DOMAIN-GENERAL VERSUS DOMAIN-SPECIFIC LEARNING MECHANISMS: NEUROCHEMICAL MECHANISMS AND RELEVANCE TO AUTISM

by

ALICIA JOHANNA RYBICKI

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

Life and Environmental Sciences

University of Birmingham

March 2022

Abstract

The theory that various features of autism spectrum disorders (ASD) can be explained by differences in the learning (or "predictive coding") process is growing in popularity. However, extant studies have focused on the domain of sensory perception, i.e., learning what to expect in the visual or auditory domains. It is thus unclear whether such models are restricted to the perceptual domain, or whether they are outlining differences in domain-general learning processes. Consequently, how such theories can explain the social and motor features of ASD is currently unclear. The first part of the current thesis asks whether autistic adults exhibit differences, compared to non-autistic adults, with respect to social learning and motor learning. The second part of this thesis focuses in detail on one of these learning types - social learning. Here I investigate the neurochemical mechanisms that underpin social learning and ask whether they are dissociable from the neurochemical mechanisms that underpin learning from one's own individual experience (individual learning). In integrating these results with the wider literature, I reflect upon the broader question of whether there are common domain-general learning mechanisms, or domain (e.g., social, motor, individual) specific learning "modules". Together the studies presented in this thesis implicate the dopaminergic neurotransmitter system in both social and individual learning. Results support the view that there are domain-general neurochemical mechanisms that support various types of learning. These results do not, however, support the view that autistic adults exhibit differences in these domain-general learning processes. That is, our empirical work showed no differences in either social or motor learning when comparing autistic and non-autistic adults. These results do not add support for impaired predictive coding as a core deficit that can explain social and motor atypicalities in autism, but rather force us to think more critically about what overarching conclusions can be drawn from studies of predictive coding in autism within the perception domain.

Acknowledgments

First, the biggest thanks go to my supervisor Jennifer Cook for supporting and motivating me throughout my PhD work. Her help and guidance have been invaluable throughout my PhD work, and it has been a great pleasure to work in the Cook lab. Jen's expertise and passion for psychology have been inspiring and she has continuously encouraged and helped me over the last few years.

Many thanks also to my second supervisor Joseph Galea, for their encouragement and help with the development and interpretation of my work.

A huge thanks to everyone in the Cook lab (both new and old members!) - Sophie Sowden, Bianca Schuster, Lydia Hickman, Connor Keating, Dagmar Fraser, Hélio Clemente Cuve and Molly Gracey. It has been great to work with you all and I hope the interesting lab meeting and collaborations can continue in the future. A special thanks to Sophie, Bianca, Lydia, and Connor for being fantastic office companions, I only wish that we had had more time in the office together.

I would like to thank MIBTP and the BBSRC for granting me a studentship and supporting the completion of this thesis. I would also like to thank all the research participants who took part in the empirical studies in this thesis, without whom this work would not have been possible.

Finally, a big thanks to all my family and friends, for their support and belief in me, and a very special thanks to Jon.

About this thesis

This thesis incorporates published papers. Chapters 2 and 6 include copies of published journal articles, available on the relevant journal websites - DOIs are included for published work. For ease of reference, and to comply with University of Birmingham guidelines, separate pagination is included for chapters containing published work and these chapters are reproduced in the journal format. Chapters 3-5, as well as the general introduction and discussion chapters, were written specifically for this thesis. All studies were approved by local research ethics committees and are in accord with the Declaration of Helsinki.

Author contributions for published work

Chapter 2: Study design was conducted by myself, Jennifer Cook and Joseph Galea. The experimental paradigm was programmed by Jennifer Cook and Joseph Galea. I carried out participant recruitment and data collection, with help from Chole Hiles, Cleo Fabian, Bianca Schuster, and Jennifer Cook. I conducted all data analysis and drafted the manuscript for publication, which was then edited by Jennifer Cook and approved by all co-authors.

Chapter 6: Study design was conducted by myself and Jennifer Cook. I created the stimuli and programmed the experimental paradigm. I carried out recruitment and data collection for this study, alongside Bianca Schuster, Sophie Sowden, and Jennifer Cook. I collected and analysed the data and data interpretation was carried out by myself and Jennifer Cook. I drafted the manuscript for publication which was then edited by Jennifer Cook and approved by all co-authors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

Acetylcholine (ACh)

Anterior cingulate cortex gyrus (ACCg)

Anterior cingulate cortex sulcus (ACCs)

Autism Diagnostic Observation Schedule, second edition (ADOS-2)

Autism spectrum disorders (ASD)

Autism Quotient (AQ)

Bayesian model selection (BMS)

Becks Depression Index (BDI)

Blood oxygenation level-dependent (BOLD)

Catechol-O-Methyltransferase (COMT)

Diagnostic and Statistical Manual of Mental Disorders (DSM–5)

Dopamine (DA)

Dopamine transporter (DAT)

Dopamine transporter gene (DAT1/SLC6A3)

Dopamine receptor 2 (D2)

Dorsomedial prefrontal cortex (dmPFC)

Functional magnetic resonance imaging (fMRI)

Haloperidol (HAL)

Hierarchical Gaussian Filter (HGF)

International Classification of Diseases, Tenth Revision (ICD-10)

Inverse Efficiency scores (IES)

Maximum-a-posteriori (MAP)

Medial prefrontal cortex (mPFC)

Methylphenidate (MPH)

Nijmegen Brain Imaging Genetics project (BIG)

Norepinephrine (NE)

Nucleus accumbens (NAcc)

Open Science Framework (OSF)

Primary motor cortex (M1)

Prediction error (PE)

Reaction time (RT)

Reinforcement learning (RL)

Rescorla-Wagner (RW)

Serotonin (5-HT)

Serotonin reuptake transporter (SERT/5-HTT)

Serotonin reuptake transporter gene (SERT1/SLC6A4)

Serotonin-transporter-linked promoter region (5-HTTLPR)

Shapiro-Wilk (SW)

Single nucleotide polymorphism (SNP)

Social learning task (SLT)

Substantia nigra (SN)

Temporoparietal junction (TPJ)

Translational Algorithms for Psychiatry-Advancing Science (TAPAS)

Variable number of tandem repeats (VNTR)

Ventral tegmental area (VTA)

Ventromedial prefrontal cortex (vmPFC)

Verbal working memory span (VWM)

Wechsler Adult Intelligence Scale | Fourth Edition (WAIS-IV)

Win-stay, lose-shift (WSLS)

Chapter 1: Introduction

1.1 General Introduction

Autism spectrum disorder (hereafter referred to as autism[1]) is a diverse and complex neurodevelopmental condition, diagnosed primarily based on the presence of socio-communicative impairments and restricted or repetitive behaviours [1]. However, despite not being prerequisite for diagnosis, there exists a growing awareness that other features, such as motor difficulties, are important characteristics of the autistic phenotype [2]. Thus, autism is characterised by many heterogeneous traits that affect different domains, with cognitive, motor, social and perceptual characteristics [3]–[5], suggesting widespread alterations in neural processing and computations. However, the underpinning computational mechanisms are yet to be understood. Indeed, most prominent theories of autism only seek to explain a specific subset of behaviours, for example, difficulties in social interaction [6],[7] or atypical sensory processing [8] [9]. Hence it is of crucial importance to investigate computations underlying atypical behaviours across all domains, i.e., link observable behaviours to underlying computational and neural processes and provide a broader explanation for phenotypic traits. This could lead to progress in improving therapeutic recommendations and appropriate strategies for autistic individuals, potentially via improvement of diagnostic criteria [10].

Bayesian and predictive coding accounts of autism propose a unifying explanation for the many heterogeneous symptoms observed in autism, under the framework of atypical prediction [11]–[13]. These accounts propose atypical predictive processing as a common domain-general impairment, raising the question of whether they can be extended to predictive processing in the social and motor domains. This chapter provides a background to these questions, first providing an overview of reward learning atypicalities in autism, before outlining Bayesian and predictive processing accounts of autism. Next, evidence for predictive atypicalities in motor and social learning is considered, and gaps in the current

[1] I will use identity-first terminology, such as 'autistic person', following the preferred language of many people on the autistic spectrum (e.g., Kenny *et al.*, 2016).

literature identified. In considering whether predictive atypicalities are a common computational atypicality in autism, the final section focuses specifically on social learning, considering evidence for and against the presence of domain-specific neural and neurochemical mechanisms for social learning.

1.2 Learning

Atypicalities in learning have been widely reported in autism and may comprise a common process in contributing to impaired development of motor and social skills [14]–[16]. While some types of learning are intact, and some indeed superior, in autistic individuals [17],[18], different learning styles have been observed, with autistic learning biased towards local, specific, information and away from context-dependent, global learning [19]. Atypical learning has been reported across different domains, such as perceptual learning [20] and sensorimotor learning [2],[21], amongst others. However, while clear differences exist, the exact nature of learning atypicalities in autism remains unclear [22]. In recent years, cognitive studies have investigated learning atypicalities in autism under different frameworks, including reward/reinforcement learning, reversal learning and implicit learning.

1.2.1 Reward learning

Learning relies on the computing and updating of internal representations of the environment's structure, facilitated by learning the signals of relevant events in order to anticipate their occurrence, and using past experiences to predict future outcomes [23]. Models of how the brain computes these representations have been described in the field of reward learning which focuses upon learning to take actions that maximize reward or value. Here, learning relies on a prediction error (PE) signal, which represents the difference between the actual and the anticipated reward. This signal allows the brain to update its predictions about the environment based on the PE signal, using this feedback to guide future decisions [24],[25]. The PE signifies the difference between the actual event (R) and the predicted event (V) in a given trial *t*:

$$PE = R\,(t) - V\,(t)$$

The PE updates the predicted value for the next trial (V (t+1)) in proportion to the learning rate α, which determines the rate of learning, by altering the weight given to more recent as compared to less recent events.

$$V(t+1) = V(t) + \alpha PE(t)$$

Therefore, PEs drive learning and inference, allowing the brain to update predictions or values based on the error signal and to refine future predictions [24]–[26]. Learning and inference are proposed to be PE-dependent across all domains, including sensory perception, motor and social processing [27]–[29], as well as during classic reward learning, i.e., when learning from reward and punishment.

1.2.2 Atypical reward learning in autism

Atypical reward learning has been observed in autism in both children and adults in the context of decision-making tasks [30] and probabilistic associative learning tasks [31],[32], as well as during perceptual learning [33]. Indeed, general reward learning of cue-outcome associations has been found to be a predictor of socio-cognitive symptoms in autistic children [34] and autistic individuals have been shown to demonstrate atypical learning from reward feedback, particularly when rewards are social in nature [35]. A recent systematic meta-analysis of functional imaging studies found atypical reward processing (including reward learning) in autistic individuals, with atypicalities associated primarily with striatal regions, including the caudate nucleus, putamen and ventral striatum (mainly the nucleus accumbens (NAcc)) [36]. However, in a study which specifically examined the neural correlates of PEs in autism, although atypical neural processing of PEs was observed, no behavioural differences in learning were found in autistic individuals [37].

Focusing on implicit reward learning, referring to learning without conscious awareness, autistic children were reported to demonstrate impaired learning, as well as an atypical neural response when learning from social feedback [38]. Similarly, in an implicit categorisation learning task, autistic individuals took longer to learn and demonstrated altered neural

activation and connectivity in frontostriatal regions [39]. However, there is also evidence for intact implicit learning in ASD with large meta-analyses finding no differences between autistic and non-autistic individuals [40],[41]. Therefore, the precise nature of implicit learning deficits in autism is still unclear.

Atypical learning in autism has been found more consistently in probabilistic reversal learning tasks, which involve a switch in cue-outcome contingencies after learning [42], allowing assessment of both reward learning and behavioural flexibility, the ability to adapt response when the environment changes [43]. Evidence for reduced flexibility during learning in autistic populations has been observed across different types of learning paradigms [38],[44]–[46]. For example, D'Cruz and colleagues found that, although initial learning did not differ between autistic and non-autistic groups, autistic participants were slower to learn new contingencies after a reversal, and made more perseverative errors [45]. Similarly, during a perceptual discrimination learning task, while autistic participants showed intact initial learning, impaired performance was observed when the target (consisting of three diagonal bars surrounded by horizontal bars) was moved to a new location, suggesting over-specificity of learning [33]. In addition, atypical flexible behaviour is consistently observed in ASD, with a large-scale study finding differences in flexible behaviour across all age groups, although the exact nature of learning atypicalities varied across the lifespan [46]. However, no differences, or differences that are unrelated to reversal phases, have also been reported [47],[48].

Overall, evidence suggests reward learning is atypical under some circumstances in autism, particularly in tasks where behavioural flexibility is required. For example, impairments are often observed in learning post-reversal, in line with accounts of inflexible and perseverative behaviour in autism [49]–[52]. However, atypicalities are not consistently observed across different studies and paradigms.

1.3 Bayesian and predictive coding accounts of Autism

Bayesian and predictive coding accounts of autism have attempted to resolve conflicting findings in the literature and provide a unifying explanation for the diverse symptoms observed in autism, under the framework of aberrant precision, or autism as a 'disorder of prediction' [11]–[13]. Here, atypical autistic perception and learning are described as stemming

from a reduction in the ability to accurately predict upcoming events and to utilise these predictions, building on accounts which explain autistic perception through a framework of atypical Bayesian inference [53],[54].

1.3.1 Bayesian inference

The Bayesian inference model describes perception and learning as relying on a combination of incoming sensory information and prior beliefs, resulting in the construction of a hierarchical probabilistic model of the environment [27],[55],[56]. The brain constructs prior predictions and updates these by combining prior knowledge, or beliefs, with incoming, 'bottom-up' sensory information. The weight given to the prior and sensory information is dependent on the reliability, or precision, of their estimation, with precision in this context referring to the inverse of variance. New 'posterior' beliefs are thus a weighted combination of priors and new (incoming sensory) evidence. Learning the associations between cues and outcomes can also be understood in line with the Bayesian inference framework, whereby the brain predicts future outcomes via computation of probabilities from past experiences [57].

Bayesian accounts of autism describe autism perception and learning as stemming from an imbalance in how prior beliefs and incoming sensory information are combined, resulting in an increased reliance on incoming evidence, although different models differ in their exact mechanistic explanations. For example, this imbalance has been described as emerging from a reduced influence of priors, or 'hypo-priors' [53] or, alternatively, from enhanced reliance on sensory, bottom-up processes, stemming from increased precision of incoming information [54]. Indeed, the majority of empirical evidence in support of Bayesian accounts of autism thus far stems from studies investigating the use of priors in autistic individuals, during sensory and perceptual processing. Many studies provide evidence in support of reduced reliance on priors in autistic individuals, across a variety of different paradigms [58]–[61]. However, and in direct contrast to the above, studies have also reported intact use of priors in autism [62]–[64], suggesting typical use of specific types of priors in autism.

1.3.2 Predictive coding

The hierarchical predictive coding framework builds on the above, providing a mechanistic explanation for how Bayesian inference could be implemented in the brain [27],[56],[65]. Here, prior beliefs are recast as higher-level predictions and integrated with incoming sensory information, with the discrepancy between both signals, or prediction error (PE), used to refine and update the high-level predictions and minimise further PEs. Learning and inference occur through minimisation of PEs, with the relative influence of PEs on predictions reliant on the relative precision, or reliability of incoming and higher-level information [66]. For example, higher precision of incoming information results in a greater weight on PEs during perception and learning. In contrast, higher precision of top-down predictions relative to PEs will result in a greater weight on top-down predictions, or priors. The relative precision of PEs and higher level predictions should flexibly vary according to the context: precision of PEs should be high when PEs signal useful changes in cue-outcome probabilities (i.e., promoting ongoing learning), and low when signalling uninformative changes, or noise [67]. Modulation of precision is hierarchical; the relative weighting of PEs at lower levels relies on the precision of higher-level predictions about the probabilistic structure of the environment [65],[68], with higher level predictions representing increasingly higher-level or abstract concepts. Different neuromodulators, such as dopamine, norepinephrine and acetylcholine, have been proposed to modulate weighting or precision estimation of PEs at different hierarchical levels, through modulation of postsynaptic gain [69]–[71].

Prediction in the context of learning

Autism, along with other conditions [72]–[74], has been described through the framework of atypical predictive coding [75],[76]. Under predictive accounts of autism, it is the relative precision of PEs and higher-level predictions that is atypical, rather than PEs and priors themselves [12],[13],[77]. For example, Lawson and colleagues explain features of autistic cognition as stemming from reduced precision of priors relative to precision of sensory information [13]. Van de Cruys and colleagues propose that flexible adjusting of precision of PEs is impaired in autism, with PEs weighted uniformly high and inflexibly, resulting in a reduced distinction between PEs that are informative and relevant (signals) and those that are irrelevant (noise). Here, constantly high precision of PEs results in overly high rates of learning from both signal and noise, leading to overfitting of predictions and reduced

generalisation [12]. Other accounts propose atypical adjustment of weighting of PEs in response to context (atypical precision-weighting) [78], and learning-based accounts propose atypicalities in predicting conditional relationships between events, leading to increased baseline uncertainty [11]. While specific details differ, these theories all converge on the main theme of atypical predictive processes as a common endophenotype in autism [79],[80], a measurable cognitive mechanism bridging genotype and observable phenotypic symptoms. In addition, atypical predictive coding could explain contrasting findings with regards to reliance on priors in autism; autistic individuals could learn priors but not apply them as broadly as non-autistic individuals.

Overall, these theoretical accounts align with other theories of autistic perception and learning, such as a bias towards local or low-level processing, over global processing [9],[81],[82], and different learning styles in autism [19]. They extend previous accounts of autism, with the aim of describing the underlying computational mechanisms, and propose explanations for common behavioural atypicalities, such as intolerance of change and a desire for predictability [83],[84], as well as sensory atypicalities. It has been suggested that atypical predictive processes, such as inflexible and high weighting of PEs, lead to increased feelings of surprise, and that repetitive behaviours and desire for sameness are a compensatory response to an overall unpredictable environment [12]. This would result in autistic individuals perceiving events as being overall more unpredictable [50] and having a reduced ability to separate signal from noise during learning [12], particularly in complex and highly uncertain social environments [12],[13]. Predictive accounts of autism, while differing in their exact hypotheses, propose similar patterns of behaviour, such as decreased weight on prior predictions and an overreliance on new, incoming information, reduced flexible adjustment of learning and an increased baseline level of surprise. Importantly, as these accounts do not propose a general impairment in learning, but rather difficulty in flexibly adjusting learning rate or atypical learning in complex environments, this framework could shed light on the mixed findings from previous studies investigating learning in autism. For example, differences might be observable when learning in a more unpredictable environment or from more complex information.

Learning and uncertainty

As described in the previous section, learning is driven by prediction errors (PE), which have been described as a surprise signal, indicating the difference between the expected and actual

outcome. Learning depends on minimising or explaining away PEs, by updating expectations or beliefs. The degree to which PEs update beliefs is reflected by the speed of learning, or learning rate, determining to what extent more recent and more distant cues are taken into account during learning [85]. In the context of predictive coding accounts, the learning rate depends on a combination of the precision of incoming, bottom-up information and top-down predictions [65],[68]. Optimal learning relies on tracking the probability of cue-outcome relationships, which are inherently noisy or variable. The uncertainty (or precision) around these relationships is commonly referred to as 'expected uncertainty' [86]. However, in dynamic environments, cue-outcome joint probabilities or associations can change over time. Uncertainty about these changes is known as "unexpected uncertainty", reflecting the (subjective) uncertainty about changes in the environment over time; i.e., about the volatility of the environment [70],[87],[88]. Therefore, the weight assigned to PEs during learning should vary depending on the background environment or volatility. For example, in a rapidly changing, or volatile environment, with high unexpected uncertainty, more weight should be given to PEs relative to prior predictions, as PEs are signalling useful information about changes in cue-outcome contingencies. Thus, according to theoretical accounts [67],[88] learning rates should be higher in volatile phases, with more recent outcomes used to update decisions. The opposite should be true, however, in stable environments, where unexpected uncertainty is low. Here, the weighting of PEs should be low, as PEs are likely to reflect noise, rather than informative changes in cue-outcome probabilities. Learning rates should therefore also be low, and more distant outcomes taken into account [89]. The capacity to adjust parameters of learning, or adapt learning rates in response to the current environment, is proposed to be mediated by neuromodulators, such as acetylcholine (ACh), dopamine (DA), norepinephrine (NE) and serotonin (5-HT) [90],[91], encoding uncertainty signals or precision around PEs [67],[92]. In line with theoretical accounts, a key role has been demonstrated for catecholamines (dopamine and norepinephrine) in the context-dependent adjustment of learning, with pharmacological manipulation of catecholamine signalling affecting adaptation to environmental volatility [93],[94].

Empirical studies support the theoretical proposal that healthy individuals adjust their learning rates to volatility [70],[88],[95], although contrasting evidence has also been found [96]. Atypical adjustment of learning rates has been observed in psychiatric conditions, such as schizophrenia and borderline personality disorder [72],[97], and correlates with trait anxiety in healthy individuals [98]. Adjustment of the relative weighting of PEs, and therefore adjustment

of learning rates or response to volatility, is proposed to be atypical in autistic individuals [12],[13],[77],[78]. For example, if precision estimation of PEs is constantly high and/or not flexibly modulated by volatility, learning rates will be inflexibly high [77]. Other accounts hypothesise high precision of incoming, relative to prior predictions, which would be more noticeable in uncertain situations where prior predictions should be more heavily relied on [13]. In addition, atypical precision estimation could result in a decrease in learning the regularities of the environment, resulting in atypical prediction and anticipation and an atypically high baseline level of surprise in autistic individuals [11]. Overall, all accounts predict atypical responses to uncertainty (both expected and unexpected uncertainty) in autism, with empirical evidence in support of this: atypical predictive processes have been observed in volatile environments with changing cue-outcome probabilities, and atypical responses to unexpected events have been observed [99]–[101].

