker, H., Robbins, T. W., & Tobler, P. N. (2015). Dopamine D2-receptor blockade enhances decoding of prefrontal signals in humans. *Journal of Neuroscience*, *35*(9), 4104–4111. https://doi.org/10.1523/JNEUROSCI.4182-14.2015

Kakade, S., & Dayan, P. (2002). Acquisition and extinction in autoshaping. *Psychological Review*, *109*(3), 533–544. https://doi.org/10.1037/0033-295X.109.3.533

Kaplan, C. M., Saha, D., Molina, J. L., Hockeimer, W. D., Postell, E. M., Apud, J. A., … Tan, H. Y. (2016). Estimating changing contexts in schizophrenia. *Brain*, *139*(7), 2082–2095. https://doi.org/10.1093/brain/aww095

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*(1), 13–23. https://doi.org/10.1176/appi.ajp.160.1.13

Kapur, S., Agid, O., Mizrahi, R., & Li, M. (2006). How antipsychotics work-from receptors to reality. *NeuroRx : The Journal of the American Society for Experimental NeuroTherapeutics*, *3*(1), 10–21. https://doi.org/10.1016/j.nurx.2005.12.003

Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it. *Molecular Psychiatry*. https://doi.org/10.1038/mp.2012.105

Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine*, *41*(6), 1143–1150. https://doi.org/10.1017/S0033291710001844

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 617. https://doi.org/10.1001/archpsyc.62.6.617

King-Casas, B. (2005). Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange. *Science*, *308*(5718), 78–83. https://doi.org/10.1126/science.1108062

Klaus, J., & Schriefers, H. (2016). Measuring verbal working memory capacity: A reading span task for laboratory and web-based use.

Koob, G. F., & Bloom, F. E. (1988). Cellular and molecular mechanisms of drug dependence. *Science*. https://doi.org/10.1126/science.2903550

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908. https://doi.org/10.1038/s41562-018-0401-9

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When Does Model-Based Control Pay Off? *PLoS Computational Biology*, *12*(8), e1005090. https://doi.org/10.1371/journal.pcbi.1005090

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Korb, S., Götzendorfer, S. J., Massaccesi, C., Sezen, P., Graf, I., Willeit, M., … Silani, G. (2020). Dopaminergic and opioidergic regulation during anticipation and consumption of social and nonsocial rewards. *ELife*, *9*, 1–22. https://doi.org/10.7554/eLife.55797

Köther, U., Veckenstedt, R., Vitzthum, F., Roesch-Ely, D., Pfueller, U., Scheu, F., & Moritz, S. (2012). "Don't give me that look" - Overconfidence in false mental state perception in schizophrenia. *Psychiatry Research*, *196*(1), 1–8. https://doi.org/10.1016/j.psychres.2012.03.004

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

Koutsouleris, N., Hauser, T. U., Skvortsova, V., & De Choudhury, M. (2022). From promise to practice: towards the realisation of AI-informed mental health care. *The Lancet Digital Health*, *4*(11), e829–e840. https://doi.org/10.1016/S2589-7500(22)00153-4

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kroemer, N. B., Lee, Y., Pooseh, S., Eppinger, B., Goschke, T., & Smolka, M. N. (2019). L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action. *NeuroImage*, *186*(February 2018), 113–125. https://doi.org/10.1016/j.neuroimage.2018.10.075

Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, second edition*. *Academic Press*. https://doi.org/10.1016/B978-0-12-405888-0.09999-2

Krystal, J. H., Karper, L. P., Seibyl, J. P., Freeman, G. K., Delaney, R., Bremner, J. D., … Charney, D. S. (1994). Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Archives of General Psychiatry*, *51*(3), 199–214.

Kwak, S., Huh, N., Seo, J. S., Lee, J. E., Han, P. L., & Jung, M. W. (2014). Role of dopamine D2 receptors in optimizing choice strategy in a dynamic and uncertain environment. *Frontiers in Behavioral Neuroscience*, *8*(October), 1–15. https://doi.org/10.3389/fnbeh.2014.00368

L., R., G.D., H., J.R.L., K., H.C., W., A.M., M., L., L., … J., H. (2010). Midbrain activation during pavlovian conditioning and delusional symptoms in schizophrenia. *Archives of General Psychiatry*, *67*(12), 1246–1254. https://doi.org/10.1001/archgenpsychiatry.2010.169

Laakso, A., Pohjalainene, T., Bergman, J., Kajander, J., Haaparanta, M., Solin, O., … Hietala, J. (2005). The A1 allele of the human D2 dopamine receptor gene is associated with increased activity of striatal L-amino acid decarboxylase in healthy subjects. *Pharmacogenetics and Genomics*, *15*(6), 387–391. https://doi.org/10.1097/01213011-200506000-00003

Lahti, A. C., Koffel, B., LaPorte, D., & Tamminga, C. A. (1995). Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. *Neuropsychopharmacology*, *13*(1), 9–19.

Langdon, A. J., Sharpe, M. J., Schoenbaum, G., & Niv, Y. (2018). Model-based predictions for dopamine. *Current Opinion in Neurobiology*, *49*, 1–7. https://doi.org/10.1016/j.conb.2017.10.006

Laurent, V., Leung, B., Maidment, N., & Balleine, B. W. (2012). μ- and δ-Opioid-related processes in the accumbens core and shell differentially mediate the influence of reward-guided and stimulus-guided decisions on choice. *Journal of Neuroscience*, *32*(5), 1875–1883. https://doi.org/10.1523/JNEUROSCI.4688-11.2012

Le Merrer, J., Becker, J. A. J., Befort, K., & Kieffer, B. L. (2009). Reward Processing by the Opioid System in the Brain. *Physiological Reviews*, *89*(4), 1379–1412. https://doi.org/10.1152/physrev.00005.2009

Lee, C. T., Palacios, J., Richards, D., Hanlon, A. K., Lynch, K., Harty, S., … Gillan, C. M. (2023). The Precision in Psychiatry (PIP) study: Testing an internet-based methodology for accelerating research in treatment prediction and personalisation. *BMC Psychiatry*, *23*(1), 1–19. https://doi.org/10.1186/s12888-022-04462-5

Lee, E., Seo, M., Dal Monte, O., & Averbeck, B. B. (2015). Injection of a dopamine type 2 receptor antagonist into the dorsal striatum disrupts choices driven by previous outcomes, but not perceptual inference. *Journal of Neuroscience*, *35*(16), 6298–6306.