1.3.3 Empirical evidence for atypical prediction in autism

The focus of this thesis will be on prediction-based learning processes, specifically the ability to learn cue-outcome contingencies. However, it is also important to examine behavioural and neural responses to predictable events. For example, evidence for atypical predictive processes in autism comes from studies on perceptual adaptation, whereby repeated exposure to a stimulus subsequently biases perception away from that particular stimulus [102]. Here, reduced adaptation is proposed to reflect a reduced impact of prior predictions on perception. Several studies have demonstrated reduced adaptation in autism in the visual domain, including face discrimination [103], gaze direction [61] and biological motion perception [104], although contrasting results have been found in studies of biological motion judgements [105] and face identity and expression in autistic adults [106]. Within the field of auditory perception, autistic adults show reduced adaptation to loudness [107] and audio-visual integration [108], with neural signals reflecting reduced auditory adaptation observed in autistic adults [109]. Contrasting evidence comes from a task investigating motion prediction in children and adolescents with autism, with no differences in predictive abilities observed between autistic and non-autistic participants [110].

Further evidence for atypical predictive processes in autism comes from research investigating behavioural habituation, or attenuation of a response subsequent to repeated exposure to a stimulus. Habituation of response is a form of single-stimulus learning and

reflects a prediction or expectation that a stimulus is constant; reduced habituation can therefore be interpreted as a reduced ability to predict an upcoming stimulus or downregulate PEs. The majority of studies providing evidence of reduced habituation in autism are in the auditory domain [109],[111]–[113]. However, there is also evidence for reduced habituation with regard to visual stimuli, with reduced behavioural and neural measures of habituation observed in response to social (facial images) [59],[114] and non-social images in both autistic adults [115] and children [116]. Finally, although no evidence of reduced habituation was observed in autistic adolescents, atypical patterns of neural activation were found in response to unexpected visual stimuli [117].

Empirical testing of atypical predictive learning in autism has mainly been in the sensory/perceptual domain, where atypicalities have been well-documented in autism [118]. Implicit learning tasks have shown reduced use of previous sensory information, or environmental statistics in autism. For example, in a statistical learning paradigm, autistic individuals were reported to show an atypical neural response to unexpected auditory events which violated previous predictions [119]. A further study found reduced reliance of previous trials on perceptual decision-making, despite no evidence of atypical weighting of PEs or overfitting of predictions [58]. In an action prediction task, while non-autistic children relied on previously learned associations between context and kinematic patterns (priors) to predict action outcomes under conditions of perceptual uncertainty, autistic children did not [120]. Further research, using eye-tracking to examine how autistic adolescents responded to unexpected events that violated learned visual associations between colour and location, reported atypical predictive behaviour in the autistic group, indexed by reduced gaze to predicted locations [121]. Similarly, Lawson and colleagues [100] used a perceptual learning task with audio-visual stimuli to investigate probabilistic associative learning in autism. Results showed that, while autistic adults could successfully learn associations between a high/low pitched tone and an image (face/house), response to unpredictable events was atypical. Autistic adults showed reduced surprise in response to unpredicted events, indexed by reduced slowing of response, and reduced pupil dilation. Computational models of learning fitted to behavioural data suggested atypical (over) estimation of volatility and subsequent atypical adaptation of learning rates in the autistic group [100]. Finally, an associative learning task comparing autistic and non-autistic adults was conducted, requiring learning of auditory-visual stimuli associations. Although both autistic and non-autistic groups demonstrated use of prior predictions in learning associations, and learning did not differ between groups, the

autistic group were slower to update priors after a reversal in associations between stimuli, suggesting less flexible updating of priors [122]. Taken together, evidence across perceptual learning studies supports atypical predictive learning in autism.

Learning has been investigated in autism in reward-learning paradigms, where correct choices are linked to receipt of reward. However, evidence for atypical predictive processes within the domain of reward-learning is mixed. First, a reward-learning paradigm comparing neural responses to social and non-social reward in adolescents found no differences in behavioural indices of learning. However, differing neural activation patterns were observed in frontostriatal regions for PEs linked to social, but not non-social, rewards, suggesting atypical predictive mechanisms when learning from social reward stimuli [123]. Using a reward learning task, Goris and colleagues [99] demonstrated a correlation between less optimal decision-making in volatile contexts and higher levels of autistic traits (indexed by autism spectrum quotient (AQ) scores [124]), in a non-autistic sample. Importantly, they used a paradigm that could distinguish between volatile and noisy contexts, via inclusion of a condition where the probabilities of cue-outcome associations were low (i.e., noisy) but stable. By comparing adjustment of learning in volatile versus noisy contexts, they showed that deficits in decision-making in volatile contexts were not driven by impaired context-dependent adjustments of learning rates, suggesting typical precision estimation during learning, despite impaired performance in volatile environments. However, although a large sample was used and therefore variation in autistic traits was captured, this study investigated learning in a neurotypical population, thus the extent to which the conclusions apply to individuals with a clinical diagnosis of autism is unclear [99]. Using a similar probabilistic reward learning task, Robic and colleagues [101] investigated the influence of social and non-social information on choice, under stable and volatile conditions. On each trial, participants were required to choose one of two boxes, with hidden reward probabilities, and received advice in the form of social or non-social advice cues, which varied in their utility. Reduced performance in the autistic group (as measured by the proportion of participants who reached a pre-defined success criterion) was found on trials with the social cue and during volatile phases, although volatility had a greater impact on task performance compared with the social cue [101]. Finally, contrasting results have been found. Manning and colleagues [125] compared the performance of autistic and non-autistic children on a probabilistic reward learning task, under both stable and volatile conditions. In stable conditions, reward probabilities were fixed, while, in volatile phases, they fluctuated, with cue-outcomes probabilities reversing

every 10-20 trials. In contrast to the main hypotheses of the study, which predicted uniformly higher, and less flexible learning rates in the autistic group, no significant group differences were observed. Autistic children showed typical learning rates and increased learning rates in volatile, relative to stable phases, indicating typical precision estimation of PEs [125].

Overall, while atypicalities in learning in autism are often observed in the perceptual domain, evidence is more mixed with regard to reward learning. Atypical performance is frequently observed in volatile or uncertain conditions in autistic individuals, but this cannot be fully ascribed to impaired adjustment of learning rates. In addition to explaining atypical sensory perception and learning, however, Bayesian and predictive processing accounts propose a general mechanism underpinning autistic processing which should apply to learning across different domains [11]. While empirical evidence in recent years has grown rapidly, it is crucial to test these theories across all domains and investigate whether predictive impairments represent a core underlying impairment in autistic learning. Extending existing theories in this way may help us to understand whether common, domain-general predictive impairments can explain frequently observed features of autism, such as social and motor atypicalities.

1.4 Motor atypicalities in autism

There is mounting evidence to suggest that motor differences, although not included in formal diagnostic criteria, are a core feature of autism [126],[127], with estimates of the prevalence of motor deficits ranging between 20 – 90 % of the autistic population [128]–[131]. Motor atypicalities can manifest as deficits in both fine and gross motor skills[132], including abnormalities in motor coordination and posture [126],[133], clumsiness of gait [134], impaired skilled motor gestures [135]–[137] and atypical motor planning [138],[139]. Motor deficits lead to difficulties in normal day-to-day functioning [140], but also affect communication and social interaction [141]. For example, delays in obtaining fundamental motor skills in autism [142] can lead to delay in obtaining skills such as writing, speaking, and playing [135],[143], and deficits in motor planning and sequencing could affect the ability to regulate movement while communicating with or imitating others [144]. Movement differences also contribute to difficulties for autistic individuals in interpreting the movements of others, leading to difficulties in social cognition [145]. Movement atypicalities may stem from motor learning

difficulties [2],[141],[146], with evidence that autistic individuals show delays in learning new skills [133] and in learning complex sequences of movements [147],[148]. In the next two sections, I review evidence for the presence of atypical predictive coding as a core mechanism underpinning both general motor atypicalities, and for motor learning in particular.

1.4.1 Atypicalities in the general motor domain

The majority of studies examining predictive processes in the motor domain have examined autistic individuals' action perception and understanding of others, meaning that movement and action are investigated within a social context, rather than with a focus on low-level motor processes. For example, research has shown that autistic adults demonstrate a reduction in anticipatory neural activity during processing of action-related sounds and words [149], reduced ability to predict the goal-directed actions underlying reaching movements of other individuals [150] and attenuated use of priors when predicting goal-directed actions of others [151]. Amoruso et al. [152] examined modulation of excitability in the primary motor cortex (M1) during action prediction in a non-autistic population. Individuals observed videos of actions in congruent and incongruent contexts. Autistic traits correlated with a reduction in the ability to downregulate M1 when observing actions in an incongruent context (i.e., when a mismatch occurred), suggesting impaired integration of predictions (goal representation) and incoming evidence (kinematics) [152]. A further study found that, when inferring intentions of others through action observation under conditions of perceptual uncertainty, autistic children showed reduced use of priors in comparison with non-autistic children [120]. Taken together, evidence suggests that predictive processes are atypical in both autistic children and adults, during action perception and understanding.

1.4.2 Motor learning

Active inference extends the predictive processing framework to include action and movement [153],[154], with actions described as a method to minimise prediction errors or surprise [155]. Similarly, motor learning is described as relying on the creation and utilisation of internal sensory models of actions to predict movement, with these predictions compared to incoming sensory (visual and proprioceptive) feedback. Here, PEs can be reduced by either updating predictions or through action [153],[156],[157]. Although empirical evidence is still

limited, studies have begun to examine prediction in motor learning in autistic individuals, with mixed results.

Empirical evidence for atypical motor learning in autism

A recent study investigating sensorimotor prediction in adults with autism used two tasks to investigate predictive mechanisms in the sensorimotor domain: a force matching task, measuring attenuation in the sensory perception of self-generated movements, and an intentional binding task; an effect whereby sensory outcomes and motor actions are perceived to be closer together temporally when actions are voluntary and self-generated, that is thought to index predictive motor processes. Thus, allowing investigation into both low-level sensory prediction and prediction of action-outcome contingencies in the same group of individuals. However, the authors found no evidence to support the presence of atypical predictive processing in either task [158]. Similarly, Arthur and colleagues [159] reported that autistic adults showed typical predictive processes for sensory perception and action during an object lifting task. Specifically, both initial force profiles and subsequent action kinematics were driven by prior perceptions of object weight in autistic individuals [159]. However, in a subsequent study, which utilised a virtual reality paradigm to measure sensorimotor prediction in a volatile environment, autistic adults showed atypical predictive behaviour, indexed by atypical movement kinematics and predictive gaze in response to surprising or unexpected events. These differences were amplified under more uncertain, or volatile conditions [160]. Taken together, evidence from sensorimotor paradigms suggests that autistic adults show typical use of predictive action models, but atypical adjustment of behaviour in volatile conditions, in line with accounts proposing, rather than higher weighting of PEs overall, atypical *modulation* of precision in autism [78].

Aside from sensorimotor learning, learning has been investigated in the motor domain in the context of implicit learning of action sequences, mainly through employment of sequence learning paradigms, such as the serial reaction time task (SRT) [161]. Here, participants are required to learn cue-action contingencies in a training phase, and then to respond to these cues as rapidly as possible by performing the associated action in the test phase. During the test phases, and unknown to the participant, cues follow a repeating pattern or sequence, providing a measure of implicit or statistical learning; participants are unaware of the underlying pattern but demonstrate increased speed for predictable events that follow the sequence. Mixed evidence has been found when investigating motor sequence learning in

autistic individuals, with some studies finding intact sequence learning [162]–[164] and some, in direct contrast, finding atypical sequence learning [148],[165],[166].

Predictive processing accounts of learning have the potential to shed light on the mechanisms underpinning motor sequence learning in autism. For example, the motor system relies on top-down predictions (or priors) to prepare for actions, in a manner that is proportional to the likelihood of the event occurring [167]. Therefore, as the precision of predictions about an upcoming action increases, reaction time (RT) decreases. This implies that, if a violation of predictions (a PE) occurs, RT will increase, due to the requirement to inhibit the prepared action and prepare and execute the unpredicted action. Predictive theories propose an atypically high baseline level of surprise in autistic individuals [78], suggesting that unpredictable or surprising events are not as surprising for autistic relative to non-autistic individuals. In addition, if prior predictions are attenuated, learning and motor planning would be impaired. Thus, in the context of sequence learning, an atypical response to surprising or unpredictable events should be observed in autistic individuals, indexed by a lack of slowing of response where slowing would be expected (i.e., reduced surprise-related slowing), and/or impaired motor preparation during sequence learning.

Evidence for reduced surprise-related slowing has been demonstrated in the sensory domain [100], with autistic adults demonstrating a reduced distinction between expected and unexpected outcomes, indexed by reduced slowing of response, relative to non-autistic adults. In the motor domain, Rinehart and colleagues showed that, when executing actions in response to changing visual patterns, autistic, compared to non-autistic children showed a significant reduction in surprise-related slowing in response to unexpected patterns [168]. Similarly, Gidley Larson and colleagues used a rotary-pursuit paradigm to examine visuomotor sequence learning in autistic children. Rotary-pursuit tasks are a measure of motor skill performance, involving learning a sequence of movements to accurately predict and track the motion of a moving target. Results showed that violations to the pattern resulted in surprise-related slowing for non-autistic, but not for autistic, children [147]. However, both studies were carried out with children, highlighting the importance of determining whether these effects persist into adulthood. To date, only one study has been carried out with adults, with no differences between groups observed with regard to surprise-related slowing [139].

In sum, a limited number of studies have investigated motor processes from a predictive coding perspective in autism, with the majority showing a focus on action perception. The

studies that do, including low-level sensorimotor prediction, provide mixed evidence. Previous work suggests atypical prediction in autistic individuals, but these studies were not developed specifically to test hypotheses from predictive coding accounts of autism, and were conducted in children and adolescents, but not in adults. It is therefore important to investigate predictive motor sequence learning in adults and look at response to surprise on a trial-by-trial basis (Chapter 2).

1.5 Social atypicalities in autism

Impairments in social communication and interaction are classified as cardinal symptoms in autism [1], with a negative impact on quality of life and functioning [169]. Bayesian and predictive processing accounts suggest common computational underpinnings for social and non-social symptoms (repetitive and restricted behaviour, interests and activities), in contrast with domain-specific theories, which propose a primary 'social' impairment, such as impaired theory of mind in autism [170], or reduced social motivation. [7]. Indeed, these accounts of autism propose that predictive deficits are enhanced in the social domain [13],[78]. For example, understanding the actions or intentions of other individuals is proposed to be more difficult, as social information is intrinsically unpredictable and ambiguous, with less defined contingencies between events. Predictive accounts thus propose difficulties for autistic individuals in inferring the causes of (often unpredictable) social stimuli, through a lack of generalisation and overfitting of incoming data [10]. In the next two sections, I review evidence for the presence of atypical predictive processing in autism, first within the general social domain, and then for social learning specifically, with a view to assessing what is and is not known regarding predictive atypicalities in the social domain.

1.5.1 Social cognitive atypicalities

Within the social domain, individuals must predict the behaviour of others, a process that has been described through the framework of predictive coding [29],[171],[172]. Predictive accounts of autism propose that deficits in predicting the actions/and or intentions of other individuals could be underpinned by atypical prediction. Indeed, there is evidence that autistic individuals show impaired use of social information when predicting the actions and/or

intentions of others, through attenuated use of prior predictions. For example, using a visual motion dot paradigm, Von der Lühe and colleagues reported that, while non-autistic adults used the communicative actions of one agent to predict the actions of an interacting agent, autistic adults did not [173]. Similarly, Chambon et al. [174] found that attenuated prior use in autism was specific to social interactions. Autistic and non-autistic matched groups were required to observe two actors manipulating objects and infer their intentions, in both a social, and non-social context. The autistic group displayed reduced reliance on prior predictions within the social context, compared to non-autistic controls [174]. Furthermore, atypical use of social priors has been observed outside of paradigms investigating action understanding. For example, in comparison with non-autistic individuals, autistic adults showed reduced use of social priors in a paradigm examining the effect of preceding dynamic facial expressions on the evaluation of a subsequent neutral expression [175]. In contrast, in a study examining response to violations of learned visual cue-outcome events, while autistic individuals showed an atypical response to surprising events, this did not vary with regard to the social/non-social nature of the stimulus [121]. Finally, a study comparing adaptation to social and non-social visual stimuli in autistic children, found reduced adaptation to social stimuli only, indexed by reduced perceptual aftereffects [60], while a similar study found diminished adaption to eye-gaze stimuli in autistic adults [176]. Taken together, these studies suggest atypical predictive processing in the general social domain in autistic individuals. In the next section, I focus specifically on social learning, summarising the evidence for atypical PE-based learning in autism, and identifying areas that require further examination.

1.5.2 Social learning

Highlighting the importance of social learning, sociocognitive ability is thought to develop to a large extent through implicit learning from observation of other individuals [16]. Furthermore, some predominantly social theories of autism, such as reduced attention to social stimuli [177], decreased motivation to attend to social stimuli [7] or difficulty in processing biological stimuli [178], propose a reduction in opportunities for, and/or abilities in social learning, which could lead to impairment in social cognitive function. Thus, sociocognitive difficulties could have atypical social learning at their root. Social learning, or learning from others, refers to the ability for an individual to obtain information or adapt their behaviour as a result of observation of another individual's actions or choices, rather than

through their own direct experience [179]. In contrast, asocial, or "individual learning" refers to learning directly through one's own experience, through trial and error, independently of advice or observation of others. Social learning has been documented across many species [180],[181], and is fundamental for the rapid acquisition of valuable information, such as learning to acquire rewards and avoid harm, without personal cost or risk [182],[183].

Before considering the evidence for atypical social learning in autistic individuals, it is important to define social learning precisely, as the term is used interchangeably for different types of prediction-based learning. For example, social learning has been used to describe individual learning *about* other individuals, prosocial learning, namely learning the impact of one's own actions on another individual [184], and for individual learning from social reward [123]. Additionally, social learning can include learning via verbal instruction or teaching [185], as well as imitation. In this thesis, I use social learning to refer solely to learning from other individuals by observing their choices or actions and subsequent outcomes [179],[186].

Empirical evidence for atypical social learning in autism

Empirical evidence for atypical social learning is limited in autistic individuals. For example, as highlighted above, 'social' prediction error-based learning is used to describe PEs both when learning *from* other individuals, but also when learning is linked to a 'social' reward, i.e., an image of a happy face. For example, altered neural activation was reported in autistic individuals during a learning task, for 'social' PEs only. However, the PEs in this study reflected the type of reward (i.e., an image of a face or an image of an object), rather than learning through observation of another individual [123]. A further study reported differences in prediction-error associated neural signals in autistic individuals when coding PEs during observation of another person's decisions. Balsters and colleagues [187] compared autistic and non-autistic individuals on a probabilistic decision making task in combination with neuroimaging. Participants were required to play on behalf of themselves (i.e., make decisions for themselves), or to observe another individual playing, and predict, from the other person's perspective, whether the observed outcome was expected or unexpected. Both reduced accuracy and reduced neural activation in the gyrus of the anterior cingulate cortex (ACCg) were observed in autistic individuals for 'social PEs' based on predictions made from another's perspective. Results showed that autistic individuals were less accurate at monitoring the predictions of others, potentially resulting in a deficit in understanding the perspectives of others. [187]. However, this 'social' prediction error requires taking the

perspective of another person, comparing actual outcomes with expected outcomes as perceived by another individual, rather than learning for one's own benefit through observation of another individual's actions or choices. In sum, while the studies described in this section suggest atypicalities in predictive processing in the social domain in autism, there is limited research directly comparing social and individual PE-based learning in autistic individuals.

Preliminary evidence for predictive differences in social learning comes from research in a non-autistic population. Sevgi et al. showed that autistic traits were correlated with difficulty in integrating social information during learning, in a task that required integration of social and non-social cues. Individuals with high autistic traits showed a reduction in the ability to accurately utilise social information to modulate the precision of individual information [188]. Using a similar learning task, Robic and colleagues [101] compared the influence of social and non-social information on a probabilistic learning task. On each trial, participants were required to choose one of two boxes, and received advice in the form of a social or non-social advice cue. Reduced performance was observed in a volatile context in autistic adults, with these deficits more pronounced when social cues were provided. However, this study did not estimate learning rates, but rather examined group differences by comparing the number of participants in each group who achieved a certain performance criterion (60% accuracy). Thus, inter-individual differences in learning rates were not taken into account and the underlying mechanisms underpinning impaired performance were not examined [101].

Taken together, there is limited evidence in the social domain for atypical predictive processes, and no studies directly comparing social and individual learning rates in an autistic population; somewhat surprising given that Bayesian/predictive theories of autism propose that atypicalities will be amplified in a social environment. Moreover, determining whether or not domain-general theories of autistic learning and cognition can provide explanations for diverse atypicalities feeds into a wider debate, concerning the domain-specificity of social processes, in both autistic [169] and neurotypical individuals [186],[189],[190]. For example, an unresolved question in the literature is the extent to which neurochemical mechanisms and neural pathways and connections are specialised for social learning. In the next section, I first describe opposing views of the mechanisms underpinning social learning, before outlining how an investigation at the level of neurochemical signalling could help to resolve outstanding questions in this field.

1.6 Domain-specificity of social learning

Social learning, learning from observing the actions and outcomes of another individual, has been widely studied, with its importance emphasised by its occurrence across many species including humans [180],[181],[191]–[193]. In addition, social learning is widely described as the basis for human culture, by enabling behaviours to spread between individuals in groups and through subsequent generations [191],[192],[194]. There remains, however, debate as to the fundamental mechanisms underlying social learning.

Behavioural ecology and evolutionary viewpoints argue that humans, and other social animals, possess specific learning mechanisms, specialised for social learning [180],[192],[195],[196]. Here, social learning is thought to rely on distinct "social" neural mechanisms, or processes, which have genetically evolved independently of those underpinning individual (asocial) learning. Thus suggesting that social learning is an 'adaptive specialisation' for living in social groups [190],[197],[198]. Evidence in support of this stems from comparative research, with social and individual learning shown to be dissociable, both within [199] and between different species [200]. For example, variability in social, but not non-social, learning between humans and non-human primates at a young age has been used as evidence for the presence of genetically inherited social learning capabilities in humans [190],[200].

Alternatively, social learning may be underpinned by the same mechanisms as individual learning, i.e., learning from one's own reward outcomes [186],[201]. This domain-general approach argues that social learning mechanisms are underpinned by the same core associative mechanisms as general reinforcement learning (RL). Evidence for this theory originally stemmed from comparative research across different species [179]. For example, social and non-social learning co-vary both within [202],[203] and across species [204] and asocial animals can also demonstrate social learning if required to do so [205],[206]. Under this framework, differences in social learning abilities across species depend on, among others, variability in the attention paid to, and motivation to attend to, social stimuli [186],[207]. Crucially, however, the underlying mechanisms are proposed to be the same as those underpinning individual learning, namely associative learning processes [179],[186],[201],[208]. Indeed, key associative learning phenomena, including blocking [209] and overshadowing have been observed during social learning [186], and a recent computational modelling study

demonstrated that many types of social learning can be modelled by domain-general associative RL learning mechanisms [210].

If both social and individual learning rely on the same underlying cognitive mechanisms, this strongly suggests that they would be underpinned by the same neurochemical signalling processes. In contrast, if social learning relies on separate, social-specific cognitive mechanisms, that have evolved specifically, social and individual learning should be dissociable at a cognitive, neural and/or neurochemical level. Thus, to determine if specific mechanisms underpin social learning, it is important to look for dissociations between learning types. Alternatively, a lack of dissociation would fail to support the adaptative specialisation view but is consistent with the presence of shared mechanisms. I here focus on the neurochemical level. The monoamine neurotransmitters dopamine and serotonin have both been implicated in learning from one's own experience (individual learning) and in social learning (learning from others). In the next two sections, I outline evidence for the role of monoamine signalling in both types of learning.

1.6.1 Neurochemical mechanisms of learning

Dopaminergic mechanisms

Individual or direct reward learning relies on phasic dopaminergic signalling, with midbrain dopaminergic neurons in the ventral tegmental area (VTA) and substantia nigra (SN) encoding a PE signal, reflecting the difference between the actual and expected reward [71],[211]–[213]. Thus, PEs drive learning, allowing the brain to update its beliefs or values based on the error signal and to refine future predictions [24]–[26]. In humans, imaging, pharmacology and genetics studies corroborate this role for dopamine in updating predictions [70],[93],[214]–[216]. For example, functional magnetic resonance imaging (fMRI) studies have shown PE related blood oxygenation level-dependent (BOLD) signals in reward-related neural regions such as the striatum [217],[218] and the VTA/SN [219]. In addition, pharmacological manipulations of dopamine signalling modulate neural correlates of PEs and behavioural choice in healthy individuals [214]. Midbrain dopaminergic signalling has also been proposed to encode the precision of PEs, rather than PEs themselves [154], with recent empirical evidence in support of this [70],[220]. Thus, dopaminergic signalling plays a crucial role in

individual learning. If social learning relies on the same mechanisms as individual learning, a role for dopaminergic signalling should be observable with regard to social learning.

Indeed, in line with accounts proposing domain-general neurochemical mechanisms for social learning, recent work has highlighted a role for dopaminergic signalling in social learning [221]–[225]. Evidence in humans (see Joiner et al. [226] for a review in the non-human animal literature) comes mostly from the neuroimaging literature, where a number of studies have demonstrated that social learning-related PEs correlate with BOLD response in dopamine-rich brain regions such as the ventral striatum [95],[223],[227]–[229]. In further support of shared dopamine-dependent mechanisms for both types of learning, there is evidence that the same prediction-based striatal computations underlie both social and individual learning [230]. Furthermore, social PEs have been shown to covary with genetic variation in genes that affect dopamine signalling, via modulation of dopamine reuptake [222]. Finally, social information is integrated based on its reliability or precision [231] in the same manner as individual information. Thus, evidence suggests that, in addition to its role in individual learning, dopamine may also be implicated in social learning, suggesting shared dopaminergic mechanisms underpinning both types of learning.

Preliminary evidence against a neurochemical dissociation between social and individual learning comes from a study by Cook and colleagues [93], wherein a pharmacological manipulation of dopamine (and norepinephrine) signalling affected learning from the primary information source only, regardless of the social nature of the information. This observation raises the intriguing possibility that dissociations observed during learning could reflect the status of the information source (i.e., whether or not it was the primary learning source), rather than the social nature of information [93]. Taken together, these studies provide support in favour of domain-general dopaminergic mechanisms for social and individual learning.

Serotoninergic mechanisms

Along with dopamine, preliminary work has implicated serotonin in individual learning [232]–[237], with a role for serotonergic signalling proposed in signalling unsigned or surprise-related PEs [238],[239]. Matias and colleagues [239], for instance, reported both positive and negative prediction error-like signals in 5-HT neurons in mice, using a pharmacogenetic approach. Whereas dopamine was initially associated with individual learning and has only more recently become linked to social learning, serotonin has long had an association with social behaviour, with a wide body of research highlighting the importance of the serotonin system

in social cognition [240],[241]. Early work with rodents and nonhuman primates identified a role for serotonin in social behaviours including social play [242]–[244], social perception [245] and the establishment of social hierarchies [246],[247]. Subsequent studies in humans have highlighted the importance of the serotonin system in many types of social cognition [240],[241], (though see Crockett and Cools [248] for discussion of potential non-social underlying mechanisms). For instance, the effects of serotoninergic enhancement or depletion have been demonstrated on social perception [249],[250], social cooperation [251], altruistic behaviour [252] and social trust [253], among others. Using a reward learning paradigm with social stimuli (emotional faces), Frey & McCabe showed that serotonin depletion resulted in alterations in both behavioural and neural markers of learning [254]. Focusing on social learning, Crişan al. found a significant effect of genetic variation in serotonergic signalling on learning from observing the fear responses of other people [255]. In sum, serotonin is implicated in both individual and social learning, suggesting shared neurochemical mechanisms. Nevertheless, this cannot be concluded with confidence from existing research, since extant studies have not directly compared the effects of variation in serotonin signalling on social and individual learning.

1.6.2 Neural mechanisms of learning

In contrast to studies suggesting domain-general learning neurochemical mechanisms for social learning, neuroimaging studies have provided evidence of dissociations between brain regions involved in individual and social learning [95],[256]–[258]. For example, Behrens and colleagues developed a decision-making paradigm where subjects were required to learn from social and individual information simultaneously, integrating both sources of information. While overlapping neural correlates were found for social and individual learning in the striatum, dissociable patterns of activation were also observed. Neural correlates of social learning were observed in the ACCg and temporoparietal junction (TPJ), while correlates for individual learning were observed in the sulcus of the ACC (ACCs) [95]. Therefore, while social learning processes could be underpinned by the same dopamine-mediated prediction-based mechanisms as individual learning [226] relying upon shared regions, such as the striatum and ventromedial prefrontal cortex (vmPFC) [95],[228],[230], social and individual learning could, at least to some extent, recruit separable brain networks/areas. For example, neural regions such as the dorsomedial prefrontal cortex (dmPFC), the ACCg [95],[228],[256]–[258] and the TPJ [95],[259], areas described as important for representing the

motivational states of others and simulation of others' actions and intentions [260],[261], arguably comprise 'social-specific' regions that have evolved for social living [189],[262],[263].

Alternative explanations for differing patterns of neural activation for social and individual learning must be considered before a brain region can be conclusively described as 'social-specific'. For example, in a task where participants had to make choices on behalf of themselves and at other times on behalf of others, neural activation patterns were observed in dmPFC that were related to preferences that were not guiding the current choice, e.g., the other individual in the 'self' condition and own preferences in the 'other' condition. In contrast, activation was observed in the vmPFC for preferences related to the current choice, suggesting a gradient of activation that, rather than representing self versus other [264], represents the current task relevance of information [265],[266]. Thus, it is possible that the observed dissociations between neural mechanisms for social and individual learning mentioned previously, could be better explained as dissociations between other factors, such as the relevance or the value of the information during learning. However, extant paradigms cannot determine which factor accounts for observed dissociations, as the social nature of information is confounded with other factors (such as whether information is the primary source of learning) in these paradigms.

In sum, research suggests that serotoninergic, as well as dopaminergic signalling, play a key role in both individual and social learning. Therefore, if social and individual learning rely on dissociable neurochemical mechanisms, it is important to look for dissociations as a function of dopamine and/or serotonin availability. To date, studies have not been able to test this hypothesis because they have not employed a design that enables mapping of variation in neurochemical signalling onto variation in both social and individual learning and, more importantly, have not accounted for confounding factors.

1.7 Aims of this thesis

The overall aims of this thesis are to shed light on the conditions wherein predictive processes/learning is, and is not, atypical in high functioning adults with autism, and, to investigate the neurochemical underpinnings of social learning in neurotypical individuals. The first two empirical chapters investigate social and motor learning in autism, by examining whether autistic adults exhibit differences, compared to non-autistic adults, in

predictive learning processes, as proposed by Bayesian and predictive coding accounts of autism. These accounts, which describe broad atypicalities in predictive processes, have not been extensively tested from a social or motor learning perspective. Therefore, I examine these theories in the context of social and motor learning, with the aim of determining whether these models are restricted to the perceptual domain, or whether they outline differences in *domain-general* learning processes. The first empirical chapter (Chapter 2) examines whether dopamine-dependent predictive processes are atypical during implicit motor sequence learning in autistic, compared to non-autistic adults, using a probabilistic motor learning paradigm. As autism is primarily associated with differences in social functioning, Chapter 3 focuses on learning in a social context. Here, I investigate whether social learning is atypical in autism spectrum disorder and quantify how autistic adults learn from individual and social information sources simultaneously in a reward-learning paradigm, using computational models of learning. This paradigm allows investigation of whether learning atypicalities in autism relate to social learning *per se*, or the *modulation* of learning as a function of environmental volatility (as hypothesised by some predictive accounts of autism). In addition to investigating learning mechanisms in autism, Chapter 3 also has the power to provide insight into the mechanisms underpinning social and individual learning. That is, differences in social, but not individual learning in autistic individuals in Chapter 3 would comprise evidence of dissociable mechanisms for social and individual learning.

The second part of this thesis focuses in detail on one of these learning types - social learning, with the aim of addressing an extant issue in this field: are the neurochemical mechanisms that underpin social learning dissociable from the neurochemical mechanisms that underpin learning from one's own individual experience (individual learning). In Chapter 4, I examine whether social and individual learning can be dissociated as a function of naturally occurring genetic variation in genes important for the regulation of monoamine signalling. Chapter 5 describes the development and piloting of an adapted version of a social learning task. In studies demonstrating dissociations between social and individual learning, learning types differ both in terms of social nature (social or individual) and rank (primary versus secondary status), meaning that social nature and rank are confounded. This chapter is therefore written with a focus on describing the design and development of a task where social versus individual and primary versus secondary status are orthogonalized. In Chapter 6, the behavioural task developed in Chapter 5 is employed in a pharmacological intervention,

where dopamine signalling is manipulated. This chapter investigates whether the dissociable effects of dopaminergic manipulation on different learning types are better explained by primary versus secondary status, than by social versus individual nature.

The final chapter will summarise and integrate all reported results and outline future directions to further investigate the ways that Bayesian and predictive coding accounts of autism can, and cannot, be extended across all domains, and the role of both dopamine and serotonin in predictive learning. Furthermore, this thesis aims to shed light on the contexts in which learning is, and is not, atypical in adults with autism.

Chapter 2: Intact predictive motor sequence learning in autism spectrum disorder

This chapter presents a published study, examining whether predictive processes are atypical during implicit motor sequence learning in autistic, compared to non-autistic adults, using a probabilistic motor learning paradigm.

Supplementary materials for this chapter can be found in Appendix 1.

Publication 1:

Intact predictive motor sequence learning in autism spectrum disorder

Alicia J. Rybicki, Joseph M. Galea, Bianca A. Schuster, Chole Hiles, Cleo Fabian, and Jennifer L. Cook

Scientific Reports **11**:20693 (2021). DOI: 10.1038/s41598-021-00173-1

Abstract

Background. Atypical motor learning has been suggested to underpin the development of motoric challenges (e.g., handwriting difficulties) in autism. Bayesian accounts of autistic cognition propose a mechanistic explanation for differences in the learning process in autism. Specifically, that autistic individuals overweight incoming, at the expense of prior, information and are thus less likely to a) build stable expectations of upcoming events and b) react to statistically surprising events. Although Bayesian accounts have been suggested to explain differences in learning across a range of domains, to date, such accounts have not been extended to motor learning.

Methods. 28 autistic and 35 non-autistic controls (IQ > 70) completed a computerised task in which they learned sequences of actions. On occasional "surprising" trials, an expected action had to be replaced with an unexpected action. Sequence learning was indexed as the reaction time difference between blocks which featured a predictable sequence and those that did not. Surprise-related slowing was indexed as the reaction time difference between surprising and unsurprising trials.

Results. No differences in sequence-learning or surprise-related slowing were observed between the groups. Bayesian statistics provided anecdotal to moderate evidence to support the conclusion that sequence learning and surprise-related slowing were comparable between the two groups.

Conclusions. We conclude that individuals with autism do not show atypicalities in response to surprising events in the context of motor sequence-learning. These data demand careful consideration of the way in which Bayesian accounts of autism can (and cannot) be extended to the domain of motor learning.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder, characterized by restricted and repetitive interests and difficulties with social communication and interaction (American Psychiatric Association, 2013). While not considered a core diagnostic feature, over recent years the study of autistic body movements has gained traction [267]–[269] and differences in the way autistic and non-autistic people move have been documented [145]. A number of studies have suggested that movement challenges in autism could stem from atypicalities in the motor learning process [136],[137]. Evidence to support this comes from serial reaction time tasks, wherein participants execute a sequence of discrete movements over repeated trials, with motor sequence learning indexed as a reduction in response time for learned sequences [161],[270]. Several studies report atypical sequence learning in autism [148],[165]. Thus, a small but growing literature suggests that differences in autistic body movements may lie, not in the execution of learned movements, but in the learning process itself.

The claim that motor learning is different in autism resonates well with a broader literature arguing for general learning atypicalities. Current prominent accounts of autism [53],[78] propose that major characteristics can be explained by differences in Bayesian inference. Under Bayesian [55], specifically predictive coding frameworks [27],[56], perception and learning are based on the construction of hierarchical probabilistic models of the environment. These models are updated when top-down prior predictions are compared with incoming, sensory information, and the difference between the two (prediction error) is used to update the prior. Relative confidence in the prediction error and prior determines how much weight is afforded to each, and thus the extent to which beliefs are updated by incoming information versus prior knowledge. Bayesian and predictive coding accounts (referred to collectively as 'Bayesian accounts' from hereon [13],[53],[54]) propose that autism is characterised by atypical weighting of prior beliefs relative to incoming sensory information [12],[13],[76],[78]. In support of this, research has demonstrated that autistic perception and learning are dominated by incoming sensory information, with less reliance on top-down priors [103],[119],[271].

In principle, Bayesian accounts detail a general mechanism underpinning autistic processing which should apply to various domains of functioning. In the motor learning domain, the hypothesis that autistic individuals show underutilization of priors leads to specific and

testable predictions. The motor system uses prior experience to prepare motor output for an event by an amount that is proportional to the probability of the event [167]. Thus, as the precision of an individual's expectations about an upcoming action increases, reaction time (RT) decreases. However, if expectations are violated, RT increases (i.e., **surprise-related slowing** occurs) due to the requirement to halt the prepared action and prepare and execute the surprising action. Bayesian accounts of autism predict that the underutilisation of priors results in an aberrantly high baseline level of surprise. Thus, surprising events, which violate expectations, are not as surprising for autistic relative to non-autistic individuals [12],[78]. According to such accounts, during motor learning surprise-related slowing should be reduced in autism (i.e. a more efficient response to surprising events should be observed), at the expense of learning a sequence and forming strong prior predictions about upcoming events [11]. To date, this has mainly been tested by demonstrating atypical surprise-related slowing with respect to *perception*. Lawson and colleagues [100], for example, found that, relative to non-autistic controls, autistic adults showed reduced surprise in response to unexpected visual stimuli. It is currently not clear, however, whether Bayesian accounts of autism apply to motor learning. If so, they would help to shed light on the computational mechanisms underpinning differences in autistic motor learning.

Preliminary evidence for atypical surprise-related slowing in autism comes from several studies: Rinehart and colleagues [168] required participants to execute button presses in response to a visual pattern, where surprising deviations from the pattern sporadically occurred. Relative to non-autistic children, autistic individuals showed a significant reduction in surprise-related slowing. Similarly, Gidley Larson and colleagues [147], required participants to learn and execute a pattern of movements to anticipate the motion of a moving target. In several trials, expectations were violated by altering the pattern, inducing surprise-related slowing for non-autistic, but not for autistic, children. In line with Bayesian predictions, these studies suggest reduced surprise-related slowing in ASD. If underutilisation of priors is a pervasive style of autistic processing that cannot be unlearned or compensated for, establishing these effects in autistic adults is of central importance. However, to date, only one study has been carried out with adults, with results showing that participants were faster to respond to expected (referred to by the authors as validly cued) compared to unexpected (invalidly cued) events, there was no interaction between group (ASD versus control) and the validity of the cue type [139]. In sum, very few studies have investigated surprise-related slowing in ASD, and a coherent pattern of results has not emerged.

Extant studies of surprise-related slowing in autism have used paradigms with two response options (i.e. respond left or right [168], trace circle or square [147], change either hand or direction of movement [139]). A disadvantage of two-option paradigms is that, compared to multi-option paradigms, sequence-learning cannot be investigated. According to Hick's Law [272], the number of response options is logarithmically related to decision time, thus average RTs for a four-option paradigm in which each option is equally likely are 500-700 milliseconds (ms), whereas a single response elicits an RT around 250-350ms [273]–[275]. If, however, there is a sequence to the responses, one can reduce a four-option paradigm to a single response (with an associated probability) using prior knowledge of the sequence [161], resulting in a greater RT reduction (i.e. a sequence-related speeding effect). The more potential options, the greater the potential speeding effect [276]. Since the extant autism literature has focused on two-option paradigms, it is not clear whether reduced surprise-related slowing (that is, a more efficient response to surprising events), comes at the expense of sequence-learning.

Here we compared the performance of autistic and non-autistic adults on a motor sequence-learning task [277]. Participants learned associations between four visual stimuli and four unique actions. In an 'easy-predictable condition', actions followed a simple sequence with occasional surprising trials where an unpredictable action was required. The same was true of the 'difficult predictable condition', although with a more challenging sequence. In the 'unpredictable condition' there was no sequence to learn. This task thus provides indices of sequence-learning, indexed by sequence-related speeding (the difference in RT between predictable and unpredictable conditions) and surprise-related slowing (the difference in RT between surprising and unsurprising trials in the predictable conditions). We predicted that 1) autistic adults would exhibit a *less* efficient response to unsurprising events, indexed by decreased sequence learning relative to non-autistic participants and 2) autistic adults would show a *more* efficient response to surprising events, indexed by a reduction in surprise-related slowing relative to non-autistic controls.

Methods

Participants

Twenty-eight adults with a clinical diagnosis of ASD (18-57 years, mean (standard deviation (SD)) age = 29.8 (10.2); 15 female), previously diagnosed by a UK National Health Service (NHS) or privately registered clinician who worked independently from our research group, according to the DSM ([1]) or ICD-10 [278] criteria, and 35 healthy non-autistic controls (18-57 years, mean (SD) age = 27.6 (10.5); 13 female) were recruited from Birmingham and surrounding areas through advertising via posters and social media (see Table 1 for full demographic details and Supplementary Methods for full clinical details). All participants were reimbursed for their time (at a rate of £10 per hour) and travel expenses. ASD diagnosis was confirmed with administration of the Autism Diagnostic Observation Schedule, second edition (ADOS-2) [279] by a trained researcher, using the current standard scores for a diagnosis of ASD, whereby a minimum score of 7 is the cut-off for designation as "on the autism spectrum," and a minimum score of 10 is the cut-off for being designated as "autistic" (see Supplementary Methods for further inclusion criteria). The study was approved by the University of Birmingham local ethics committee (ERN_160281AP1R) and was conducted in accordance with the Declaration of Helsinki.

Table 2-1 Demographic information

	Control group (n = 35)	ASD group (n = 28)			
	Mean (SD)	Mean (SD)	t (1,61)	X^2 (1, N = 63)	p
Gender (*n* males: *n* females)	22:13	13:15		1.700	0.192
Age	27.6 (10.5)	29.8 (10.2)	1.496		0.140
Training trials	178.86 (177.16)	222.86 (133.19)	1.090		0.280
2-subscale IQ	107.51 (13.17)	108.679 (16.31)	0.314		0.755
Autism-Quotient (AQ)	15.09 (8.42)	36.46 (8.08)	10.196		<0.001
Toronto Alexithymia Scale- II (TAS-2)	43.03 (10.67)	64.11(10.51)	7.842		<0.001
ADOS total scores		7.64 (3.29)			

Note: IQ was assessed with the Wechsler Abbreviated Scale of Intelligence-2 (WASI-2). SD refers to standard deviation. Training trials, IQ and gender did not significantly differ between the groups.

General procedure

In a single session, participants first provided written, informed consent; second, completed the Autism-Quotient (AQ) questionnaire [124], Toronto Alexithymia Scale- II (TAS-20)[280] and an Intelligence Quotient (IQ) test (two-item subscale of the Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II) [281], administered by a trained researcher (1

hour); participants finally completed the serial reaction time task (1 hour) followed by ADOS administration (1-2 hours). All interviews were filmed for validation and training purposes.

Serial reaction time task

Participants completed a probabilistic serial reaction time task, widely used to investigate motor sequence-learning [277],[282]. Participants were instructed to place their index, middle, ring and little fingers on the keyboard letters V, B, N and M respectively. Subsequently, participants were required to learn associations between four imperative stimuli (IS) and four specific finger press actions (Fig. 1a) such that, for example, the appearance of a particular IS became associated with pressing the index finger down on the letter V. Participants were instructed to respond as quickly as possible to the IS. IS order followed different sequences depending on the condition, with predictable and unpredictable sequences presented in different blocks (Fig. 1b). A training period preceded the main experiment, in which participants learned associations between the IS and the specific actions. Subsequently participants completed seven blocks of 100 trials with self-paced rest intervals between the blocks. The task required approximately 45 minutes to complete in total. For the unpredictable condition, there was an equal probability of each IS appearing on each trial (Fig. 1c). For the easy-predictable condition the sequence followed a pattern in which IS order 1-2-3-4 occurred with high probability (Fig. 1d). For the difficult-predictable condition stimuli followed a more complicated pattern whereby the stimuli order 1-4-2-3 occurred with high probability. Surprising/unpredictable stimuli, which violated the sequence, occurred at a low probability, forcing participants to respond against their prior knowledge of the sequence, i.e., replace an expected action with an unexpected action. Surprising trials only occurred during the predictable blocks. The pattern or sequence in each block was not explicitly described. See Supplementary Methods for a more detailed task description and the participant instruction script.