Lee, M. C., Wagner, H. N., Tanada, S., Frost, J. J., Bice, A. N., & Dannals, D. F. (1988). Duration of occupancy of opiate receptors by naltrexone. *Journal of Nuclear Medicine*.

Lengersdorff, L., Wagner, I., & Lamm, C. (2020). When implicit prosociality trumps selfishness: the neural valuation system underpins more optimal choices when learning to avoid harm to others than to oneself. *Preprint*. https://doi.org/https://doi.org/10.31234/osf.io/q6psx

Leucht, S., Leucht, C., Huhn, M., Chaimani, A., Mavridis, D., Helfer, B., … Cipriani, A. (2017). Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, Bayesian meta-analysis, and meta-regression of efficacy predictors. *American Journal of Psychiatry*, *174*(10), 927–942.

Lieberman, J. A., Kane, J. M., & Alvir, J. (1987). Provocative tests with psychostimulant drugs in schizophrenia. *Psychopharmacology*, *91*(4), 415–433. https://doi.org/10.1007/BF00216006

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*. https://doi.org/10.1016/j.neuron.2005.05.002

Loosen, A. M., Seow, T. X. F., & Hauser, T. U. (2022). Consistency within change: Evaluating the psychometric properties of a widely-used predictive-inference task. *PsyArXiv*, 62.

Maia, T. V., & Frank, M. J. (2017). An Integrative Perspective on the Role of Dopamine in Schizophrenia. *Biological Psychiatry*, *81*(1), 52–66. https://doi.org/10.1016/j.biopsych.2016.05.021

Mannisto, P. T., & Kaakkola, S. (1999). Catechol-O-methyltransferase (COMT): Biochemistry, Molecular Biology, Pharmacology, and Clinical Efficacy of the New Selective COMT Inhibitors. *Pharmacol. Rev.*, *51*(4), 593–628.

Männistö, P. T., & Kaakkola, S. (1999). Catechol-O-methyltransferase (COMT): Biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. *Pharmacological Reviews*.

Marshall, L., Mathys, C. D., Ruge, D., de Berker, A. O., Dayan, P., Stephan, K. E., & Bestmann, S. (2016). Pharmacological Fingerprints of Contextual Uncertainty. *PLoS Biology*, *14*(11), 1–31. https://doi.org/10.1371/journal.pbio.1002575

Martins, D., Mehta, M. A., & Prata, D. P. (2017). The "highs and lows" of the human brain on dopaminergics: Evidence from neuropharmacology. *Neuroscience and Biobehavioral Reviews*, *80*(April), 351–371. https://doi.org/10.1016/j.neubiorev.2017.06.003

Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy, *78*, 293–296. https://doi.org/10.1016/j.schres.2005.06.020

Mathys, C. D. (2016). How Could We Get Nosology from Computation? *Computational Psychiatry*, *20*. https://doi.org/10.7551/mitpress/9780262035422.003.0007

Mathys, C. D., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, *5*(May), 39. https://doi.org/10.3389/fnhum.2011.00039

Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00825

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–841. https://doi.org/10.1038/nature08028

McCutcheon, R. A., Abi-Dargham, A., & Howes, O. D. (2019). Schizophrenia, Dopamine and the Striatum: From Biology to Symptoms. *Trends in Neurosciences*, *42*(3), 205–220. https://doi.org/10.1016/j.tins.2018.12.004

McElreath, R. (2018). *Statistical rethinking: A bayesian course with examples in* R *and stan*. *Statistical Rethinking: A Bayesian Course with Examples in* R *and Stan*. https://doi.org/10.1201/9781315372495

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

McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2017). Association of the jumping to conclusions and evidence integration biases with delusions in psychosis: A detailed meta-analysis. *Schizophrenia Bulletin*, *43*(2), 344–354. https://doi.org/10.1093/schbul/sbw056

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*(1), 103–107. https://doi.org/10.1038/nn2024

Mehta, M. A., Manes, F. F., Magnolfi, G., Sahakian, B. J., & Robbins, T. W. (2004). Impaired set-shifting and dissociable effects on tests of spatial working memory following the dopamine D2 receptor antagonist sulpiride in human volunteers. *Psychopharmacology*, *176*(3–4), 331–342. https://doi.org/10.1007/s00213-004-1899-2

Mehta, M. A., Montgomery, A. J., Kitamura, Y., & Grasby, P. M. (2008). Dopamine D2 receptor occupancy levels of acute sulpiride challenges that produce working memory and learning impairments in healthy volunteers. *Psychopharmacology*, *196*(1), 157–165. https://doi.org/10.1007/s00213-007-0947-0

Mehta, M. A., Sahakian, B. J., McKenna, P. J., & Robbins, T. W. (1999). Systemic sulpiride in young adult volunteers simulates the profile of cognitive deficits in Parkinson's disease. *Psychopharmacology*, *146*(2), 162–174. https://doi.org/10.1007/s002130051102

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

Meyer-Lindenberg, A., Miletich, R. S., Kohn, P. D., Esposito, G., Carson, R. E., Quarantelli, M., … Berman, K. F. (2002). Reduced prefrontal activity predicts exaggerated striatal dopaminergic function in schizophrenia. *Nature Neuroscience*, *5*(3), 267–271. https://doi.org/10.1038/nn804