Figure 2.1 Behavioural task

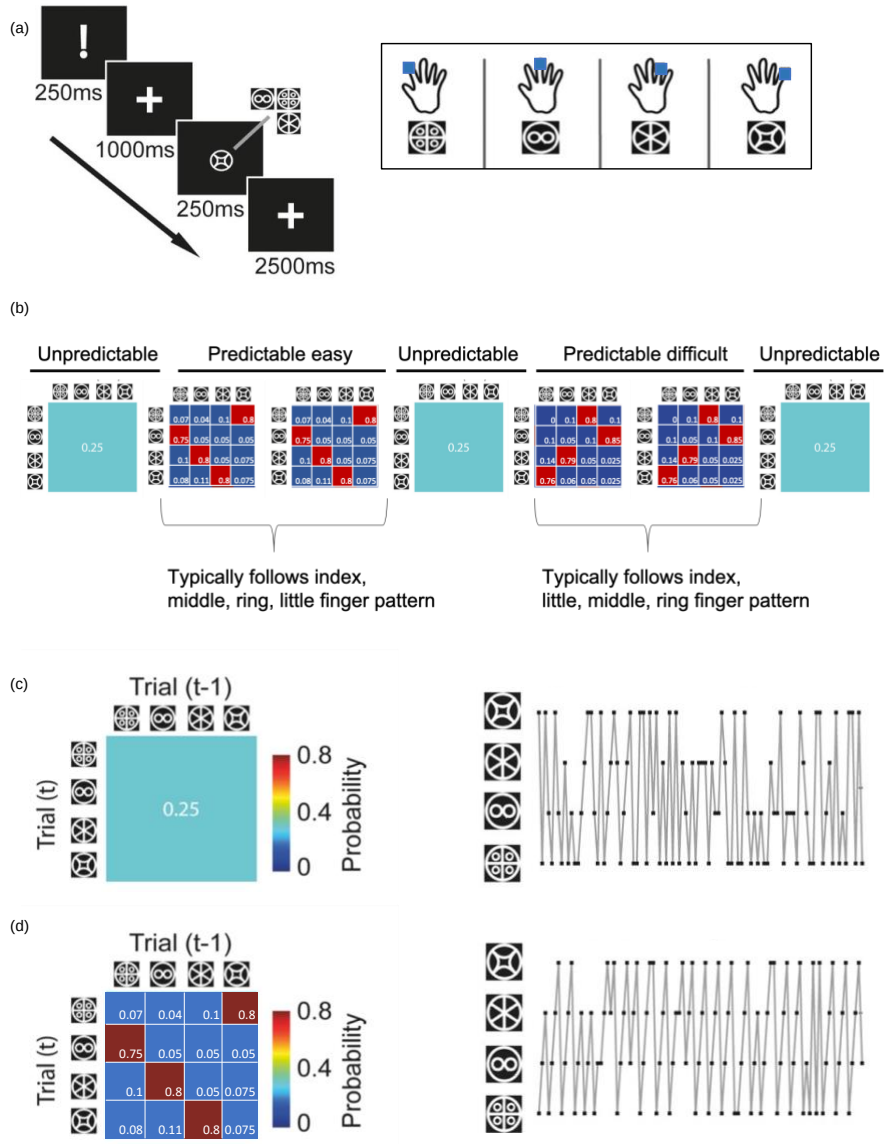


Figure 1. Behavioural task (a) Representation of a single trial. Participants observed a warning signal, followed by a fixation cross, then one of four different imperative stimuli (IS) and another fixation cross. Participants were advised to react as quickly as possible without sacrificing accuracy. Each of the four imperative stimuli corresponded to a specific finger press. (b) Overall task structure. Each participant completed seven blocks. Block order was the same for all participants. (c) Unpredictable condition. Transition matrix: All IS followed each other with equal probability, resulting in an unpredictable sequence. (d) Predictable condition. Sequences were generated from a first-order Markov sequence whereby numbers within the probability matrix represent the transition probabilities (16 possible combinations) and determined the relationship between the IS on trial t and trial t-1. The easy-predictable sequence is displayed here. Adapted from "Action reprogramming in Parkinson's disease: response to prediction error is modulated by levels of dopamine" by J. M. Galea et al, 2012, Journal of Neuroscience, 32(2):542-50. Copyright (2012) Galea et al.

Statistical analysis

All analyses were conducted with MATLAB 2018b (MathWorks, Inc) and JASP (JASP Team, 2019). Raw and collated data and analysis scripts can be accessed at OSF (tiny.cc/58oxsz). RT was calculated as the time in milliseconds (ms) between onset of the IS and response (button press), for correct responses only. Since error rates and log-transformed error rates violated the assumptions for parametric testing (Levene's test scores significantly differed from zero ($p = 0.010$)), Inverse Efficiency scores [IES; 283] were instead used to account for possible speed-accuracy trade-offs. IES scores comprised the RT divided by 1 minus the proportion of correct responses (Supplementary Methods). To test whether autistic and non-autistic adults exhibit comparable sequence learning, RTs and IES were averaged for each condition and submitted to repeated-measures analysis of variance (RM-ANOVA), with the within-subject factor condition (easy-predictable, difficult-predictable and unpredictable) and between-subjects factor group (ASD, control). To test whether autistic and non-autistic adults exhibit comparable surprise-related slowing, RTs were calculated separately for surprising (probability > 0.75) and unsurprising (probability < 0.75) trials (Fig. 1d) separately for the predictable easy and difficult conditions and submitted to a RM-ANOVA with within-subject factors surprise (surprising, unsurprising), condition (easy-, difficult-predictable) and between-subjects factor group (ASD, control). To investigate whether the temporal evolution of surprise-related slowing differed between groups we modelled the effects of surprise on RT on a trial-by-trial basis. That is, alongside trial number and group, trial-by-trial surprise was included as a factor in a multiple regression analysis with RT as the dependent variable. The (trial specific) surprise of observing IS type *i* on trial *t* after experiencing IS type *j* on trial *t*-1 was calculated as the negative log of the IS pair's (*ji*) predicted joint probability, with the joint probability of an IS pair on a given trial estimated from the number of previous occurrences of this IS pair on the preceding trials. Surprise thus represented the unexpectedness of the current IS (*i*) given the previous IS (*j*) based on the number of previous trials in which *j* preceded *i* [277] (Supplementary Methods). Multiple regression was performed for each participant for each condition, with standardised β values then averaged across all participants within each group.

A threshold of $p<0.05$ was used for all statistical tests, with significant effects investigated with Bonferroni-corrected post-hoc t-tests. For all analyses, Bayesian statistical testing was

implemented as a supplement to null hypothesis significance tests. Bayes inclusion factors (BF_{incl}) were included for all RM-ANOVAs, representing the evidence given the observed data for including a certain predictor in the model (see Supplementary information for a full description of Bayesian analyses). For example, an inclusion Bayes factor for an effect of 3 for a given predictor *i* can be interpreted as stating that models which include the predictor *i* are 3 times more likely to describe the observed data than models without the predictor.

An *a priori* power analysis was carried out, based on research by Galea and colleagues [277]. Results showed that, based on an achieved effect size of $\eta^2p = 0.193$, a minimum sample size of 22 participants per group was required. We complemented our *a priori* power analysis with *post hoc* sequential Bayesian testing. For our main effect of interest (surprise-related slowing), we continued data collection (and accumulation of evidence) until we had sufficient certainty about the absence of a group difference, i.e., the relative evidence for H0 plateaued above 3 (Suppl. Figure 1), representing moderate evidence for no group differences.

Results

Autistic and non-autistic adults exhibit comparable sequence learning

To test our hypothesis that we would observe reduced sequence learning in the autistic group, as indexed by a reduced RT difference between predictable and unpredictable conditions relative to the control group, we submitted RTs to a RM-ANOVA with within-subject factor condition (easy-predictable, difficult-predictable, and unpredictable) and between-subjects factor group (ASD, control). This revealed a significant main effect of condition ($F(2,122) = 28.804$, $p < 0.001$, $\eta^2 = 0.028$, $BF_{incl} = 1.862e+8$) (Fig. 2a), no main effect of group ($F(1,61) = 0.518$, $p = 0.474$, $\eta^2 = 0.008$, $BF_{incl} = 0.423$) and no interaction between condition and group ($F(2,122) = 0.429$, $p = 0.652$, $\eta^2 = 0.000$, $BF_{incl} = 0.182$). Post-hoc Bonferroni-corrected t-tests demonstrated that RT was significantly lower for the predictable-easy (mean (standard error) $\bar{x}$ ($\sigma_{\bar{x}}$) = 660.350 (11.438)) compared to the predictable-difficult ($\bar{x}$ ($\sigma_{\bar{x}}$) = 697.318 (12.062), $t(61) = -6.707$, $p < 0.001$, $d = -0.845$) and unpredictable conditions ($\bar{x}$ ($\sigma_{\bar{x}}$) = 687.624 (11.316), $t(61) = -5.156$, $p < 0.001$, $d = -0.65$). Although RTs for the unpredictable and predictable-difficult conditions differed - with lower RTs for the unpredictable condition – this difference did not reach statistical significance ($t(61) = 2.018$, $p_{bonf} = 0.137$). In sum, we observed lower RTs for the predictable-easy compared to the unpredictable condition,

suggesting that sequence learning enabled participants to speed their responses. However, we observed no significant speeding for the predictable-difficult condition, thus raising the possibility that the sequence was too challenging for participants to learn. Finally, the lack of a main effect of group suggested that the groups do not differ with respect to sequence-learning. This result was strengthened by Bayesian independent t-tests, with the BF indicating moderate evidence for H_0 for the easy ($BF_{01} = 3.193$) and anecdotal evidence for the difficult ($BF_{01} = 2.371$) predictable conditions (Figs. 3a-3b).

Figure 2.2 Reaction time and Inverse Efficiency Scores

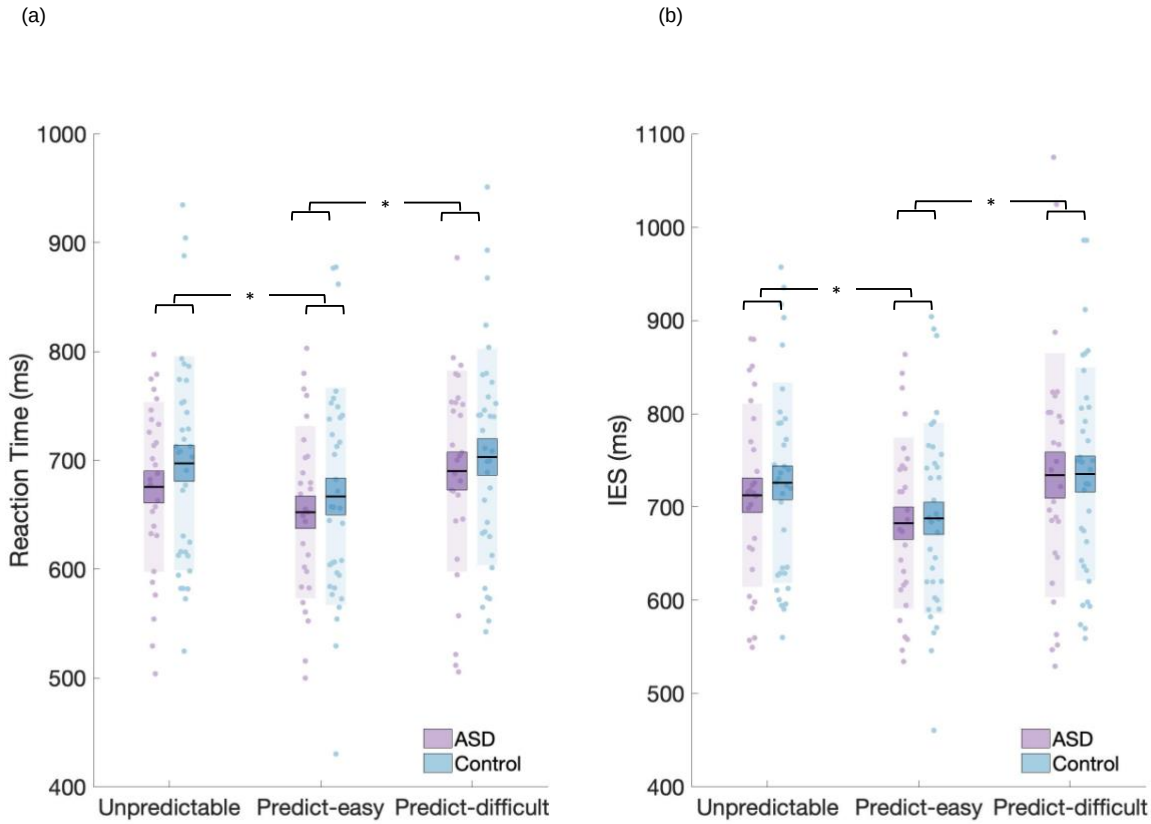


Figure 2. (a) Reaction time (RT). A significant difference in RT was observed between both the easy-predictable condition and the difficult-predictable and unpredictable conditions. ASD and control groups did not significantly differ in RT for any of the three conditions. (b) Inverse Efficiency Scores (IES). IES scores varied as a function of condition; no differences between groups were observed. Data points indicate individual participants. The mean is the thick black horizontal line and 1 standard error of the mean (SEM) is represented by the shaded box around the mean. Standard deviation (SD) is the shaded region.

Figure 2.3 Bayesian statistical testing

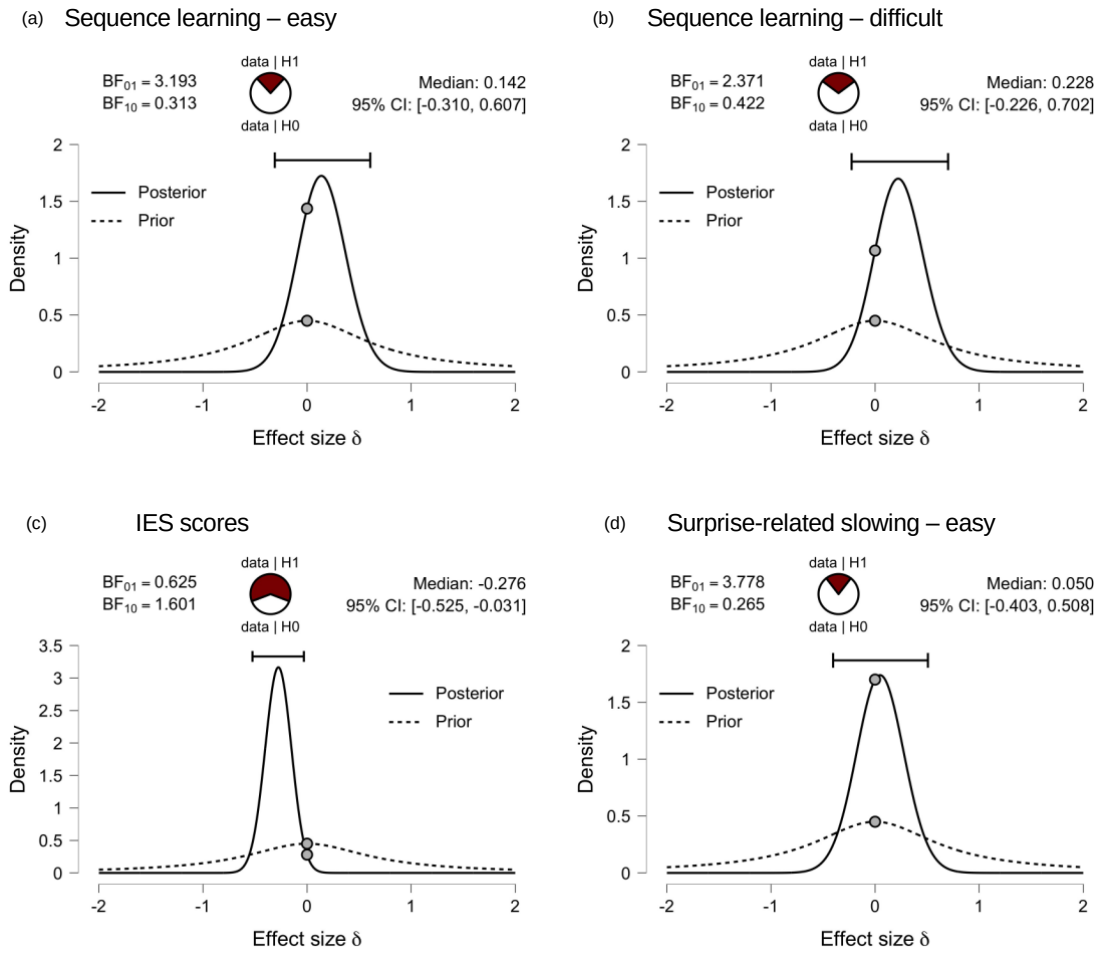


Figure 3. Bayesian statistical testing. The Bayes factor (BF_{01}) indicates the evidence for the null hypothesis of no difference between groups. The density distribution displays the prior and posterior distribution for the population effect size, with the median effect size estimated, and a 95% credible interval which contains the median effect size. **(a)** Sequence learning – easy**:** BF_{01} = 3.193, meaning that the data are over three times more likely under H0 and provide moderate support for null hypothesis of no difference between groups **(b)** Sequence learning – difficult: BF_{01} = 2.371, meaning that the data are over two times more likely under H0 and provide anecdotal support for the null hypothesis of no difference between groups. **(c**) Bayesian paired t-test for IES scores in the predictable difficult condition compared with the unpredictable condition. BF_{01} = 0.625, meaning that the data are more likely under H1, providing weak support for the alternative hypothesis of a difference between conditions. **(d)** Surprise-related slowing – easy condition**.** BF_{01} = 3.778, meaning that the data are over three times more likely under H0 and provide moderate support for null hypothesis of no difference between groups.

A RM-ANOVA with within-subject factor condition (easy-predictable, difficult-predictable and unpredictable) and between-subjects factor group (ASD, control) and IES as dependent

variable revealed a main effect of condition ($F(2,122) = 25.078$, $p < 0.001$, $\eta^2 = 0.036$, $BF_{incl} = 1.585e+7$) and no main effect of group ($F(1,61) = 0.064$, $p = 0.801$, $\eta^2 = 0.001$, $BF_{incl} = 0.509$) or group by condition interaction ($F(2,122) = 0.377$, $p = 0.687$, $\eta^2 = 0.001$, $BF_{incl} = 0.137$) (Fig. 2b). Post-hoc tests demonstrated that IES were significantly lower for the predictable-easy ($\bar{x}$ ($\sigma_{\bar{x}}$) = 685.502 (12.222)) compared to both the unpredictable ($\bar{x}$ ($\sigma_{\bar{x}}$) = 719.985 (12.938), $t(61) = -5.357$, $p < 0.001$) and the predictable-difficult condition ($\bar{x}$ ($\sigma_{\bar{x}}$) = 734.790 (15.226), $t(61) = -5.942$, $p < 0.001$). IES for the unpredictable compared to predictable difficult conditions did not significantly differ ($t(61) = -2.155$, $p = 0.099$, Cohen's $d = 0.271$). However, a Bayesian paired t-test provided weak evidence that IES for the unpredictable blocks differed from the predictable difficult condition ($BF_{01} = 0.625$, Fig. 3c). Thus, RT, after correcting for number of errors, was higher during the predictable difficult, relative to the unpredictable, condition adding support for a lack of sequence learning during the difficult-predictable condition. The lack of a main effect of group, or interaction between group and condition suggests that the groups did not differ in the ability to execute the appropriate action. In sum, we did not find evidence to support the hypothesis that, relative to controls, autistic adults exhibited decreased sequence learning.

Autistic and non-autistic adults exhibit comparable surprise-related slowing

To test our second hypothesis that, relative to non-autistic controls, autistic participants would exhibit a reduction in surprise-related slowing - as indexed by a reduced RT difference between surprising and unsurprising trials - we submitted RTs to a RM-ANOVA with within-subject factors surprise (surprising trials, unsurprising trials) and condition (easy-predictable, difficult-predictable), and between-subjects factor group (ASD, control). We observed main effects of surprise ($F(1,61) = 34.144$, $p < 0.001$, $\eta^2 = 0.017$, $BF_{incl} = 1.166e+13$), condition ($F(1,61) = 34.144$, $p < 0.001$, $\eta^2 = 0.017$, $BF_{incl} = 4.526e+13$) and a surprise by condition interaction ($F(1,61) = 72.325$, $p < 0.001$, $\eta^2 = 0.022$, $BF_{incl} = 3.714e+8$). Post-hoc comparisons revealed an increase in RT for surprising compared to unsurprising trials for the easy-predictable (surprising: $\bar{x}$ ($\sigma_{\bar{x}}$) = 703.782 (11.744), unsurprising: $\bar{x}$ ($\sigma_{\bar{x}}$) = 649.278 (11.575), mean difference = 54.340; $t(61) = 9.850$, $p < 0.001$, $d = 1.241$) (Fig. 4a) but not the difficult-predictable condition (surprising: $\bar{x}$ ($\sigma_{\bar{x}}$) = 694.875 (12.011), unsurprising: $\bar{x}$ ($\sigma_{\bar{x}}$)= 697.835 (12.241), mean difference = -3.656; $t(61) = -0.663$, $p = 1.000$, $d = -0.083$) (Fig. 4b). Crucially, no main effect of group ($F(1,61) = 0.493$, $p = 0.485$, $\eta^2 = 0.008$, $BF_{incl} = 0.326$),

surprise by group ($F(1,61) = 0.797$, $p = 0.375$, $\eta^2 = 0.000$, $BF_{incl} = 0.267$), condition by group ($F(1,61) = 0.795$, $p = 0.980$, $\eta^2 = 0.000$, $BF_{incl} = 0.199$) or surprise by condition by group ($F(1,61) = 0.493$, $p = 0.485$, $\eta^2 = 0.000$, $BF_{incl} = 0.088$) interactions were observed. To ensure that the lack of group difference could not be attributed to differences in baseline speed, we re-ran the analysis with baseline-corrected mean RT scores. This did not change the observed pattern of results, with no main/interaction effect(s) of group observed (all p-values > 0.05, all $\eta^2 < 0.001$, all $BF_{incl} < 1$). Indeed, no differences in motor execution overall were observed between groups (Supplementary Results).