Meyer-Lindenberg, A., & Tost, H. (2012). Neural mechanisms of social risk for psychiatric disorders. *Nature Neuroscience*. https://doi.org/10.1038/nn.3083

Mikhael, J. G., & Bogacz, R. (2016). Learning Reward Uncertainty in the Basal Ganglia. *PLoS Computational Biology*, *12*(9), 1–28. https://doi.org/10.1371/journal.pcbi.1005062

Mikus, N., Eisenegger, C., Mathys, C. D., Clark, L., Müller, U., Robbins, T. W., … Naef, M. (2022). Blocking D2/D3 dopamine receptors increases volatility of beliefs when we learn to trust others. *BioRxiv*, 2022.06.21.496956. https://doi.org/10.1101/2022.06.21.496956

Mikus, N., Korb, S., Massaccesi, C., Gausterer, C., Graf, I., Willeit, M., … Mathys, C. D. (2022). Effects of dopamine D2/3 and opioid receptor antagonism on the trade-off between model-based and model-free behaviour in healthy volunteers. *ELife*, *11*, 2022.03.03.482871. https://doi.org/10.7554/eLife.79661

Mikus, N., Mancinelli, F., Lamm, C., & Mathys, C. D. (n.d.). A novel task for measuring sensitivity to loss of control.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic predictive Hebbian learning. *Journal of Neuroscience*.

Montague, P. R., Dolan, J. R., Friston, K. J., & Dayan, P. (2012). Computational Psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80. https://doi.org/10.1017/CBO9781107415324.004

Morawetz, C., Bode, S., Baudewig, J., Jacobs, A. M., & Heekeren, H. R. (2016). Neural representation of emotion regulation goals. *Human Brain Mapping*, *37*(2), 600–620.

https://doi.org/10.1002/hbm.23053

Moritz, S., Ramdani, N., Klass, H., Andreou, C., Jungclaussen, D., Eifler, S., … Zink, M. (2014). Overconfidence in incorrect perceptual judgments in patients with schizophrenia. *Schizophrenia Research: Cognition*, *1*(4), 165–170. https://doi.org/10.1016/j.scog.2014.09.003

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

Murray, G. K., Cheng, F., Clark, L., Barnett, J. H., Blackwell, A. D., Fletcher, P. C., … Jones, P. B. (2008). Reinforcement and reversal learning in first-episode psychosis. *Schizophrenia Bulletin*, *34*(5), 848–855. https://doi.org/10.1093/schbul/sbn078

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., … Fletcher, P. C. (2008). Substantia nigra / ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, *13*(3), 1–18. https://doi.org/10.1038/sj.mp.4002058.Substantia

Naef, M., Müller, U., Linssen, A., Clark, L., Robbins, T. W., & Eisenegger, C. (2017). Effects of dopamine D2/D3 receptor antagonism on human planning and spatial working memory. *Translational Psychiatry*, *7*(4), e1107. https://doi.org/10.1038/tp.2017.56

Nalborczyk, L., Batailler, C., Loevenbruck, H., Vilain, A., & Bürkner, P.-C. (2019). An introduction to bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language, and Hearing Research*, *62*(5), 1225–1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006

Nassar, M. R., Waltz, J. A., Albrecht, M. A., Gold, J. M., & Frank, M. J. (2021). All or nothing belief updating in patients with schizophrenia reduces precision and flexibility of beliefs. *Brain*, 1–17. https://doi.org/10.1093/brain/awaa453

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. *Journal of Neuroscience*, *30*(37), 12366–12378. https://doi.org/10.1523/jneurosci.0822-10.2010

Niv, Y., Daw, N. D., & Joel, D. (2006). Tonic dopamine : opportunity costs and the control of response vigor. https://doi.org/10.1007/s00213-006-0502-4

Nour, M. M., Dahoun, T., Schwartenbeck, P., Adams, R. A., FitzGerald, T. H. B., Coello, C., … Howes, O. D. (2018). Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. *Proceedings of the National Academy of Sciences*, *115*(43). https://doi.org/10.1073/pnas.1809298115

Otto, A. R., Gershman, S. J., & Daw, N. D. (2013). The Curse of Planning: Dissecting multiple reinforcement learning systems by taxing the central executive, *24*(5). https://doi.org/10.1177/0956797612463080.The

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

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports*, *10*(3), 799–812.

Pani, L., Villagrán, J. M., Kontaxakis, V. P., & Alptekin, K. (2008). Practical issues with amisulpride in the management of patients with schizophrenia. *Clinical Drug Investigation*. https://doi.org/10.2165/00044011-200828080-00001

Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational Phenotyping: Using Models to Understand Individual Differences in Personality, Development, and Mental Illness. *Personality Neuroscience*, *1*, e18. https://doi.org/10.1017/pen.2018.14

Patzelt, E. H., Kool, W., Millner, A. J., & Gershman, S. J. (2018). Incentives Boost Model-Based Control Across a Range of Severity on Several Psychiatric Constructs. *Biological Psychiatry*, 1–9. https://doi.org/10.1016/j.biopsych.2018.06.018

Paulus, M. P., Huys, Q. J. M., & Maia, T. V. (2016). A Roadmap for the Development of Applied Computational Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2016.05.001

Peciña, S. (2008). Opioid reward "liking" and "wanting" in the nucleus accumbens. *Physiology and Behavior*, *94*(5), 675–680. https://doi.org/10.1016/j.physbeh.2008.04.006

Peciña, S., & Berridge, K. C. (2013). Dopamine or opioid stimulation of nucleus accumbens similarly amplify cue-triggered "wanting" for reward: Entire core and medial shell mapped as substrates for PIT enhancement. *European Journal of Neuroscience*, *37*(9), 1529–1540. https://doi.org/10.1111/ejn.12174