Figure 2.4 Surprise-related slowing

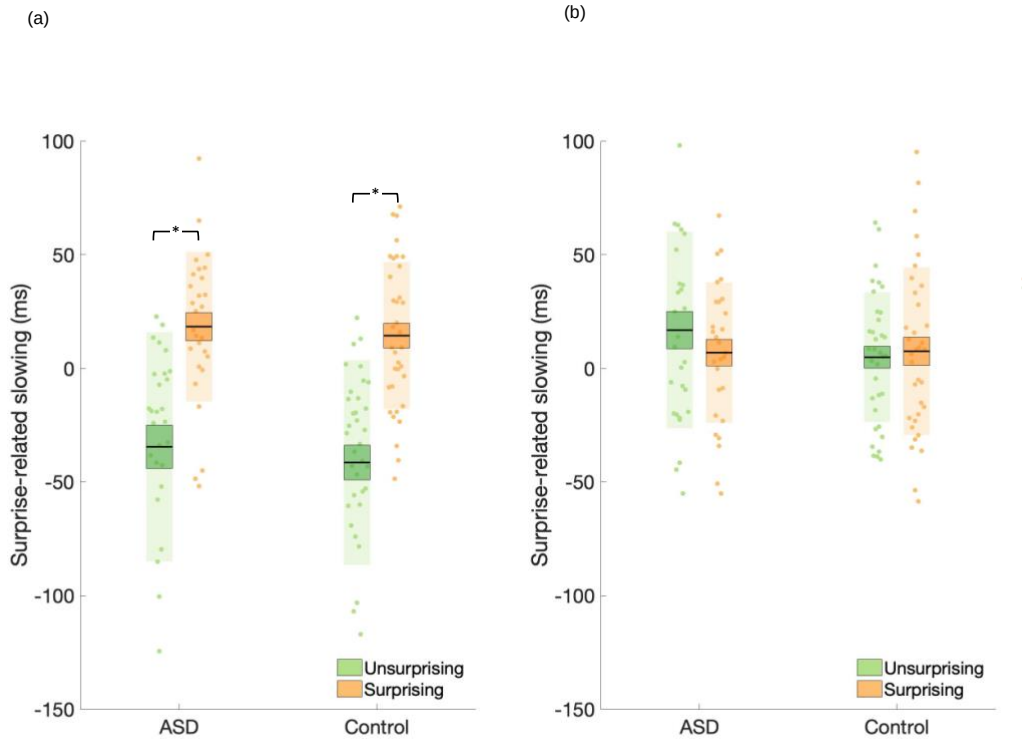


Figure 4. Surprise-related slowing (a) Easy-predictable condition**.** Data represent the difference between the mean reaction time (RT) for the unpredictable conditions and RT for surprising (orange) and unsurprising (green) trials. RT was significantly greater for surprising compared to unsurprising trials**. (b)** Difficult-predictable condition. There was no difference in RT between surprising and unsurprising trials in the difficult condition. No differences between the ASD and control group were observed in either condition. Data points indicate individual participants. The mean is the thick black horizontal line and 1 standard error of the mean (SEM) is represented by the shaded box around the mean. Standard deviation (SD) is the shaded region.

Surprise-related slowing scores for the two groups in the easy-predictable condition were compared using a Bayesian independent t-test. The BF_{01} was equal to 3.778, indicating the

data were approximately 3.8 times more likely under the hypothesis that groups did not differ with respect to surprise-related slowing and providing moderate evidence for H_0 (Fig. 3d).

Finally, although IES scores varied significantly across condition ($F(1,61) = 16.538$, $p < 0.001$, $\eta^2 = 0.011$, $BF_{incl} = 106.780$), surprise ($F(1,61) = 43.141$, $p < 0.001$, $\eta^2 = 0.031$, $BF_{incl} = 1.905e+7$) and condition by surprise ($F(1,61) = 56.538$, $p < 0.001$, $\eta^2 = 0.026$, $BF_{incl} = 2.649e+7$), no main/interaction effect(s) of group were observed (all p-values > 0.05, all $\eta^2 < 0.001$, all $BF_{incl} < 1$) (Suppl. Figs. 2a-2b).

Trial-by-trial surprise did not differ between groups

The observed results demonstrated typical surprise-related slowing in autistic individuals. However, the above analyses collapse data across all trials within each condition and cannot detect differences in the temporal progression of surprise-related slowing, nor reveal differences between the groups in the speed of acquisition of surprise-related slowing. Trial-by-trial surprise was therefore included as a predictor in a multiple regression analysis, alongside trial number and condition. Standardized beta values (β) for the main and interaction effect(s) of predictors (Table 2) were compared using one-sample t-tests to determine if they were significant predictors of RT. β values that significantly differed from zero were averaged across each group and compared using standard and Bayesian independent sample t-tests. If differences in the temporal progression of surprise-related slowing existed between groups, we would expect to observe a significant difference in β values relating to the interaction between surprise and trial number. However, no differences in β values were observed between groups for this interaction ($t(61) = 1.130$, $p = 0.263$, $d = 0.287$, $BF_{01} = 2.260$), nor for β values for condition ($t(61) = -0.022$, $p = 0.983$, $d = -0.005$, $BF_{01} = 3.868$), trial-by-trial surprise ($t(61) = 0.905$, $p = 0.369$, $d = 0.229$, $BF_{01} = 2.739$) or surprise by condition ($t(61) = 1.191$, $p = 0.238$, $d = 0.302$, $BF_{01} = 2.130$). In summary, no differences in the temporal evolution of surprise-related slowing were observed between groups.

Table 2-2 Predictors of reaction time

	Standardized β values	Standard Error β values	t (62)	p	Cohen's d
Time	1.430	1.776	0.805	0.424	0.101
Condition	18.793	2.821	6.663	< 0.001***	0.839
Surprise	9.707	1.598	6.076	< 0.001***	0.766
Time x condition	0.026	2.247	0.012	0.991	0.001
Time x surprise	1.116	1.120	0.996	0.323	0.126
Condition x surprise	-10.432	1.414	-7.3766	< 0.001***	-0.929
Time x condition x surprise	1.612	1.300	1.239	0.220	0.156

Note: Standardized beta values (β) are displayed for significant and non-significant predictors of reaction time.

ADOS severity scores as a predictor of surprise-related slowing and sequence-learning

Focusing specifically on the easy-predictable condition, where surprise-related slowing and sequence-learning effects were observed, correlation analysis showed that ADOS scores were not a significant predictor of surprise-related slowing ($r = -0.324$, $F(1,26) = 3.042$, $p = 0.093$) or sequence-learning ($r = 0.070$, $F(1, 26) = 0.129$, $p = 0.723$). Furthermore, neither AQ nor TAS scores were significant predictors of behavioural measures (Supplementary Results).

Discussion

Here we investigated the underutilisation of priors in ASD in the context of a motor sequence-learning task. In predictable conditions, actions largely followed a pre-defined

sequence with infrequent surprising violations of this sequence. In the unpredictable condition, there was no sequence to learn. In line with Bayesian accounts of autism, we hypothesised that autistic adults would show a more efficient response to surprising events at a cost to sequence learning, indexed by a reduction, relative to non-autistic controls, in surprise-related slowing alongside decreased sequence-learning. Contrary to our predictions, there were no significant differences between autistic and non-autistic adults in terms of surprise-related slowing or sequence learning. Furthermore, Bayesian statistics provided anecdotal to moderate evidence to support the conclusion that the groups were comparable with respect to both measures.

The lack of a difference between the groups departs from the predictions of Bayesian accounts of autism. One potential explanation for this conflict is that our sample might not be representative of the populations typically used to test these accounts. In opposition to this, we argue that our sample is comparable in terms of age, IQ and average ADOS score to a number of studies that have found evidence in support of the underutilization of priors in ASD [100],[107],[108],[173]. Thus, suggesting that the level of autism symptomatology, age or IQ of our participants is unlikely to explain the observed null results, though we note that comparison between studies is challenging due to the use of different paradigms. In addition, Bayesian analyses revealed that we had anecdotal evidence to support the null hypothesis that there is no correlation between ADOS scores and the extent of surprise-related slowing. This suggests that recruiting a more diverse sample is unlikely to alter the observed results. Indeed, if the relationship between ADOS and surprise-related slowing is linear, we would not expect different results with a broader sample. Nevertheless, we acknowledge that this claim requires empirical testing since at present we have only anecdotal evidence for the lack of a relationship between ADOS and surprise-related slowing and, furthermore, it is possible that a non-linear relationship exists (e.g., there could be a step change in surprise-related slowing with increasing ADOS score). Consequently, we can confidently conclude, based on our Bayesian and frequentist analyses that in our sample (with age, AQ and ADOS ranges of 18-57, 17-48 and 2-14 respectively) there is no evidence of underutilisation of priors. Further empirical testing would be necessary to be confident that this conclusion extends to samples of the autistic population with different characteristics.

A further potential explanation for the conflict between the current results and the predictions from Bayesian accounts of autism is that our task does not really index the process of evaluating and updating priors predicted by Bayesian accounts of autism. The motor

sequence featured in the easy-predictable condition is easily executable and could potentially be explicit in nature. Some authors have argued that explicit motor learning relies less on priors and prediction errors and more on target-driven error derived from an explicit strategy, although results thus far derive from sensorimotor adaptation tasks [284],[285]. However, the current paradigm employed a probabilistic sequence learning structure, frequently used as a measure of *implicit* learning, whereby a predictable sequence is frequently interspersed with "surprising" stimuli [286]. Such surprising trials decrease explicit knowledge of the sequence in comparison to a fixed or deterministic sequence [287],[288]. Additionally, Galea and colleagues [277],[282], demonstrated that the (dopamine-dependent) prediction error process is central to this task, observing increased surprise-related slowing in the context of the same motor sequence learning task in adults with Parkinson's disease when off- compared to on-medication [277] and under dopamine antagonism in healthy adults [282]. Consequently, we believe that our task provides a good measure of the utilisation of priors.

One might ask whether the logical conclusion from our results is that Bayesian accounts of autism do not apply in the motor domain. Indeed, research relating to Bayesian accounts of autism has primarily focused on sensory/perceptual processing [58],[103],[119],[271],[289], leading to the possibility that Bayesian accounts of autism are restricted to these domains. However, Bayesian accounts have been proposed as a general principle of information processing across various domains including motor function [157]. Furthermore, attenuated priors have been suggested to account for difficulties in movement preparation and planning in autism [78] and reduced slowing for unexpected movements has been demonstrated for autistic children relative to controls [147],[168]. Thus, there is little reason to believe that Bayesian accounts would be restricted to the sensory/perceptual domain.

This lack of a theoretical or empirical basis to support the conclusion that Bayesian accounts of autism do not extend to the motor domain forces us to consider alternative explanations. For example, our results clearly contrast with recent research showing attenuated surprise-related slowing (albeit in the context of learning auditory-visual as opposed to motor-motor associations) in autistic adults [100]. However, a notable difference between the current paradigm and the one employed by Lawson and colleagues is that the latter concerned a learning environment containing multiple levels of uncertainty including, probabilistic uncertainty (i.e. the auditory stimulus could be weakly, strongly or not at all predictive of the visual stimulus) and crucially, variation in the uncertainty of the learning environment itself (i.e. "environmental volatility") such that some periods featured frequent reversals in learned

associations, and others rarely featured reversals. Lawson and colleagues argue that group differences in surprise-related slowing stem from an overestimation of environmental volatility. Results were in accordance with recent updates to predictive coding accounts of autism, which propose that, while general learning is unaffected, meta-learning (learning about learning, as when one learns about the statistics (e.g. volatility) of a learning environment), is atypical [12],[77],[78]. The current paradigm contained only probabilistic uncertainty (i.e., the current action could be weakly or strongly predictive of the action in the subsequent trial), with no requirement to learn higher-order statistics about the environment. Therefore, it is possible that our task did not tap into the (meta-learning related) predictive processes that are thought to be a key point of difference between autistic and non-autistic individuals. Work from Manning and colleagues [125], however, casts doubt on this potential explanation. Using a probabilistic reward-learning paradigm which demanded learning of environmental volatility, Manning and colleagues demonstrated that autistic children successfully adapted their learning rate to suit the level of environmental volatility. To investigate whether the current (null) results are due to a lack of variation in environmental uncertainty, an adapted version of our paradigm that demands learning about environmental volatility - such as that developed by Marshall et al. [290] - could be employed.

Finally, contrasting findings could be related to different networks of brain regions recruited across different tasks. Sequence learning tasks have relatively low motor demands and do not require the acquisition of a novel movement, thus they predominantly rely on connections within the motor cortical and subcortical regions [291]. In contrast, several tasks in which performance is atypical in ASD require integration between distinct brain regions [100],[108] and thus rely on long-range connectivity. Neuroimaging studies have linked autism to alterations in the coordinated activity of distant brain regions [292]–[294]. Indeed, autism has been associated with underconnectivity for long-range cortico-cortical connections [295],[296] and theoretical accounts have linked this to the underutilisation of priors [297],[298]. Furthermore, with respect to motor function, Gowen and Hamilton [2] have argued that motor learning *per se* is not atypical in autism, however, complications arise when cross-modal integration, which relies on long-range connectivity, is required. It is, therefore, possible that the influence of top-down priors is predominantly attenuated in tasks that rely on long-range connectivity between cortical regions (e.g., [100],[299],[300]).

In summary, after considering both frequentist and Bayesian statistics, we did not find evidence for differences in surprise-related slowing or sequence learning between autistic and

non-autistic adults in a motor sequence learning task. Our results fail to provide evidence in support of straightforward predictions from Bayesian accounts of autism in the context of motor learning. Consequently, these data highlight that more nuanced Bayesian accounts of autism (potentially considering the role of factors such as meta-learning or long-range connectivity demands) are required if such accounts are to be extended to the domain of motor learning.

Chapter 3: Intact social reward learning in autistic adults

The first empirical chapter did not find predictive processing differences in autistic adults in the motor domain. This provides preliminary evidence suggesting that predictive accounts do not extend beyond sensory domains, and/or are restricted to paradigms where volatility is manipulated. To test these ideas, this chapter presents a study investigating whether autistic adults show atypical social or individual learning or atypical adjustment of learning rate to the current environmental volatility.

Supplementary materials for this chapter are in Appendix 2.

3.1 Introduction

Socio-communicative atypicalities are a core feature of autism spectrum disorder (ASD) and a prerequisite for diagnosis [1]. The development of socio-cognitive differences in autistic children has been hypothesised to stem from reduced attention to and/or motivation for social stimuli [7],[36],[301],[302]. These could lead to reduced or atypical social learning, and subsequent downstream changes in social cognition, i.e., atypicalities in social cognition could develop from differences in the social learning process itself. It is therefore of high importance to investigate whether social learning is atypical in autistic individuals, with research into social learning in autism having the potential to result in new therapies or learning aids [10].

Recent Bayesian and predictive coding accounts of autistic cognition and learning have described core symptoms as stemming from an imbalance in the precision weighting of incoming and prior information, leading to a reduction in the ability to accurately predict upcoming events and to utilise these predictions during learning [11]–[13],[50],[53]. These accounts result in testable predictions, such as an impairment in adjusting the weight of prediction errors (PEs) during learning [12] or an overreliance on incoming information [13]. In theory, these accounts describe common computational mechanisms underpinning autistic processing, i.e., domain-general learning atypicalities. Thus, atypicalities in predictive processing should be observable with regard to social learning. Indeed, some studies find evidence for atypical predictive processing in the broader social domain. For example, autistic individuals show atypical use of prior predictions both when using social information to predict action sequences of other individuals [173] and when interpreting the goals of actions within a social context [174].

When examining reward learning more generally, evidence has been found for atypical response to reward in autistic populations [31],[32],[37], particularly when learning from social (e.g. emotional faces) as opposed to non-social (monetary/objects) rewards [35],[123]. However, social learning (learning from other individuals) has not been thoroughly investigated in autism and there is limited research directly examining social learning from a predictive perspective. Preliminary evidence for atypical social learning in autism comes from research reporting that autistic traits were correlated with difficulty in integrating social information during learning. Individuals with high autistic traits (AQ scores) showed a reduction in the ability to accurately utilise social information to modulate individual learning, with this deficit magnified during volatile phases. However, this research was carried out in a non-

autistic population [188]. Hence with respect to social learning in a clinical group, it is still unclear as to whether difficulties in integrating social information are present during learning.

Aside from difficulties in social cognitive abilities, autistic individuals demonstrate intolerance to uncertainty and difficulty in modulating responses to unexpected events [83],[84]. In line with this, predictive coding accounts of autism propose that autistic individuals will have increased difficulty learning when the environment is unpredictable, with this deficit magnified when processing complex, rapidly changing cues such as social information [12],[13],[50],[76]. During learning, the precision of PEs should vary depending on the underlying environment, with PE precision (and therefore learning rates) high in volatile environments, as PEs are signalling informative changes in reward-outcome contingencies. In stable environments, however, learning rates should be low, as discrepancies are likely due to noise [67],[88]. Therefore, if flexible adjustment of precision is impaired in autism, autistic individuals should show difficulty in flexibly adjusting their learning rates to the environmental context [12],[77].

Empirical work has shown mixed results. First, altered learning was observed in autistic participants when stimulus-outcomes cues were rapidly changing (i.e., in a volatile environment), coupled with decreased behavioural measures of surprise in response to unexpected events [100]. Similar results were found using a decision-making task in combination with neuroimaging, to track neural correlates of PEs when predicting rewards for oneself (self) and for another individual (other-related PEs). Participants were required to indicate if the outcome was expected or unexpected, for both self and other conditions. The autistic group reported more probable events as unexpected and less probable events as expected, suggesting atypical learning of the underlying regularities of the learning environment, with this difference particularly pronounced when taking the perspective of another individual. In addition, other-related PE signals were observed in specific neural regions in the control, but not in the autistic group [187]. In contrast, using a probabilistic reward learning paradigm, Manning and colleagues reported no differences between autistic and non-autistic children when adapting learning rates to environmental volatility [125]. Finally, Robic et al. used a decision-making task to investigate the influence of social and non-social information on choice in an environment where volatility was manipulated. On each trial, a cue representing social (a video of an actor visually indicating an option) or non-social advice (an arrow indicating an option) was given, which varied in its validity. Autistic participants showed impaired performance relative to non-autistic participants during volatile,

but not stable phases, with a stronger behavioural deficit observed with regard to social cues [101].

Taken together, preliminary research suggests that social learning may be atypical in autistic individuals, particularly under volatile conditions. However, most paradigms investigating social learning employed overtly social stimuli, such as pictures or videos of faces [35],[101],[123],[126],[188]. Thus, introducing confounds, since both atypical processing of, and attention to, explicitly social stimuli have been observed in autistic populations [303]–[308]. Therefore, extant studies cannot determine whether differences in the use of social information in autistic individuals stem from differences in processing/ attending to social cues or differences in social learning *per se*. Considering the above, the current study employed a paradigm where social information was presented abstractly through cues, with no requirement for face or biological motion processing.

We examined computational learning parameters in social and individual learning simultaneously in autistic adults and a matched group of non-autistic control participants. An adapted version of a probabilistic social learning task (SLT) was employed, allowing indexing of learning in stable and volatile environments, separately for social and individual information [93],[95],[309]. Learning rates were modelled using an adapted Rescorla-Wagner (RW) learning model [24], which allowed social and individual learning rates to be estimated simultaneously, separately for volatile and stable phases. Adjustment to volatility was calculated as the increase in learning rate during volatile, versus stable, conditions. To supplement computational analyses, an analysis of win-stay lose-shift scores was conducted.

We hypothesised that, despite the lack of a requirement for face/biological motion processing in the current paradigm, autistic individuals would show deficits in social learning, relative to the non-autistic group. Thus, we predicted a group by information interaction for learning rates, with reduced learning from social information in the autistic group. This pattern of results would imply that autistic individuals show social-specific learning difficulties that cannot be explained by the complex/unpredictable nature of social stimuli or impaired face/biological motion processing. We further hypothesised that the autistic group would show atypical adaptation to volatility, indexed by the difference in learning rates in stable, versus volatile phases, and that this would be more pronounced with regard to social information, in line with previous research reporting atypical social learning under volatile

conditions. Specifically, we predicted a group by volatility and/or a group by information by volatility interaction.

3.2 Methods

3.2.1 Participants

Subjects (n = 73, aged 18 to 70 years, mean (SD) = 37.3 (13.1); 7 female) were recruited from a database held by the Autism Research Groups at City University London and Oxford University, and non-autistic control participants were recruited online. All subjects provided written, informed consent to participate. Autistic participants had previously been diagnosed by a UK National Health Service (NHS) or privately registered independent clinician, according to the DSM ([1]) or ICD-10 [278] criteria. All participants were reimbursed for their time (at a rate of £10 per hour) and travel expenses. ASD diagnosis was confirmed with administration of the Autism Diagnostic Observation Schedule, second edition ADOS-2; [279] by a trained researcher, using the current standard scores for a diagnosis of ASD. The study was approved by the City University London local ethics committee and was conducted in accordance with the Declaration of Helsinki. Autistic (ASD) and non-autistic control (CTRL) groups were matched on IQ and gender (Table 3.1).

3.2.2 General Procedure

Participants first provided written, informed consent; second, completed the Autism-Quotient (AQ) questionnaire [124] and the vocabulary and reasoning subscales of the WAIS-IV (if not already on file in the database), administered by a trained researcher (1 hour); finally, participants completed 120 trials of the social learning task (SLT) (approx. 35 minutes).