Pelloux, Y., Dilleen, R., Economidou, D., Theobald, D., & Everitt, B. J. (2012). Reduced forebrain serotonin transmission is causally involved in the development of compulsive cocaine seeking in rats. *Neuropsychopharmacology*, *37*(11), 2505–2514.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, J. R., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. https://doi.org/10.1038/nature05051

Peters, E., Joseph, S., Day, S., & Garety, P. A. (2004). Measuring delusional ideation: The 21-item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin*, *30*(4), 1005–1022. https://doi.org/10.1093/oxfordjournals.schbul.a007116

Petersen, N., Rapkin, A. J., Okita, K., Kinney, K. R., Mizuno, T., Mandelkern, M. A., & London, E. D. (2021). Striatal dopamine D2-type receptor availability and peripheral 17β-estradiol. *Molecular Psychiatry*, *26*(6), 2038–2047. https://doi.org/10.1038/s41380-020-01000-1

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2007). Linear and nonlinear mixed effects models. *R Package Version*, *3*(57), 1–89.

Piray, P., & Daw, N. D. (2021). A model for learning based on the joint estimation of stochasticity and volatility. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-26731-9

Potkin, S. G., Kane, J. M., Correll, C. U., Lindenmayer, J. P., Agid, O., Marder, S. R., … Howes, O. D. (2020). The neurobiology of treatment-resistant schizophrenia: paths to antipsychotic resistance and a roadmap for future research. *Npj Schizophrenia*, *6*(1). https://doi.org/10.1038/s41537-019-0090-z

Pulcu, E., & Browning, M. (2019). The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences*, *23*(10), 865–875. https://doi.org/10.1016/j.tics.2019.07.007

Quednow, B. B., & Herdener, M. (2016). Human pharmacology for addiction medicine: From evidence to clinical recommendations. In *Progress in Brain Research*. https://doi.org/10.1016/bs.pbr.2015.07.017

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Reed, E. J., Uddenberg, S., Suthaharan, P., Mathys, C. D., Taylor, J. R., Groman, S. M., & Corlett, P. R. (2020). Paranoia as a deficit in non-social belief updating. *ELife*, *9*, 1–55. https://doi.org/10.7554/eLife.56345

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: current research and theory, Vol. 2*. https://doi.org/10.1101/gr.110528.110

Richfield, E. K., Penney, J. B., & Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience*. https://doi.org/10.1016/0306-4522(89)90168-1

Riederer, P., Laux, G., Nagatsu, T., Le, W., & Riederer, C. (2022). *NeuroPsychopharmacotherapy*. Springer Nature.

Roiser, J. P., Howes, O. D., Chaddock, C. A., Joyce, E. M., & McGuire, P. (2013). Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophrenia Bulletin*, *39*(6), 1328–1336. https://doi.org/10.1093/schbul/sbs147

Roiser, J. P., Stephan, K. E., Den Ouden, H., Friston, K. J., & Joyce, E. M. (2010). Adaptive and aberrant reward prediction signals in the human brain. *NeuroImage*, *50*(2), 657–664. https://doi.org/10.1016/j.neuroimage.2009.11.075

Rosenberger, L. A., Eisenegger, C., Naef, M., Terburg, D., Fourie, J., Stein, D. J., & van Honk, J. (2019). The Human Basolateral Amygdala Is Indispensable for Social Experiential Learning. *Current Biology*, *29*(20), 3532–3537.

Rosenzweig, P., Canal, M., Patat, A., Bergougnan, L., Zieleniuk, I., & Bianchetti, G. (2002). A review of the pharmacokinetics, tolerability and pharmacodynamics of amisulpride in healthy volunteers. *Human Psychopharmacology*, *17*(1), 1–13. https://doi.org/10.1002/hup.320

Rossi-Goldthorpe, R. A., Leong, Y. C., Leptourgos, P., & Corlett, P. R. (2021). Paranoia, self-deception and overconfidence. *PLOS Computational Biology*, *17*(10), e1009453. https://doi.org/10.1371/journal.pcbi.1009453

Rush, A. J., Trivedi, M., Carmody, T. J., Biggs, M. M., Shores-Wilson, K., Ibrahim, H., & Crismon, M. L. (2004). One-year clinical outcomes of depressed public sector outpatients: A benchmark for subsequent studies. *Biological Psychiatry*, *56*(1), 46–53. https://doi.org/10.1016/j.biopsych.2004.04.005

Rush, C. R., Stoops, W. W., Hays, L. R., Glaser, P. E. A., & Hays, L. S. (2003). Risperidone attenuates the discriminative-stimulus effects of d-amphetamine in humans. *Journal of Pharmacology and Experimental Therapeutics*, *306*(1), 195–204.

Rutledge, R. B., Chekroud, A. M., & Huys, Q. J. M. (2019). Machine learning and big data in psychiatry: toward clinical applications. *Current Opinion in Neurobiology*, *55*, 152–159. https://doi.org/10.1016/j.conb.2019.02.006

Salamone, J. D., Correa, M., Yohn, S., Lopez Cruz, L., San Miguel, N., & Alatorre, L. (2016). The pharmacology of effort-related choice behavior: Dopamine, depression, and individual differences. *Behavioural Processes*, *127*, 3–17. https://doi.org/10.1016/j.beproc.2016.02.008

Sawaguchi, T., & Goldman-Rakic, P. (1991). D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science*. https://doi.org/10.1126/science.1825731

Schennach, R., Riedel, M., Obermeier, M., Spellmann, I., Musil, R., Jäger, M., … Möller, H. J. (2015). What are residual symptoms in schizophrenia spectrum disorder? Clinical description and 1-year persistence within a naturalistic trial. *European Archives of Psychiatry and Clinical Neuroscience*, *265*(2), 107–116. https://doi.org/10.1007/s00406-014-0528-2

Schlagenhauf, F., Huys, Q. J. M., Deserno, L., Rapp, M. A., Beck, A., Heinze, H. J., … Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *NeuroImage*, *89*, 171–180. https://doi.org/10.1016/j.neuroimage.2013.11.034

Schlagenhauf, F., Rapp, M. A., Huys, Q. J. M., Beck, A., Wüstenberg, T., Deserno, L., … Heinz, A. (2013). Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence. *Human Brain Mapping*, *34*(6), 1490–1499. https://doi.org/10.1002/hbm.22000

Schmack, K., Schnack, A., Priller, J., & Sterzer, P. (2015). Perceptual instability in schizophrenia: Probing predictive coding accounts of delusions with ambiguous stimuli. *Schizophrenia Research: Cognition*, *2*(2), 72–77. https://doi.org/10.1016/j.scog.2015.03.005

Schoemaker, H., Claustre, Y., Fage, D., Rouquier, L., Chergui, K., Curet, O., … Scatton, B. (1997). Neurochemical characteristics of amisulpride, an atypical dopamine D2/D3 receptor antagonist with both presynaptic and limbic selectivity. *The Journal of Pharmacology and Experimental Therapeutics*, *280*(1), 83–97. https://doi.org/10.1002/mc.20769

Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology*, (80), 1–

27.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*. https://doi.org/10.1126/science.275.5306.1593

Schultz, W., Preuschoff, K., Camerer, C. F., Hsu, M., Fiorillo, C. D., Tobler, P. N., & Bossaerts, P. (2008). Explicit neural signals reflecting reward uncertainty. *Philos Trans R Soc Lond B Biol Sci*, *363*(1511), 3801–3811. https://doi.org/10.1098/rstb.2008.0152

Schwartenbeck, P., FitzGerald, T. H. B., Mathys, C. D., Dolan, J. R., & Friston, K. J. (2015). The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cerebral Cortex*, *25*(10), 3434–3445. https://doi.org/10.1093/cercor/bhu159

Seamans, J. K., & Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in Neurobiology*, *74*(1), 1–58. https://doi.org/10.1016/j.pneurobio.2004.05.006

Seow, T. X. F., & Gillan, C. M. (2020). Transdiagnostic Phenotyping Reveals a Host of Metacognitive Deficits Implicated in Compulsivity. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-59646-4

Seth, P., Scholl, L., Rudd, R. A., & Bacon, S. (2018). Overdose deaths involving opioids, cocaine, and psychostimulants — United States, 2015–2016. *American Journal of Transplantation*, *18*(6), 1556–1568. https://doi.org/10.1111/ajt.14905

Sharp, C., Monterosso, J., & Montague, P. R. (2012). Neuroeconomics: A bridge for translational research. *Biological Psychiatry*, *4*(72), 87–92. https://doi.org/10.1126/scisignal.2001449.Engineering

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

Siegel, J. Z., Mathys, C. D., Rutledge, R. B., & Crockett, M. J. (2018). Beliefs about bad people are volatile. *Nature Human Behaviour*, *2*(10), 750–756. https://doi.org/10.1038/s41562-018-0425-1

Slifstein, M., Kolachana, B., Simpson, E. H., Tabares, P., Cheng, B., Duvall, M., … Abi-Dargham, A. (2008). COMT genotype predicts cortical-limbic D1 receptor availability measured with [11C]NNC112 and PET. *Molecular Psychiatry*. https://doi.org/10.1038/mp.2008.19

Smith, C. T., Dang, L. C., Buckholtz, J. W., Tetreault, A. M., Cowan, R. L., Kessler, R. M., & Zald, D. H. (2017). The impact of common dopamine D2 receptor gene polymorphisms on D2/3 receptor availability: C957T as a key determinant in putamen and ventral striatum. *Translational Psychiatry*, *7*(4). https://doi.org/10.1038/tp.2017.45

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., & Dolan, J. R. (2013). Disruption of Dorsolateral Prefrontal Cortex Decreases Model-Based in Favor of Model-free Control in Humans. *Neuron*, *80*(4), 914–919. https://doi.org/10.1016/j.neuron.2013.08.009

Snyder, S. H. (1976). The dopamine hypothesis of schizophrenia: focus on the dopamine receptor. *The American Journal of Psychiatry*, *133*(2), 197–202.

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., … Rodrigues, A. J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. *Nature Communications*, *7*(May). https://doi.org/10.1038/ncomms11829

Soares-Cunha, C., Coimbra, B., Sousa, N., & Rodrigues, A. J. (2016). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neuroscience and Biobehavioral Reviews*, *68*, 370–386. https://doi.org/10.1016/j.neubiorev.2016.05.021

Soutschek, A., Weber, S. C., Kahnt, T., Quednow, B. B., & Tobler, P. N. (2021). Opioid antagonism modulates wanting-related frontostriatal connectivity. *ELife*, *10*, 1–17. https://doi.org/10.7554/eLife.71077

Sridharan, D., Prashanth, P. S., & Chakravarthy, V. S. (2006). The role of the basal ganglia in exploration

in A neural model based on reinforcement learning. *International Journal of Neural Systems*, *16*(2), 111–124. https://doi.org/10.1142/S0129065706000548

Stalnaker, T. A., Howard, J. D., Takahashi, Y. K., Gershman, S. J., Kahnt, X. T., & Schoenbaum, G. (2019). Dopamine neuron ensembles signal the content of sensory prediction errors. *ELife*, *8*, 1–16. https://doi.org/10.7554/eLife.49315

Stephan, K. E., Bach, D. R., Fletcher, P. C., Flint, J., Frank, M. J., Friston, K. J., … Breakspear, M. (2016). Charting the landscape of priority problems in psychiatry, part 1: Classification and diagnosis. *The Lancet Psychiatry*, *3*(1), 77–83. https://doi.org/10.1016/S2215-0366(15)00361-2

Stephan, K. E., Baldeweg, T., & Friston, K. J. (2006). Synaptic Plasticity and Dysconnection in Schizophrenia. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2005.10.005