Social learning task

On each trial participants had to make a choice between a blue and a green shape, with correct choices rewarded (Fig. 3.1A). The probability of reward (individual reward) associated with each shape (blue or green) varied throughout the task, requiring participants to track the changing probabilities of either shape being correct. Participants were informed that the task followed phases, with sometimes the blue, and sometimes the green shape more

likely to result in reward, and that phases could change at any time. In addition to the individual reward history, a second source of information was available to participants: on each trial, a red frame surrounded one of the shapes. Participants were informed that this frame represented 'social' advice: the most popular choice made by a group of people, who had previously completed the task. The probability of the social advice also varied throughout. Participants received reward feedback in the form of an outcome indicator, a green or blue box in the middle of the screen, directly indicating which shape (blue/green) was correct. Participants could then infer from reward feedback whether the social information (the red frame) was correct or incorrect. Participants were randomly allocated to one of four different groups, with group membership determining the probabilistic schedule underpinning both reward outcomes (blue/green) and the veracity of the group advice (correct/incorrect), with differing stable and volatile phases for each (Fig. 3.1B). For stable phases, the probability of reward and "advice" was constant and during volatile phases, probabilities changed every 10-20 trials. The randomisation schedule for group 1 was the same as in the task developed by Behrens et al. [88]. For groups 2, 3, and 4, the schedules were inverted and counterbalanced versions of schedule 1 (see Appendix 2.1 for a full description of the behavioural task and randomisation schedules).

Figure 3.1 Behavioural task

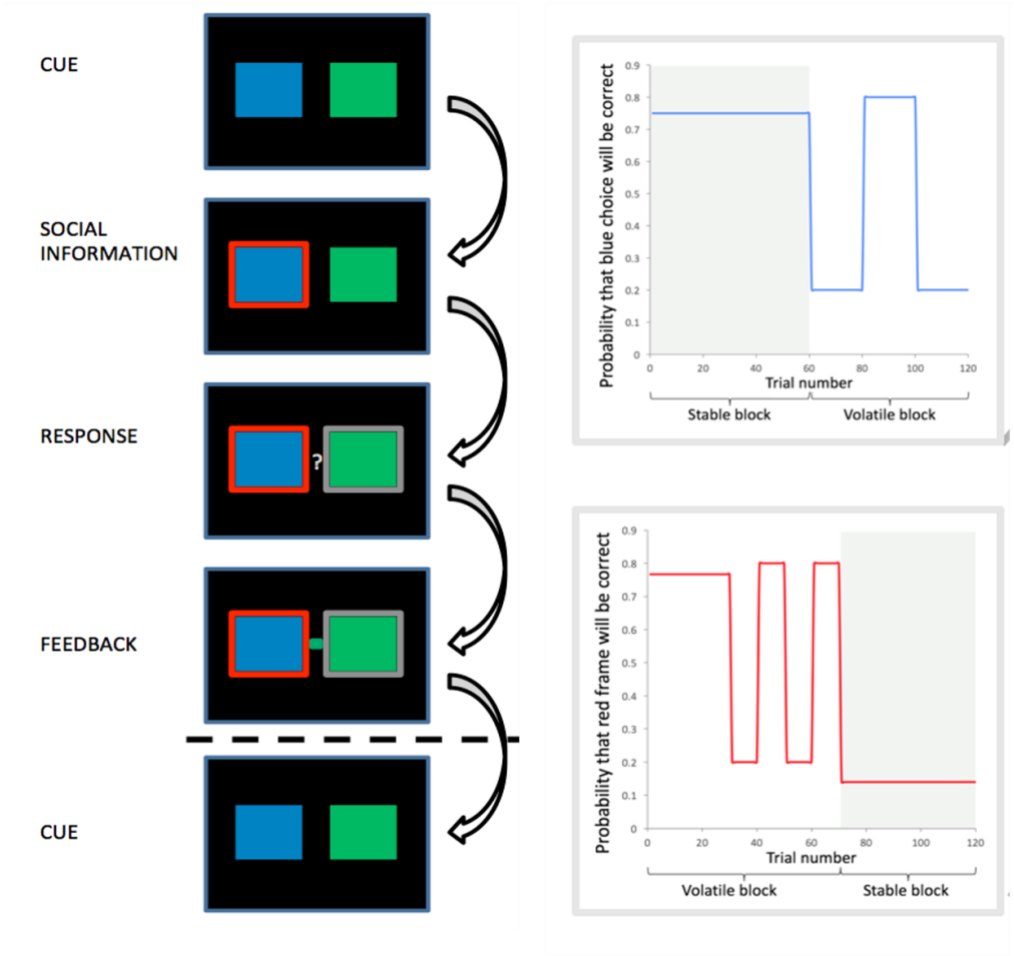


Figure 3.1. Behavioural task. A. Participants selected between a blue and a green box to gain points. On each trial, the blue and green boxes were presented first. After 1-4 seconds (s), one of the boxes was highlighted with a red frame, representing the social information. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a green or blue box in the middle of the screen. **B.** The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of the blue box/frame being correct (top) and the probability of the red (social) box/frame being correct (bottom).

Table 3-1 Demographic information

n = 64	CTRL group (n = 35)	ASD group (n = 29)			
	Mean (SD)	Mean (SD)	t (1,63)	X^2 (1, N = 64)	p
Gender (n males: n females)	31:4	26:3		0.019	0.890
Age	34.8 (12.9)	41.4 (13.2)	-2.031		0.047
2-subscale IQ	110.727 (13.588)	114.481 (13.704)	-1.061		0.293
Autism-Quotient (AQ)	17.382 (6.893)	36.276 (7.769)	-10.228		<0.001
BDI	6.333 (5.861)	15.733 (9.369)			
ADOS-2 total scores		9.760 (2.934)			

Note: IQ was assessed with the Wechsler Abbreviated Scale of Intelligence-2 (WASI-2). Depression was indexed with the Becks Depression Index (BDI). ADOS-2 refers to the Autism Diagnostic Observation Schedule, second edition. CTRL refers to the non-autistic control group, ASD refers to the autistic group. Due to missing data, BDI and ADOS could not be compared between groups. SD refers to standard deviation. IQ and gender did not significantly differ between the groups. Age and AQ differed significantly between groups.

3.2.3 Behavioural measures

Accuracy and Reaction Time

Accuracy was defined as the proportion of correct responses. Reaction time (RT) was calculated as time from stimulus presentation to response in milliseconds (ms).

Win-stay, lose-shift analysis

Win-stay, lose-shift scores were calculated separately for individual and social information. For individual learning, a win-stay was counted if the participant chose the correct answer on the previous trial and if their response on the current trial was the same (e.g., pick blue and win, and subsequently pick blue again). A trial was denoted as a lose-shift if the participant had lost on the previous trial and subsequently shifted their response. For the social information, a trial was counted as a win-stay if the participant correctly conformed to or disregarded the group choice and if they were correct, repeated this decision on the next trial (e.g., go with the group and win, and subsequently chose to go with the group again). A lose-shift trail was counted if the participant wrongly conformed to or disregarded the group's choice and then made the opposite decision on the next trial. Win-stay and lose-shift trials were summed and divided by the total number of win/lose trials, separately for volatile and stable trials.

Data pre-processing

Participant data was excluded based on the following: accuracy under chance level (< 50%), choose the same side or colour on over 80% trials or conformed to the group answer on over 80% of trials, incomplete datasets (less than 120 trials completed), resulting in a final sample of n = 64. As in the previous chapter, we supplemented all analyses with Bayesian statistical analyses (Appendix 2.2).

Computational model

An adapted Rescorla-Wagner (RW) learning model [24] was fitted to participants' choice data, providing estimates, for each participant, of $\alpha, \beta,$ and ζ. The learning rate (α) controls the weighting of prediction errors on each trial. A high α favours recent over (outdated) historical outcomes, while a low α suggests a more equal weighting of recent and more distant trials. α was estimated separately for volatile and stable phases for both individual and social learning, resulting in four α estimates: $\alpha_{\text{individual\_volatile}}$, $\alpha_{\text{individual\_stable}}$, $\alpha_{\text{social\_volatile}}$, $\alpha_{\text{social\_stable}}$. β captures

the extent to which learned probabilities determine choice, with a larger β meaning that choices are more deterministic with regard to the learned probabilities. ζ represents the relative weighting of primary and secondary sources of information, with higher values indicating a bias towards the over-weighting of social relative to individual information (see Appendix 2.3 for full model details and Appendix 2.4 for model comparison and validation methods).

Estimated α values were compared to optimal α estimates. An optimal learner model, with the same architecture and priors as the model employed in the current task, was fit to 100 synthetic datasets, resulting in average optimal learning rates: $\alpha_{\text{optimal\_primary\_stable}} = 0.16$, $\alpha_{\text{optimal\_primary\_volatile}} = 0.21$, $\alpha_{\text{optimal\_secondary\_stable}} = 0.17$, $\alpha_{\text{optimal\_secondary\_volatile}} = 0.19$. Scores representing the difference between α estimates and optimal α scores were calculated (α_{diff}: $\alpha - \alpha_{\text{optimal}}$).

3.3 Results

Accuracy and reaction time

No differences were observed in accuracy between ASD (mean (standard error) $\bar{x}$ ($\sigma_{\bar{x}}$) = 0.607 (0.009)), and CTRL groups ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.617 (0.009); F (1,56) = 0.676 p = 0.414, η_p^2 = 0.012). Furthermore, accuracy was significantly greater than zero for both the autistic (t(28) = 69.770, p <0.001, d = 12.956) and the non-autistic group (t(34) = 77.755, p <0.001, d = 13.143)), suggesting both groups could successfully perform the task. Additionally, mean reaction time (RT) did not significantly vary between ASD ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.455 (0.130) and CTRL groups ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.240 (0.130); F (1,62) = 1.605, p = 0.210, η_p^2 = 0.025) (see Appendix 2.5 for extended statistical analyses).

Analysis of computational modelling parameters

Mathematical models of learning suggest that beliefs are updated in proportion to prediction errors – the difference between predicted and actual outcomes – which are modulated by a learning rate (α). Previous work predicts either a difference between ASD and CTRL groups in the change in learning rate between volatile and stable environments, and/or a difference in

learning rate regarding social relative to individual sources of information during learning. To test these predictions, we compared estimated α values from a RW model of learning, between individual and social information sources and for volatile and stable learning environments across groups. A RM-ANOVA was conducted on (square-root transformed) α values (see Appendix 2.5 for untransformed values), with factors volatility (stable, volatile), information source (individual, social), and group (ASD, CTRL). A significant main effect of information source was observed ($F(1, 62) = 67.111$, $p < 0.001$, $\eta_p^2 = 0.520$) with $\alpha_{individual}$ ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.572 (0.018)) significantly higher than α_{social} ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.349 (0.018)). No other main or interaction effects were observed. Importantly, and in contrast to our main predictions, no main or interaction effect(s) were observed involving the factor group (all $p > 0.05$). Specifically, we did not observe an interaction between group and information source ($F(1, 62) = 0.798$, $p = 0.375$, $\eta_p^2 = 0.013$, $BF_{excl} = 2.549$) or group and volatility ($F(1, 62) = 0.867$, $p = 0.356$, $\eta_p^2 = 0.014$, $BF_{excl} = 3.989$). Thus, higher learning rates were observed for individual information, with no differences between ASD and CTRL groups.

Separate univariate ANOVAs were carried out to compare decision parameters β, controlling the extent to which learned probabilities determine behaviour, and ζ, controlling the weighting of social relative to individual sources of information, across groups. No effect of group was observed on β values ($F(1, 62) = 0.133$, $p = 0.716$, $\eta_p^2 = 0.002$, $BF_{excl} = 3.411$) or ζ values ($F(1, 62) = 1.302$, $p = 0.258$, $\eta_p^2 = 0.021$, $BF_{excl} = 2.252$). Results suggest that the relative weighting of social versus individual information did not differ between groups.

Optimal learning rates

We predicted differences between the ASD and CTRL groups in adjusting learning rate to current volatility, i.e., atypical adjustment of learning rate within the ASD group. However, analysis of learning rates did not find evidence to support our hypotheses, with no group differences observed. Furthermore, although our winning perceptual model included separate learning rates for volatile and stable phases, no main effect of volatility was found on α values overall ($F(1, 62) = 0.293$, $p = 0.590$, $\eta_p^2 = 0.005$), suggesting that participants, in both groups, were not modulating learning rate according to environmental volatility. Thus, raising the possibility that, although the chosen computational model was the (relative) best fitting model (Appendix 2.4), this model was not accurately describing participant behaviour. To

investigate this further, we compared our estimated α values with optimal α estimates, allowing measurement of how estimated learning rates differed from optimal values. Scores representing the difference between (untransformed) α estimates and optimal α scores (α_{diff}: $\alpha - \alpha_{\text{optimal}}$) were submitted to one-sample t-tests, with difference scores for $\alpha_{\text{individual\_volatile}}$ ($t(63) = 6.696$, $p < 0.001$), $\alpha_{\text{individual\_stable}}$ ($t(63) = 7.163$, $p < 0.001$) and $\alpha_{\text{social\_volatile}}$ ($t(63) = -4.323$, $p < 0.001$) significantly differing from optimal estimates (Fig. 3.2).

Figure 3.2 Learning rate estimates minus optimal learning rates

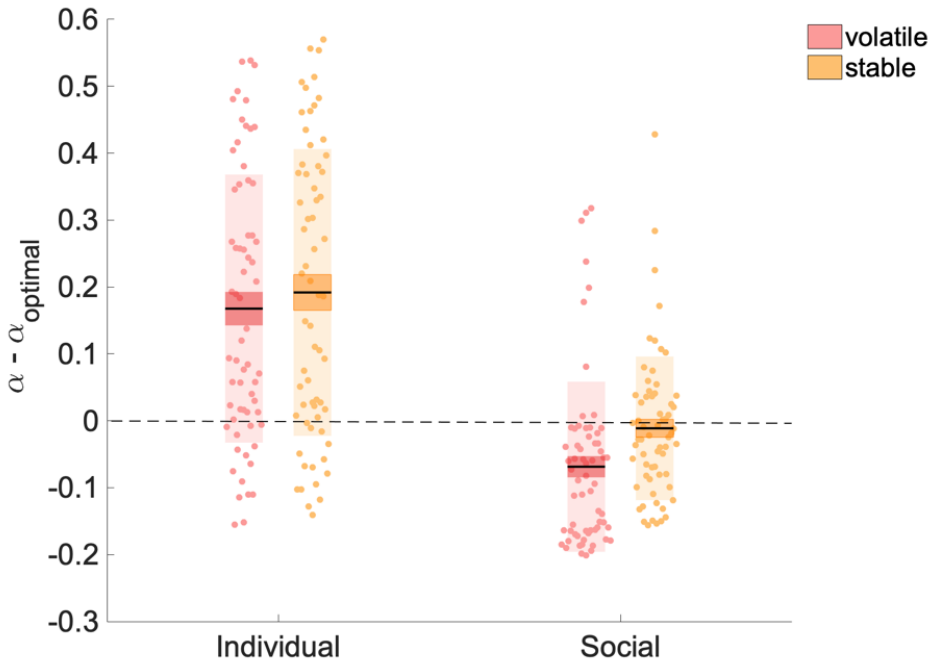


Figure 3.2. Learning rate estimates minus optimal learning rates. The dashed line indicates optimal α values. Data points indicate $\alpha - \alpha_{\text{optimal}}$ values for individual participants (n = 64) across both groups, boxes = standard error of the mean, shaded region = standard deviation.

Analysis of win-stay, lose-shift behaviour

Learning rates significantly differed from optimal estimates, raising the possibility that participants were not using a reinforcement learning (RL) strategy. We subsequently performed a behavioural analysis of learning, comparing the use of win-stay, lose-shift

(WSLS) behaviour between groups. Win-stay (WS) behaviour measures how often a participant stays with the previously rewarded choice, while lose-shift (LS) behaviour involves shifting to the alternative option after a loss. In this task, the utility of a WSLS strategy varies according to environmental volatility. For example, increased WSLS is theoretically adaptive in volatile phases but maladaptive in stable phases where optimal performance depends on ignoring misleading probabilistic feedback.

WSLS scores were calculated separately for individual and social, and volatile and stable phases and submitted to a RM-ANOVA with within-subjects factors volatility (stable, volatile), information source (individual, social) and index (win-stay, lose-shift), and between-subjects factor group (ASD, CTRL). WSLS scores were significantly correlated with learning rates for individual and social information for both volatile (individual: $r = 0.769$, $p < 0.001$, social: $r = 0.558$, $p < 0.001$) and stable phases (individual: $r = 0.818$, $p < 0.001$, social: $r = 0.323$, $p = 0.009$). A main effect of information source was observed ($F(1,62) = 55.250$, $p < 0.001$, $\eta_p^2 = 0.471$), with higher WSLS behaviour for individual ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.716 (0.012)) versus social information ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.571 (0.012)), as well as a main effect of index ($F(1,62) = 110.273$, $p < 0.001$, $\eta_p^2 = 0.640$), with significantly more win-staying (WS) ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.721 (0.010)) versus lose-shifting (LS) ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.571 (0.010)). As in our analysis of learning rates, no main effect of volatility was observed ($p > 0.05$). However, a significant group by index by volatility interaction was observed ($F(1,62) = 4.936$, $p = 0.030$, $\eta_p^2 = 0.074$), although analysis of effects using a Bayesian RM-ANOVA provided weak evidence *against* the presence of this interaction effect ($BF_{excl} = 1.386$). To unpack this interaction, separate RM-ANOVAs were conducted for WS and LS scores, collapsed across individual and social learning. For LS scores, no main/interaction effect(s) were observed (all $p > 0.05$). For WS scores, however, a significant volatility by group interaction was observed ($F(1, 62) = 4.327$, $p = 0.042$, $\eta_p^2 = 0.065$) (Fig. 3.3). Although holm-corrected post hoc comparisons did not reach significance (all $p_{holm} > 0.05$), the pattern of results suggested that the ASD group were showing more WS behaviour in volatile ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.734 (0.015)) compared with stable phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.709 (0.015); $t(64) = 1.098$, $p_{holm} = 0.691$, uncorrected $p = 0.114$). In contrast, the CTRL group were showing more WS in stable ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.741 (0.016)) compared with volatile phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.703 (0.014); $t(64) = 1.884$, $p_{holm} = 0.386$, uncorrected $p = 0.127$).

In sum, win-stay behaviour varied in response to volatility between groups, providing tentative support for altered response to volatility in the autistic group. There were no differences in response to social/individual information.

Figure 3.3 Win-stay behaviour as a function of volatility

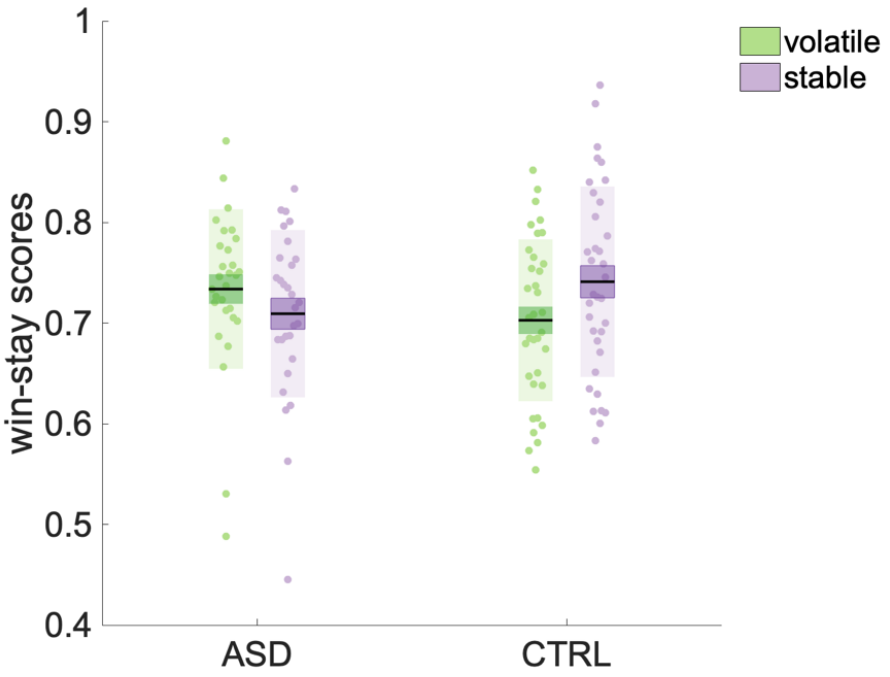


Figure 3.3. Win-stay (WS) scores (collapsed across individual and social learning). WS varied as a function of volatility between ASD and CTRL groups. Data points indicate mean WS values for individual participants (n = 64) across groups, boxes = standard error of the mean, shaded region = standard deviation.

3.4 Discussion

The current study investigated learning from social versus individual information and adaptation to volatility in adults with autism. In line with predictive processing and Bayesian accounts of autism, we predicted that learning from social information would be atypical in the autistic group. In addition, we predicted that the autistic group would show reduced adaptation to volatility during learning, indexed by atypical adjustment of learning rates. In contrast to our predictions, results showed no evidence for group differences with respect to social, or indeed individual, learning. Furthermore, no evidence for atypical adaptation of

learning rate in response to changes in environmental volatility was observed in the autistic, compared with the non-autistic, group. Bayesian analyses provided moderate evidence in support of no differences between groups as a function of learning type or volatility. When we further explored the data, group differences were observed for win-stay scores: in contrast with the non-autistic group, who showed more win-staying in stable, as compared to volatile phases, autistic participants showed the opposite pattern of behaviour. Results suggest subtle differences in response to volatility in autism; in a volatile environment, while non-autistic participants became more flexible with regard to repeating a previously rewarded response or shifting responses after a loss, autistic participants did not, but rather stayed with the previously rewarded response. Overall, results do not support predictions from accounts proposing altered prediction-based learning in social contexts in autism [13],[50], nor atypical adjustment of learning rates to volatility [77],[78]. However, analyses of win-stay scores add tentative support for altered choice behaviour after reward in volatile environments in autistic individuals.