Stephan, K. E., & Mathys, C. D. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology*, *25*, 85–92. https://doi.org/10.1016/j.conb.2013.12.007

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017. https://doi.org/10.1016/j.neuroimage.2009.03.025

Sterzer, P., Adams, R. A., Fletcher, P. C., Frith, C. D., Lawrie, S. M., Muckli, L., … Corlett, P. R. (2018). The Predictive Coding Account of Psychosis. *Biological Psychiatry*, *84*(9), 634–643. https://doi.org/10.1016/j.biopsych.2018.05.015

Strauss, G. P., Frank, M. J., Waltz, J. A., Kasanova, Z., Herbener, E. S., & Gold, J. M. (2011). Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biological Psychiatry*, *69*(5), 424–431. https://doi.org/10.1016/j.biopsych.2010.10.015

Sutton, R. S. (1992). Gain adaptation beats least squares. *Proceedings of the 7th Yale Workshop on* {…}, 161–166. Retrieved from http://webdocs.cs.ualberta.ca/~sutton/papers/sutton-92b.pdf

Sutton, R. S., & Barto, A. G. (2017). Reinforcement learning: an introduction. *UCL,Computer Science Department, Reinforcement Learning Lectures*, 1054. https://doi.org/10.1109/TNN.1998.712192

Svenningsson, P., Nishi, A., Fisone, G., Girault, J.-A., Nairn, A. C., & Greengard, P. (2004). DARPP-32: An Integrator of Neurotransmission. *Annual Review of Pharmacology and Toxicology*. https://doi.org/10.1146/annurev.pharmtox.44.101802.121415

Takano, A., Suhara, T., Yasuno, F., Suzuki, K., Takahashi, H., Morimoto, T., … Okubo, Y. (2006). The antipsychotic sultopride is overdosed - A PET study of drug-induced receptor occupancy in comparison with sulpiride. *International Journal of Neuropsychopharmacology*, *9*(5), 539–545. https://doi.org/10.1017/S1461145705006103

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

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642–1645. https://doi.org/10.1126/science.1105370

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*. https://doi.org/10.1037/h0061626

Tomassini, A., Ruge, D., Galea, J. M., Penny, W., & Bestmann, S. (2016). The role of dopamine in temporal uncertainty. *Journal of Cognitive Neuroscience*, *28*(1), 96–110. https://doi.org/10.1162/jocn_a_00880

Torous, J., Myrick, K., & Aguilera, A. (2023). The need for a new generation of digital mental health tools to support more accessible, effective and equitable care. *World Psychiatry*, *22*(1), 1–2. https://doi.org/10.1002/wps.21058

Trantham-Davidson, H., Neely, L. C., Lavin, A., & Seamans, J. K. (2004). Mechanisms underlying differential D1 versus D2 dopamine receptor regulation of inhibition in prefrontal cortex. *Journal of Neuroscience*, *24*(47), 10652–10659. https://doi.org/10.1523/JNEUROSCI.3179-04.2004

Trøstheim, M., Eikemo, M., Haaker, J., Frost, J. J., & Leknes, S. (2022). Opioid Antagonism in Humans: A Primer on Optimal Dose and Timing for Central Mu-Opioid Receptor Blockade. *BioRxiv*, 2022.02.25.481943. https://doi.org/10.1101/2022.02.25.481943

Upthegrove, R., Marwaha, S., & Birchwood, M. (2017). Depression and Schizophrenia: Cause, Consequence, or Trans-diagnostic Issue? *Schizophrenia Bulletin*, *43*(2), 240–244. https://doi.org/10.1093/schbul/sbw097

Valenti, O., Cifelli, P., Gill, K. M., & Grace, A. A. (2011). Antipsychotic drugs rapidly induce dopamine neuron depolarization block in a developmental rat model of schizophrenia. *Journal of Neuroscience*, *31*(34), 12330–12338. https://doi.org/10.1523/JNEUROSCI.2808-11.2011

van Opheusden, B., Kuperwajs, I., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2023). Expertise increases planning depth in human gameplay. *Nature*, 1–6.

van Schouwenburg, M., Aaron, ;, & Cools, R. (2010). Dopaminergic Modulation of Cognitive Control: Distinct Roles for the Prefrontal Cortex and the Basal Ganglia. *Current Pharmaceutical Design*, *16*(18), 2026–2032. https://doi.org/10.2174/138161210791293097

van Steenbergen, H., Eikemo, M., & Leknes, S. (2019). The role of the opioid system in decision making and cognitive control: A review. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(3), 435–458. https://doi.org/10.3758/s13415-019-00710-6

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

Volavka, J., Dornbush, R., Mallya, A., & Cho, D. (1979). Naloxone fails to affect short-term memory in man. *Psychiatry Research*, *1*(1), 89–92.

Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln* (Vol. 9). Voss.

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., … Bullmore, E. T. (2015). Disorders of compulsivity: a common bias towards learning habits. *Molecular Psychiatry*, *20*(3), 345–352. https://doi.org/10.1038/mp.2014.44

Voon, V., Joutsa, J., Majuri, J., Baek, K., Nord, C. L., Arponen, E., … Kaasinen, V. (2020). The neurochemical substrates of habitual and goal-directed control. *Translational Psychiatry*, *10*(1). https://doi.org/10.1038/s41398-020-0762-5

Voon, V., Reiter, A. M. F., Sebold, M., & Groman, S. M. (2016). Model-Based Control in Dimensional Psychiatry. *Biological Psychiatry*, *82*(6), 391–400. https://doi.org/10.1016/j.biopsych.2017.04.006

Waltmann, M., Schlagenhauf, F., & Deserno, L. (2022). Sufficient reliability of the behavioral and computational readouts of a probabilistic reversal learning task. *Behavior Research Methods*, *54*(6), 2993–3014. https://doi.org/10.3758/s13428-021-01739-7

Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: Further evidence of orbitofrontal dysfunction. *Schizophrenia Research*, *93*(1–3), 296–303. https://doi.org/10.1016/j.schres.2007.03.010

Wassum, K. M., Cely, I. C., Maidment, N. T., & Balleine, B. W. (2009). Disruption of endogenous opioid activity during instrumental learning enhances habit acquisition. *Neuroscience*, *163*(3), 770–780. https://doi.org/https://doi.org/10.1016/j.neuroscience.2009.06.071

Weber, L. A., Waade, P. T., Legrand, N., Møller, A. H., Stephan, K. E., & Mathys, C. D. (2023). The generalized Hierarchical Gaussian Filter, 1–38. Retrieved from http://arxiv.org/abs/2305.10937

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. *Translational Psychiatry*, *6*(7), e850. https://doi.org/10.1038/tp.2016.113

Weilnhammer, V., Röd, L., Eckert, A. L., Stuke, H., Heinz, A., & Sterzer, P. (2020). Psychotic experiences in schizophrenia and sensitivity to sensory evidence. *Schizophrenia Bulletin*, *46*(4), 927–936. https://doi.org/10.1093/schbul/sbaa003

Wellstein, K. V., Diaconescu, A. O., Bischof, M., Rüesch, A., Paolini, G., Aponte, E. A., … Stephan, K. E. (2020). Inflexible social inference in individuals with subclinical persecutory delusional tendencies. *Schizophrenia Research*, *215*, 344–351. https://doi.org/10.1016/j.schres.2019.08.031

Westbrook, A., Bosch, R. van den, 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. *Science*, *367*(6484), 1362–1366. https://doi.org/10.1126/SCIENCE.AAZ5891

Westbrook, A., Frank, M. J., & Cools, R. (2021). A mosaic of cost–benefit control over cortico-striatal circuitry. *Trends in Cognitive Sciences*, *25*(8), 710–721. https://doi.org/10.1016/j.tics.2021.04.007

Wiesel, F.-A., Alfredsson, G., Ehrnebo, M., & Sedvall, G. (1982). Prolactin response following intravenous and oral sulpiride in healthy human subjects in relation to sulpiride concentrations. *Psychopharmacology*, *76*(1), 44–47.

Williams, G. V., & Goldman-Rakic, P. S. (1995). Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature*. https://doi.org/10.1038/376572a0

Winton-Brown, T. T., Fusar-Poli, P., Ungless, M. A., & Howes, O. D. (2014). Dopaminergic basis of salience dysregulation in psychosis. *Trends in Neurosciences*, *37*(2), 85–94. https://doi.org/10.1016/j.tins.2013.11.003

Woodward, T. S., Moritz, S., Menon, M., & Klinge, R. (2008). Belief inflexibility in schizophrenia. *Cognitive Neuropsychiatry*, *13*(3), 267–277. https://doi.org/10.1080/13546800802099033

Worbe, Y., Savulich, G., de Wit, S., Fernandez-Egea, E., & Robbins, T. W. (2015). Tryptophan depletion promotes habitual over goal-directed control of appetitive responding in humans. *International Journal of Neuropsychopharmacology*, *18*(10).

World Medical Association Declaration of Helsinki. (2013). *JAMA*, *310*(20), 2191. https://doi.org/10.1001/jama.2013.281053

Wunderlich, K., Smittenaar, P., & Dolan, J. R. (2012). Dopamine Enhances Model-Based over Model-Free Choice Behavior. *Neuron*, *75*(3), 418–424. https://doi.org/10.1016/j.neuron.2012.03.042

Yao, W. D., Spealman, R. D., & Zhang, J. (2008). Dopaminergic signaling in dendritic spines. *Biochemical Pharmacology*. https://doi.org/10.1016/j.bcp.2008.01.018

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using Stacking to Average Bayesian Predictive Distributions (with Discussion). *Bayesian Analysis*. https://doi.org/10.1214/17-ba1091

Zacny, J. P., Coalson, D. W., Lichtor, J. L., Yajnik, S., & Thapar, P. (1994). Effects of naloxene on the subjective and psychomotor effects of nitrous oxide in humans. *Pharmacology Biochemistry and Behavior*, *49*(3), 573–578.

Zakzanis, K. K., & Hansen, K. T. (1998). Dopamine D2 densities and the schizophrenic brain. *Schizophrenia Research*, *32*(3), 201–206.

Zech, H., Waltmann, M., Lee, Y., Reichert, M., Bedder, R. L., Rutledge, R. B., … Deserno, L. (2022). Measuring self-regulation in everyday life: Reliability and validity of smartphone-based experiments in alcohol use disorder. *Behavior Research Methods*, (0123456789). https://doi.org/10.3758/s13428-022-02019-8

Zhang, J.-P., Lencz, T., & Malhotra, A. K. (2010). D2 receptor genetic variation and clinical response to antipsychotic drug treatment: A meta-analysis. *American Journal of Psychiatry*, *167*(7), 763–772. https://doi.org/10.1176/appi.ajp.2009.09040598

Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls, and suggestions. *Social Cognitive and Affective Neuroscience*.

Annexes

Abstract

Psychiatric drugs are widely prescribed in every corner of the world and are an essential part of most psychiatric treatment plans. However, we have a very poor understanding of how and why they work. As a result, there are no validated tests in clinical practice to guide treatment selection or enable treatment monitoring. Through generative computational models we can formalize a mechanistic link between mental health symptoms and the molecular targets of psychiatric drugs and elucidate how neurobiological alterations contribute to psychopathology. In the work presented in this dissertation, we have described several computational mechanisms that underlie false inferences and beliefs in individuals with psychotic experiences and investigated how these mechanisms are affected by antipsychotic medication in healthy volunteers.