There were no differences in learning rates between the groups with regard to adjustment of learning to environmental volatility. However, and in contrast to previous research showing higher learning rates in volatile, relative to stable, phases in the general population [98],[230], no main effect of volatility was observed on learning rates when collapsed across groups. A potential explanation for the lack of a main effect of volatility is that our paradigm, which demands learning from two dissociable, and sometimes opposing sources of information, has high learning demands, particularly during volatile phases. Thus, raising the possibility that participants were disengaging during volatile phases. However, learning rates were significantly higher than zero in both stable and volatile phases, and accuracy did not decrease significantly during volatile phases. Furthermore, model comparison showed that the model that best fit the data included separate learning rates for learning in volatile and stable phases. Thus, we argue that participants were engaged with the task in volatile phases and were responding to the volatility manipulation.

A further possibility is that, due to high learning demands, participants were not integrating the value of trials over the course of the experiment, in line with an RL model of learning, but rather, using a WSLS strategy, taking only the outcome on the previous trial into account. WSLS and RL strategies are closely related, with WSLS scores often used as a proxy measure of learning rate [93], although they can also be dissociated [310]. For example, continually following a WSLS strategy would correspond to a learning rate of 1 within an RL

framework, meaning that only the most recent trial (*t-1*) would be taken into account on the current trial *t*. As WSLS scores and learning rates were significantly correlated in our paradigm, we examined WSLS behaviour in this task.

An exploratory analysis of WSLS scores revealed a group difference with regard to win-staying in volatile trials; no group differences were observed for trials where a loss was experienced. That is, while non-autistic participants showed a greater tendency to shift response in volatile phases, autistic participants showed the opposite pattern of behaviour. Thus, it is possible that the WSLS analysis was detecting subtle differences in response to trials that resulted in a win or loss, which could not be indexed using a RW learning model. Indeed, previous work has found differences in behaviour in volatile environments, unrelated to differences in learning rates. Using a reward learning task with non-autistic individuals, Goris et al. reported that, while impaired performance in volatile environments was correlated with a higher level of autistic traits, no impairment in adjustment of learning rate to volatility was found. Instead, impaired performance was observed after a reversal in cue-outcome contingencies [99]. This raises the question of whether atypical performance in autistic individuals in volatile environments can be ascribed to impairment in adaptation to volatility, or instead, to impaired performance post-reversal, as has been previously reported [44],[45],[311]. To unpack this further, a paradigm and a model of learning that can discriminate between initial and post-reversal learned reward contingencies could be used in future research. However, it is important to note that, although a significant interaction effect on win-stay scores was observed, post-hoc tests did not reach significance and Bayesian evidence provided weak evidence *against* the inclusion of this interaction. Furthermore, this effect does not translate into group differences in accuracy, in contrast with observations of decreased performance in volatile phases in previous work [99],[101]. Therefore, this exploratory analysis of WSLS scores should be treated with caution.

Predictive processing accounts of autism, in theory, describe common mechanisms underpinning autistic cognitive processing. We, therefore, expected to observe atypical learning in the autistic group, particularly in volatile phases, in accordance with previous research [101]. However, no evidence for differences in social learning between groups was observed. Moreover, Bayesian analyses provided moderate evidence to support the conclusion that the groups were comparable with respect to learning from social versus individual information. Finally, no differences were found in the relative weighting of social and individual information, as indexed by the decision-weighting parameter ζ. The current

results raise the possibility that predictive atypicalities are not present during social learning and stand in contrast to previous findings [59],[123],[188].

A potential explanation for the contrasting results reported here is the lack of requirement for face processing or biological motion processing in our paradigm. Predictive processes are proposed to be atypical in the social domain, in part due to the complex and unpredictable nature of social cues [13]. In contrast with the social information in the research mentioned above, the majority of which used images or videos of faces (for example, representing the social cue through eye gaze direction or pointing [101],[188]), the social information in our paradigm was equally as predictable as the individual information, and represented by a simple cue (a red frame). The current results demonstrate that autistic adults engaged in social learning to the same extent as non-autistic adults, suggesting that differences in social learning observed in previous work were underpinned by differences in face/biological motion processing, rather than in predictive learning specifically.

Our results are relevant both for refining predictive accounts of autism and for the development of better tools to improve social learning in autistic individuals. For example, it is important to determine whether aids or therapies to improve social cognitive processes should be social-specific or rather based on domain-general associative learning principles. Our findings demonstrate that autistic individuals do not show a specific impairment in social relative to individual learning, compared with non-autistic individuals; consistent with a domain-general view of social learning. However, across both groups, learning rates were higher for individual compared with social information, indicating that mechanisms underlying social and individual learning might be dissociable. Social learning has been reported to follow the same computational and neural principles as individual reward learning, relying on the same prediction-error mediated processes [230] and associated with the same neural regions [221],[225],[227]. However, other research supports the existence of social-specific learning mechanisms [95],[187],[257]. Thus, future work should unpack what makes social learning specifically 'social', at a neurochemical level.

To conclude, using a combination of frequentist and Bayesian statistics, no evidence was found to support differences in social or individual learning between autistic and non-autistic adults in a probabilistic social learning task. Furthermore, no group differences were found in adaption of learning rate to volatility. We did, however, find a small and unpredicted group difference in choice behaviour after trials which resulted in reward in volatile phases. Taken

together, results fail to provide evidence in support of predictions from Bayesian accounts of autism in the context of social learning and adjustment to volatility and have the potential to help in refining predictive coding accounts of autism.

Chapter 4: The influence of genetic variation in monoamine transporters on social and individual learning

In the previous chapter, although no differences in learning rates between autistic and non-autistic adults were observed, dissociations between social and individual learning were found. Thus, highlighting the importance of finding out whether social learning is underpinned by social-specific neurochemical mechanisms. In this chapter, a large-scale behavioural genetics approach is employed to investigate whether there is evidence to support the existence of dissociable genetic contributions to social and individual learning. Building on their previously documented roles in learning (see Chapter 1), I here focus on genes related to the monoamine neurotransmitters dopamine and serotonin, both implicated in learning from one's own experience (individual learning) and in social learning (learning from others).

Supplementary materials for this chapter are in Appendix 3.

4.1 Introduction

The existence in the human brain of neural and/or neurochemical pathways that are specialised for learning from social information and from individual experience respectively is the topic of much debate [186],[189],[201],[312], and of great importance to theories of cultural evolution [313]. If humans learn from others using the same, domain-general mechanisms as when learning from any other source of information in the environment [201], social and individual learning should not be dissociable at a neural/neurochemical level. In contrast, if social and individual learning can be dissociated, this implies the presence of specific mechanisms for social learning. Building on their previously documented roles in both individual and social learning, this chapter focuses on genes related to the monoamine neurotransmitters dopamine and serotonin.

Dopamine (DA) is perhaps best known for its role in individual learning [211],[217]. Indeed, a number of studies have shown that individual learning covaries with genetic polymorphisms that affect the dopamine system [236],[314],[315]. Den Ouden and colleagues, for example, showed that genetic variation in the dopamine transporter (DAT1) polymorphism altered both perseveration and the influence of past choices on behaviour in the reversal phase of an associative learning task [236]. Thus, supporting a role for dopamine in learning and demonstrating that naturally occurring (e.g., genetic) variation in the dopamine system can account for variation in behavioural indices of learning. Furthermore, although dopamine has predominately been investigated in the context of individual learning, recent work in both animals and humans suggests a role for dopaminergic signalling in social learning [222],[223],[225],[226],[229],[316],[317]. For example, Diaconescu et al. found that the magnitude of striatal social learning-related prediction error signals co-varied with dopamine-related genotype - specifically polymorphisms of the Catechol-O-Methyltransferase (COMT) gene - in a sample of 82 individuals [222]. Thus, preliminary work suggests that, in addition to its role in individual learning, dopamine may also be implicated in social learning.

Serotonin (5-HT) signalling is strongly associated with social behaviour, with a wide body of research highlighting the importance of the serotonin system in social cognition [240],[241]. With respect to social learning more specifically, Crişan et al [255] found a significant effect of genetic variation in serotonergic signalling on learning from observing the fear responses of other people during an aversive conditioning task. Specifically, variation in the serotonin reuptake transporter gene (SERT) covaried with participants' autonomic skin conductance

response, both when observing another individual receiving aversive feedback and when presented with the conditioned stimuli themselves.

Importantly, in addition to its role in social behaviour, serotonin has been implicated in individual learning [232]–[237]. Den Ouden and colleagues [236], for example, showed that the extent to which behaviour was adapted after punishment - indexed by the degree to which participants shifted response after receiving negative feedback - was predicted by SERT genotype. More recently, a role for serotonergic signalling has been proposed in signalling prediction errors [238],[239]. Consequently, preliminary work implicates serotonin in both individual and social learning.

The literature, to date, implicates dopamine and serotonin in both individual and social learning. Based on the importance of serotonin in social learning, and dopamine in individual learning, social learning might predominately rely upon serotonergic mechanisms and individual learning predominately on dopaminergic mechanisms. For example, if social learning relies more heavily upon serotonergic mechanisms (compared to individual learning), this raises the possibility of a social-specific neurochemical learning pathway that is especially serotonin-dependent. However, extant studies have not been able to test this because they have not adopted a design enabling the mapping of variation in monoamine systems onto variation in both social and individual learning within the same individuals. Here we leverage work showing that genetic variation in genes involved in dopamine and serotonin signalling can result in naturally occurring variation in monoamine availability [318]. Building upon this, the current study asked whether social and individual learning share common dopaminergic and serotonergic mechanisms. Specifically, whether genetic variation in dopamine- and/or serotonin-related single nucleotide polymorphisms (SNPs) has different effects on social versus individual learning.

Variation in social and individual learning was indexed using an online version of a widely used social learning task (SLT) [93],[95],[309]. Participants (n = 803) were required to choose between a blue or green shape to gain points, receiving reward feedback in the form of a blue or green indicator, which directly informed them about the utility of their choice. In addition, participants saw an orange frame, representing social information, which surrounded one of the boxes. Participants were informed that the orange frame represented social information (the most popular choice made by previous participants) which could vary between being predominantly correct, and predominantly incorrect. Successful performance thus required

participants to engage in individual learning (about the value associated with the blue and green shapes) and social learning (about the value of the social information). The majority of the commonly used formalisations (and conceptualisations) of human decision-making assume that there is a prediction error-driven learning process by which individuals learn the value (either by individual or social means) of stimuli/actions, and a subsequent decision-making process by which individuals evaluate the extent to which their behaviour will be driven by these learned values. These processes have been referred to, respectively, as the critic and actor, the perceptual and response model, and/or the learning model and decision rule [24],[25],[319]. Here we used a computational model with free parameters, which can be estimated participant-wise, in both the learning model and decision rule. With respect to the estimated parameters from the learning model, we predicted an interaction such that serotonin-related genetic variation would account for significantly more variation in social as opposed to individual learning-related estimates, whereas dopamine-related genetic variation would account for significantly more variation in individual as opposed to social learning-related estimates. Theories of cultural learning postulate social-specific *learning* mechanisms, however, to the best of our knowledge these theories do not predict social-specific biases in decision rules. Thus, as far as we are aware there is no *a priori* reason to predict that parameter estimates relating to the decision-rule would covary with dopamine- or serotonin-related genes.

Participants were genotyped for variation in both dopamine- (DAT1, COMT) and serotonin-related (SERT) genes. Both frequentist and Bayesian statistical models were used to test whether variation in dopamine- and/or serotonin-related genes predicted variation in indices of social and/or individual learning.

4.2 Materials and Methods

4.2.1 Participants

Participants were part of the Nijmegen Brain Imaging Genetics (BIG) project (Donders Institute, Nijmegen, The Netherlands), comprising imaging, genetic and cognitive data from healthy volunteer subjects (https://www.ru.nl/donders/research/research-facilities-projects/cognomics/big-project/). Participants were recruited from among the BIG cohort and invited to complete an online learning task. All participants had given written informed

consent to participate in the BIG study prior to participation and the study was approved by the local medical ethics committee. 803 healthy participants for whom genetic data was available, aged 18-80 years (57.9% female, mean age = 28.17 ± 10.07 years) completed the online behavioural task (described below).

Table 4-1 Demographic information

DAT		***9/9***	***9/10***	***10/10***
	n	32	208	330
	Gender (male: female)	16:16	87:121	147:183
	χ^2 (570) p		0.902 0.637	
	Age (mean (SD))	30.4 (13.2)	26.3 (6.8)	28.0 (9.8)
	F(2,567) p		3.793 0.023*	
SERT		***s/s***	***s/l***	***l/l***
	n	185	262	123
	Gender (male: female)	86:99	110:152	54:69
	χ^2 (570) p		0.893 0.640	
	Age (mean (SD))	27.6 (9.8)	26.8 (7.6)	28.8 (10.7)
	F(2,567) p		1.975 0.140	
COMT		***v/v***	***v/m***	***m/m***
	n	166	270	134
	Gender (male: female)	80:86	118:152	52:82
	χ^2 (570) p		2.658 0.265	
	Age (mean (SD))	27.6 (9.8)	26.8 (7.6)	28.8 (10.7)
	F(2,567) p		0.430 0.651	

Note: Genotype frequencies (number), gender (number) and age (mean and standard deviation (SD)) are reported. Statistical tests compare the genotype groups on demographic variables age and gender. Gender did not significantly differ as a function of DAT, SERT or COMT genotype. Age significantly differed as a function of DAT genotype.

4.2.2 Genetic Analysis

All genetic analyses were performed at the Department of Human Genetics, Radboud University Nijmegen Medical Centre, with details of DNA collection, extraction and genotyping described in full in Appendix 3.1. Three different genes were investigated, DAT1, SERT1 and COMT. For DAT1 (SLC6A3), two alleles were analysed: 9R and 10R. Genotypes included 9:9, 9:10 and 10:10. S/S, S/L and L/L genotypes were included for SERT (SLC6A4). For the COMT gene, val/val, val/met and met/met genotypes were included. All genotypes were in Hardy Weinberg equilibrium (all p-values > 0.1).

Dopamine-related genetic variation

The dopamine transporter (DAT) is a membrane protein that regulates dopamine transmission by reuptake from the presynaptic cleft. A 40-base pair variable number of tandem repeats (VNTR) polymorphism located in the 3′-untranslated region (3′UTR) of the DAT gene (DAT1/SLC6A3) produces two common alleles with 9- and 10-repeats (9R and 10R) [320]. Although the polymorphism causes variation in DAT protein expression, results are mixed as to the direction of the effects on DAT expression and DA signalling [320]–[322]. Consequently, our predictions were only that indices of learning would covary with DAT genotype; we had no directional predictions regarding the effect of DAT genotype. We did, however, predict an interaction between DAT genotype and learning source, such that DAT genotype would account for significantly more variation in individual as opposed to social learning-related estimates.

Variation in the gene encoding Catechol-O-Methyltransferase (COMT) was also investigated. COMT plays a key role in synaptic dopamine removal, particularly in the cortex where DAT expression is low [323],[324]. A single nucleotide polymorphism of the COMT gene results in a single amino acid substitution of methionine (met) for valine (val), with the val allele associated with higher activity [325] and the met allele associated with lower activity and higher prefrontal dopamine concentrations [314]. Variation in the gene encoding COMT has been linked to the ability to flexibly adapt decisions during a probabilistic reversal learning task [326] and found to influence prediction error responses in the midbrain during social learning tasks [222]. We predicted an interaction between COMT genotype and learning source, such that COMT genotype would account for significantly more variation in individual as opposed to social learning-related estimates.

Serotonin-related genetic variation

A key regulator of serotonergic transmission is the high-affinity 5-HT reuptake transporter, SERT/5-HTT, encoded by a single gene, SLC6A4. An insertion or deletion in the promoter region of the gene, the SLC6A4-linked polymorphic region (5-HTTLPR), results in a short (S) and long (L) allele respectively, resulting in variation in both transcriptional and translational levels [327],[328]. The s-allele has been associated with a decrease in transcriptional efficacy, which putatively results in less 5-HT reuptake and increased extracellular 5-HT levels relative to the long allele [329], as well as reduced 5-HT binding [322], and a greater sensitivity to social cues [240],[330]. We predicted an interaction between SERT genotype and learning source, such that SERT variation would account for significantly more variation in social as opposed to individual learning-related estimates, and that s-allele carriers would show greater sensitivity to social information compared with l-allele carriers.

4.2.3 Behavioural task

Participants completed an online, modified version of the task employed in Chapter 3. Although the task structure was the same, the task was visually different, employing different shapes for the blue and green choices, and an orange frame for the social information. Furthermore, choices were made with a mouse, rather than with button presses, and to progress to the next trial, the mouse had to be re-centred to allow the next trial to start. All other details were the same: on each trial, participants had to make a choice between a blue and a green shape, with correct choices rewarded (Fig. 4.1). In addition to the individual reward history, a second source of information was available to participants: on each trial, an orange frame surrounded one of the shapes, representing the social information (see Appendix 3 for task instructions). Both the probability of reward and the veracity of the social information switched between being correct and incorrect, with stable and volatile phases for each.

Figure 4.1 Behavioural task

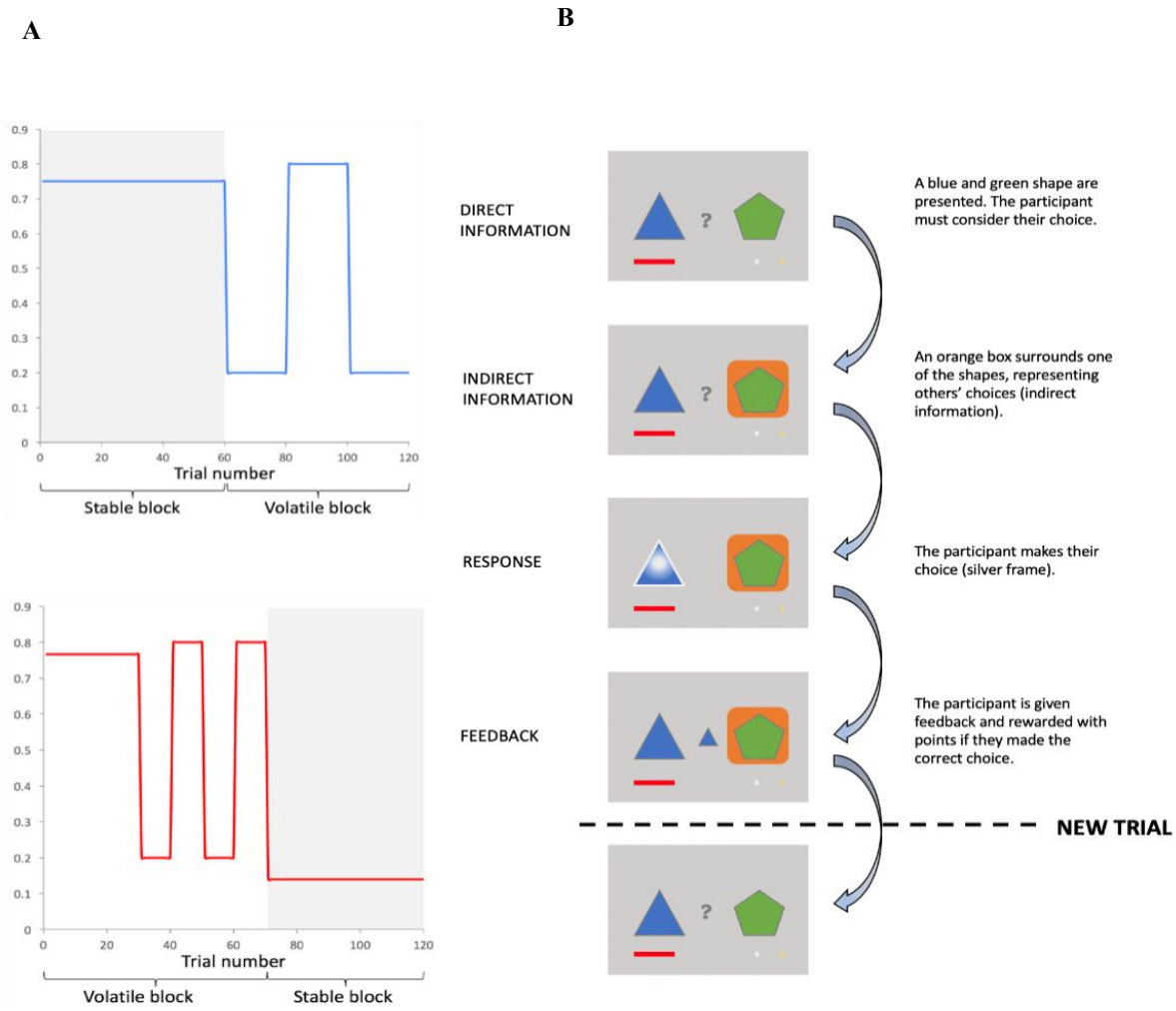


Figure. 4.1.A. Stable and volatile phases. The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of blue being correct (individual information) and the probability of the orange box indicating the correct answer (social information). **B. Task structure.** Participants selected between a blue and a green shape to gain points. One of the shapes was highlighted with an orange box. Participants were instructed that this box represented the most popular choice made by a group of four participants who had completed the task previously. After participants selected their response, a silver frame surrounded their choice. After a 0.5–2 s interval, participants received feedback in the form of a green or blue shape in the middle of the screen. If participants had chosen the correct answer, a red reward bar at the bottom of the screen progressed toward the silver and gold goals.

4.2.4 Data analysis

Statistical analyses were conducted using MATLAB R2017b (The MathWorks, Natick, MA) and JASP (JASP Team (2020). JASP (Version 0.14) [Computer software]). We used the standard $p < .05$ criteria for determining if significant effects were observed, with a Holm

correction applied for multiple comparisons, to control for type I family-wise errors. All analyses were repeated with and without the inclusion of age as a covariate. Where appropriate, data were transformed to meet assumptions of normality for parametric testing. The effects of different genotypes were tested separately for DAT, COMT and SERT, as small samples for some genotypes resulted in empty cells in the design when analysed simultaneously. The study had a mixed design with between-subjects factors COMT (A/A; A/G; G/G), DAT (9:9; 9:10; 10:10) and SERT (S/S; S/L; L/L) genotypes and within-subjects factors information type (individual, social information) and volatility (volatile, stable).