Psychotic symptoms are characterized by a profound loss of touch with reality, rigid delusional beliefs, and hallucinations. All currently approved antipsychotics block D2 dopamine receptors, but how this leads to symptom relief is largely an open question. In the first paper of this thesis, we investigated how psychotic symptoms in patients and psychotic-like experiences in healthy volunteers relate to various computational phenotypes that describe inter-individual differences in how participants update beliefs about a changing variable from noisy observations. Using a novel hierarchical Bayesian model, we found that psychotic traits across the clinical severity spectrum were related to increased uncertainty about the volatility of the task, suggesting an impaired ability to build or use higher-order models for prediction and planning. In contrast, negative symptoms in patients and traits related to depression and anxiety correlated with rigid beliefs about the underlying mean and noise in observation.

In two subsequent studies we then investigated how selective D2 receptor antagonists affect belief updating and model-based planning in healthy volunteers. First, we found that blocking D2 receptors reduced the rigidity of beliefs when learning about the prosocial attitudes of others in a repeated Trust game. This effect was particularly pronounced in individuals with genetic markers for higher dopamine availability. In the following study, we used an experimental paradigm designed to measure the capacity for model-based learning, and directly compared the effects of a D2 antagonist with those of an antagonist of the opioid system – a system known to influence habit formation. We found that blocking dopamine D2 receptors, but not opioid receptors improved model-based behavior.

These studies suggest that D2 antagonism may allow more flexible adaptation of higher-order representations. D2 receptor occupancy by antipsychotic drugs may therefore render delusional beliefs more malleable while promoting the formation of healthier world models. The thesis provides a proof of concept of how computational models can bridge multiple levels of description and define clinically relevant subgroups, while demonstrating how computational techniques can be leveraged to improve our understanding of how pharmacological treatment leads to symptom relief.

Zusammenfassung

Psychopharmaka werden weltweit verschrieben und sind ein wesentlicher Bestandteil der meisten psychiatrischen Behandlungspläne. Wir wissen jedoch sehr wenig darüber, wie und warum sie wirken. Daher gibt es in der klinischen Praxis keine validierten Tests, die bei der Auswahl der Behandlung helfen oder die Überwachung der Behandlung ermöglichen. Mit Hilfe generativer Computermodelle können wir einen mechanistischen Zusammenhang zwischen psychischen Symptomen und den molekularen Angriffspunkten von Psychopharmaka herstellen und aufklären, wie neurobiologische Veränderungen zur Psychopathologie beitragen. In dieser Dissertation haben wir mehrere Rechenmechanismen beschrieben, die falschen Schlussfolgerungen und Überzeugungen bei Personen mit psychotischen Erfahrungen zugrunde liegen, und untersucht, wie diese Mechanismen durch antipsychotische Medikamente bei gesunden Freiwilligen beeinflusst werden.

Psychotische Symptome sind durch einen tiefgreifenden Realitätsverlust, Wahnvorstellungen und Halluzinationen gekennzeichnet. Alle derzeit zugelassenen Antipsychotika blockieren die D2-Dopaminrezeptoren, aber wie dies zu einer Linderung der Symptome führt, ist weitgehend ungeklärt. In der ersten Arbeit dieser Dissertation untersuchten wir, wie psychotische Symptome bei Patienten und psychoseähnliche Erfahrungen bei gesunden Probanden mit verschiedenen computergestützten Phänotypen zusammenhängen, die interindividuelle Unterschiede in der Art und Weise beschreiben, wie Teilnehmer ihre Überzeugungen über eine sich verändernde Variable aus verrauschten Beobachtungen aktualisieren. Unter Verwendung eines neuartigen hierarchischen Bayesschen Modells fanden wir, dass psychotische Merkmale über das gesamte klinische Schweregradspektrum hinweg mit erhöhter Unsicherheit über die Volatilität der Aufgabe assoziiert sind. Im Gegensatz dazu korrelierten negative Patientensymptome und Merkmale im Zusammenhang mit Depression und Angst mit starren Überzeugungen über den zugrunde liegenden Mittelwert und Rauschen in der Beobachtung.

In zwei weiteren Studien untersuchten wir, wie selektive D2-Rezeptor-Antagonisten die Aktualisierung von Überzeugungen und die modellbasierte Planung bei gesunden Freiwilligen beeinflussen. Zunächst stellten wir fest, dass die Blockade der D2-Rezeptoren die Rigidität der Überzeugungen beim Lernen über die prosozialen Einstellungen anderer in einem wiederholten Vertrauensspiel verringerte. Dieser Effekt war besonders ausgeprägt bei Personen mit genetischen Markern für eine erhöhte Dopaminverfügbarkeit. In der folgenden Studie verwendeten wir ein experimentelles Paradigma, um die Fähigkeit zum modellbasierten Lernen zu messen, und verglichen die Effekte eines D2-Antagonisten direkt mit denen eines Antagonisten des Opioidsystems - ein System, von dem bekannt ist, dass es die Gewohnheitsbildung beeinflusst. Wir fanden heraus, dass die Blockade der Dopamin-D2-Rezeptoren, nicht aber der Opioidrezeptoren, das modellbasierte Verhalten verbesserte.

Diese Studien legen nahe, dass der D2-Antagonismus eine flexiblere Anpassung von Repräsentationen höherer Ordnung ermöglichen könnte. Die Besetzung von D2-Rezeptoren durch antipsychotische Medikamente könnte daher wahnhafte Überzeugungen formbarer machen und gleichzeitig die Bildung gesünderer Weltmodelle fördern. Die Arbeit liefert einen Konzeptnachweis dafür, wie Computermodelle mehrere Beschreibungsebenen überbrücken und klinisch relevante Untergruppen definieren können, und zeigt gleichzeitig, wie Computertechniken eingesetzt werden können, um unser Verständnis dafür zu verbessern, wie pharmakologische Behandlung zu Symptomlinderung führt.