Exclusions

Only participants with the three most common genotypes for each gene were included (n = 693). Further to this, participant data was excluded based on the following measures of performance on the behavioural task: accuracy under chance level (< 50%), choose the same side or colour on over 80% of trials or conformed to the group answer on over 80% of trials, incomplete datasets (less than 120 trials completed), mean reaction time (RT) greater than three standard deviations (SD) from the mean RT, resulting in a final sample of n = 570 (56.1% female, mean age = 27.50 ± 9.1 years, see Table 4.1 for demographic details of the final sample).

4.2.5 Computational modelling framework

We modelled participant response using a Rescorla-Wagner learning model [24], consisting of a perceptual model and an action selector. The model relies on the assumption that updates to choice behaviour are based on prediction errors, i.e., the difference between an expected and the actual outcome. Participants were assumed to update their beliefs about reward outcomes, or the state of the environment based on sensory feedback in the form of reward, and to use this feedback to make decisions about the next action (response model). The Rescorla-Wagner predictors used in our learning models consisted of a modified version of a simple learning model (see Appendix 3.2 for full model details). Parameters were fitted separately for each participant's choice data and provided estimates of $\alpha, \beta,$ and ζ. Learning rate (α) was estimated for each participant, separately for individual and social information and volatile and stable phases, resulting in four estimated learning rates per participant. The free parameter β captured the extent to which learned probabilities determined participant choice

behaviour, with a larger β when more deterministic choices were observed, with regard to the learned probabilities. ζ represented the relative weighting of individual and social sources of information, with higher values indicating a bias towards the over-weighting of social relative to individual information.

4.3 Results

We employed three separate repeated-measures analysis of variance (RM-ANOVA) with fixed factors information source (individual, social), environmental volatility (volatile, stable), and genotype (DAT: 9/9, 9/10, 10/10, or SERT: S/S, S/L, L/L, or COMT: val/val, val/met, met/met) on dependent variables α, β, and ζ. For α we predicted an interaction between SERT and information source such that serotonin-related genetic variation would account for significantly more variation in social as opposed to individual learning-related estimates. Furthermore, we predicted that dopamine-related genetic variation (DAT or COMT) would account for significantly more variation in individual as opposed to social learning-related estimates. We conducted separate ANOVAs for DAT, SERT and COMT. Including randomisation schedule as a factor in all analyses did not change the pattern of results.

Effects of serotonin-related genetic variation on learning model-related parameters

The RM-ANOVA with α as the DV revealed a significant main effect of information type ($F(1,567) = 256.521$, $p < 0.001$, $\eta^2p = 0.311$, $BF_{incl} > 100$), with higher (sqrt-transformed) $\alpha_{individual}$ (mean (standard error) ($\bar{x}(\sigma_{\bar{x}})$ = 0.507 (0.006)) compared to α_{social} ($\bar{x}(\sigma_{\bar{x}})$ = 0.374 (0.006)). A significant volatility by SERT interaction was observed ($F(2,567) = 4.535$, $p = 0.011$, $\eta^2p = 0.016$; Fig. 4.2), with post hoc tests revealing a significantly lower α for volatile ($\alpha_{volatile}$: $\bar{x}(\sigma_{\bar{x}})$ = 0.425 (0.008)) compared to stable phases (α_{stable}: $\bar{x}(\sigma_{\bar{x}})$ = 0.452 (0.008)) for the S/L genotype ($t(567) = -3.437$, $p_{holm} = 0.009$, $BF_{10} = 38.950$). This difference was not observed for the S/S ($t(567) = 0.948$, $p_{holm} = 1.000$, $BF_{10} = 0.139$) or L/L genotypes ($t(567) = -0.309$, $p_{holm} = 1.000$, $BF_{10} = 0.111$), with Bayesian analysis providing moderate to strong evidence to support the null hypothesis of no difference between $\alpha_{volatile}$ and α_{stable} for these groups. Indeed, an ANOVA conducted on difference scores ($\alpha_{volatile}$ –

α_{stable}) revealed that the signed difference between $\alpha_{volatile}$ and α_{stable} was significantly smaller (i.e. more negative) for S/L carriers ($\bar{x}(\sigma_{\bar{x}})$ = -0.028 (0.008)) compared with S/S carriers ($\bar{x}(\sigma_{\bar{x}})$ = 0.006 (0.008); t (567) = -2.560, p_{holm} = 0.032) and (non-significantly) lower than for L/L carriers ($\bar{x}(\sigma_{\bar{x}})$ = -0.007 (0.009); t (567) = -1.655, p_{holm} = 0.197; Fig. 4.2). No other main/interaction effect(s) of SERT were observed. These results show that, unlike the S/S and L/L groups, participants with the S/L SERT allele were showing altered adaptation of learning rate to environmental volatility, with a higher learning rate in stable, compared to volatile, phases.

Figure 4.2. Learning rate for volatile and stable phases as a function of SERT genotype

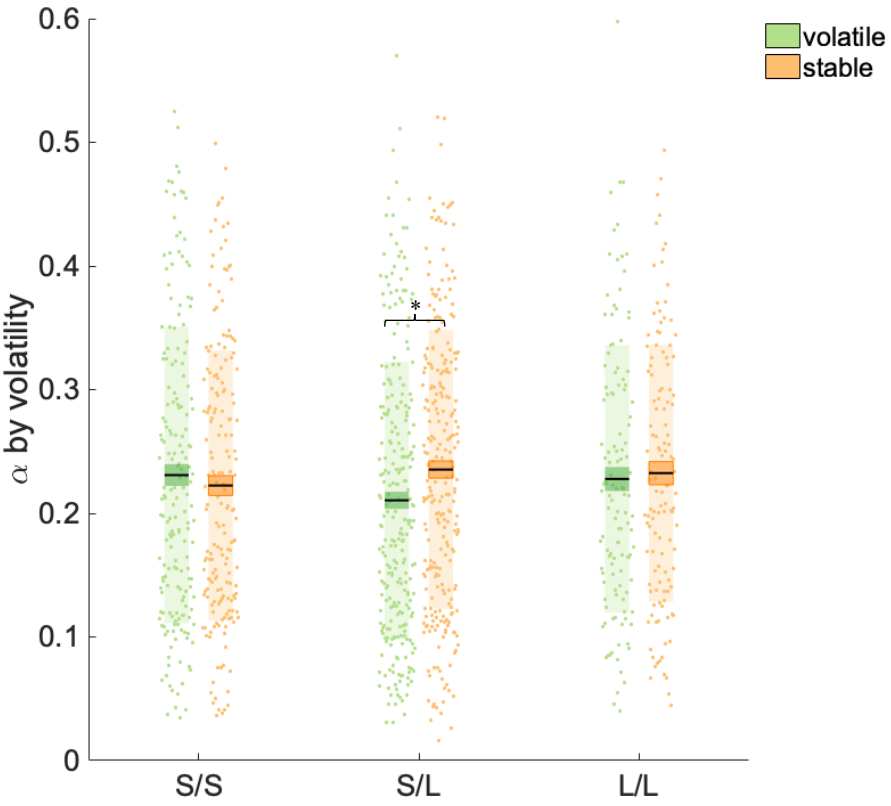


Figure 4.2. Learning rate (α) estimates for volatile and stable phases (collapsed across individual and social learning) as a function of SERT genotype. α estimates for S/L carriers varied significantly between volatile and stable phases. Data points indicate α estimates for individual participants (n = 570), boxes = standard error of the mean, shaded region = standard deviation, * indicates statistical significance ($p < 0.05$).

We had predicted an interaction such that serotonin-related genetic variation would account for significantly more variation in social as opposed to individual learning-related estimates. However, we did not observe a significant SERT genotype by information type interaction ($F(2,567) = 1.019$, $p = 0.361$, $\eta^2p = 0.004$). The BF_{excl} value for this interaction was 18.891, illustrating that the observed data is 18.9 times more likely under a model that excludes the interaction term compared to a model that includes the interaction term. Furthermore, we did not observe an interaction between SERT genotype, information type and volatility ($F(2,567) = 0.064$, $p = 0.938$, $\eta^2p > 0.001$, $BF_{excl} = 44.760$). Thus, the volatility-dependent effect of SERT genotype on α did not significantly differ for social and individual learning, with Bayesian analysis providing strong evidence ($BF_{excl} = 44.760$) against the presence of an interaction.

Effects of serotonin-related genetic variation on decision rule-related parameters

A separate ANOVA revealed a main effect of SERT genotype on ζ values ($F(2,567) = 6.326$, $p = 0.002$, $\eta^2p = 0.022$) with Bayesian analysis providing moderate evidence for this effect ($BF_{incl} = 7.681$). ζ values were significantly higher for the S/L ($\bar{x}(\sigma_{\bar{x}}) = 0.477\ (0.016)$) compared with the L/L genotype ($\bar{x}(\sigma_{\bar{x}}) = 0.376\ (0.024)$; $t(567) = 3.507$, $p_{holm} = 0.001$), and the S/S ($\bar{x}(\sigma_{\bar{x}}) = 0.459\ (0.019)$) compared with L/L groups ($t(567) = 2.701$, $p_{holm} = 0.014$) (Fig. 4.3). Values between the S/L and S/S groups did not significantly differ ($t(567) = 0.720$, $p_{holm} = 0.472$). These results suggest a bias, specifically for s-allele carriers (S/S and S/L groups), towards more highly weighting social, as opposed to individual, information. No main effect of SERT genotype on β values was observed ($F(2,567) = 1.153$, $p = 0.316$, $\eta^2p = 0.004$).

Figure 4.3 ζ values as a function of SERT

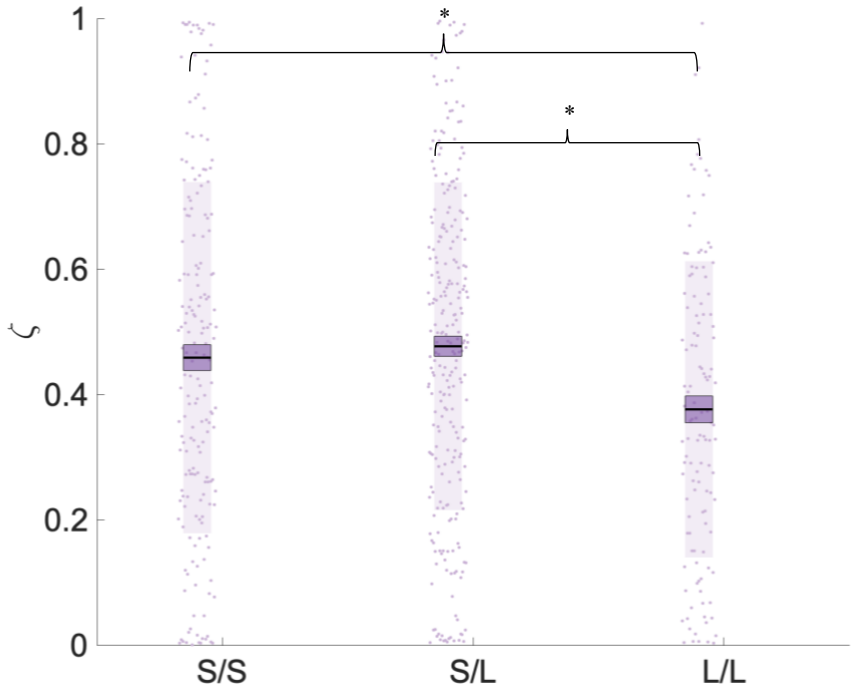


Figure 4.3. ζ values as a function of SERT genotype. ζ estimates for L/L carriers were significantly lower than both S/S and S/L groups. Data points indicate ζ estimates for individual participants (n = 570), boxes = standard error of the mean, shaded region = standard deviation, * indicates statistical significance ($p < 0.05$).

Effects of dopamine-related genetic variation on learning model-related parameters

For analysis with (square-root transformed) α as the DV, and DAT genotype as the BS factor, a main effect of information type was observed ($F(1,567) = 163.541$, $p < 0.001$, $\eta^2p = 0.224$, $BF_{incl} > 100$). $\alpha_{individual}$ ($\bar{x}(\sigma_{\bar{x}}) = 0.520$ (0.008)) was significantly higher than α_{social} ($\bar{x}(\sigma_{\bar{x}}) = 0.361$ (0.008)). In addition, there was a significant genotype by information interaction ($F(2,567) = 3.571$, $p = 0.029$, $\eta^2p = 0.012$, $BF_{incl} = 2.686$). Holm-corrected post hoc tests did not reach significance. However, an ANOVA conducted on difference scores ($\alpha_{individual} - \alpha_{social}$) revealed a main effect of DAT ($F(2,567) = 3.073$, $p = 0.047$, $\eta^2p = 0.011$), with the difference between $\alpha_{individual}$ and α_{social} significantly larger for 9/9 carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.210$ (0.032)) compared with 9/10 carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.127$ (0.012); $t(567) = 2.445$, $p_{holm} = 0.044$, $BF_{10} = 2.719$), and 10/10 carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.132$ (0.010); $t(567) = 2.361$, $p_{holm} = 0.044$, $BF_{10} = 2.386$). 9/10 carriers and 10/10 carriers did not differ ($t(567)$ -0.306, $p_{holm} = 0.760$, $BF_{10} =$

0.103) (Fig. 4.4). Thus, results demonstrate that both $\alpha_{individual}$ and α_{social} varied as a function of DAT genotype.

Figure 4.4 α_social and α_individual as a function of DAT genotype

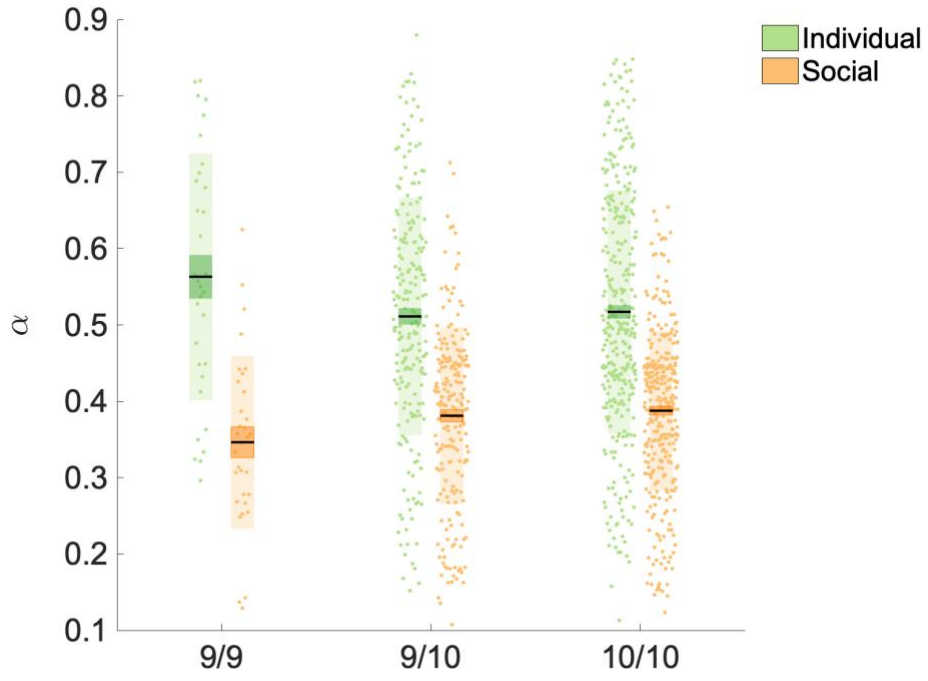


Figure 4.4. α_social and α_individual as a function of DAT genotype. The 9/9 group showed a significantly higher difference between α_{social} and $\alpha_{individual}$ compared with the 9/10 and 10/10 groups. Datapoints indicate α estimates for individual participants (n = 570), boxes = standard error of the mean, shaded region = standard deviation.

Additionally, a significant genotype by volatility interaction was observed (F (2,567) = 3.421, p = 0.033, η^2p = 0.012, BF_{incl} = 0.158). However, evidence for this was weak since the data were only 0.2 times more likely under a model that included, versus excluded, this interaction. Participants in the 9/9 group showed higher $\alpha_{volatile}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.253 (0.020)) compared with α_{stable} scores ($\bar{x}(\sigma_{\bar{x}})$ = 0.222 (0.019); t (567) = 1.990, p_{holm} = 0.660, uncorrected p = 0.046, BF_{10} = 1.25), whereas there was no significant difference between $\alpha_{volatile}$ and α_{stable} for 9/10 and 10/10 carriers ($t < 0.5, p > 0.05$) (Fig. 4.5). However, it should be noted that the difference between $\alpha_{volatile}$ and α_{stable} in the 9/9 group did not survive correction for post-hoc testing and Bayesian statistics provided weak evidence in

support of this difference. No other main/interaction effect(s) of DAT were observed. There was no significant DAT x information x volatility interaction ($F(2,567) = 0.312$, $p = 0.732$, $\eta^2p = 0.001$) with Bayesian statistics providing strong evidence against an interaction effect ($BF_{excl} = 16.950$).

Figure 4.5 Learning rate α as a function of volatility and DAT genotype.

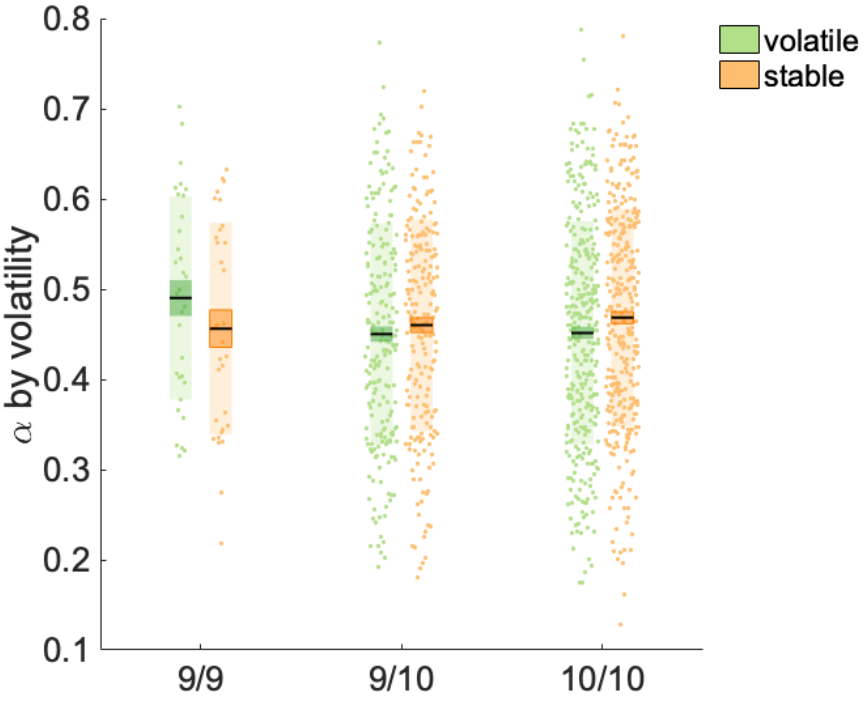


Figure 4.5. Learning rate α as a function of volatility and DAT genotype. $\alpha_{volatile}$ and α_{stable} differed for 9/9 carriers only. Data points indicate α estimates for individual participants (n = 570), boxes = standard error of the mean, shaded region = standard deviation, * indicates statistical significance ($p < 0.05$).

Finally, our analysis with (square-root transformed) α as the DV, and COMT genotype, revealed no main or interactions effects as a function of COMT genotype (all $p > 0.05$).

Effects of dopamine-related genetic variation on decision rule-related parameters

Separate ANOVAs revealed no main effect of DAT genotype on ζ values ($F(2,567) = 1.001$, $p = 0.368$, $\eta^2p = 0.004$) or β values ($F(2,567) = 0.098$, $p = 0.907$, $\eta^2p < 0.001$). A similar pattern of results was observed with COMT genotype: no main effect of COMT genotype on

ζ values (F (2,567) = 0.478, p = 0.621, η^2p = 0.002) or β values (F (2,567) = 1.307, p = 0.272 η^2p = 0.005).

Relationship between accuracy and model parameters

Participants with the S/L SERT allele showed altered adaptation of learning rate to environmental volatility, with a higher learning rate in stable, compared to volatile, phases. In addition, S/L allele carriers showed significantly higher accuracy on the behavioural task ($\bar{x}(\sigma_{\bar{x}})$ = 0.626 (0.003)) compared to the S/S carriers ($\bar{x}(\sigma_{\bar{x}})$ = 0.613 (0.004), t(567) = 2.643, p_{holm} = 0.025) and numerically higher compared with L/L carriers ($\bar{x}(\sigma_{\bar{x}})$ = 0.618 (0.005), t(567) = 1.467, p_{holm} = 0.286) (see Appendix 3.3 for an extended analysis of accuracy). We therefore explored how our model parameters were related to accuracy scores. We deployed a backwards regression model, with ζ , β, $\alpha_{\_volatile}$ and $\alpha_{\_stable}$ (collapsed across individual and social) as predictors and accuracy as the dependent variable. Accuracy was significantly predicted by the full model (R = 0.322, F (4, 565) = 16.318, p < 0.001), with $\alpha_{\_stable}$ (t (565) = 5.627, p < 0.001) and β (t (565) = 4.162, p < 0.001) significant positive predictors of accuracy. Removing ζ did not significantly improve the fit of the model (R^2 change = -0.000, F change (1,566) >0.001, p = 1.000).

Thus, higher accuracy correlated with higher α in stable compared with volatile phases, with subjects with the S/L genotype differing from the S/S and L/L groups. In contrast to predictions of a higher learning rate in volatile phases being more optimal than in stable, here we found the opposite; higher α in stable as compared with volatile phases, coupled with higher task accuracy for the S/L group. To determine if this pattern of results should statistically result in higher accuracy, an optimal learner model, with the same architecture and priors as the model employed in the current task, was fit to 100 synthetic datasets, resulting in average optimal learning rates: $\alpha_{optimal\_individual\_stable}$ = 0.16, $\alpha_{optimal\_individual\_volatile}$ = 0.21, $\alpha_{optimal\_social\_stable}$ = 0.17, $\alpha_{optimal\_social\_volatile}$ = 0.19. Scores representing the difference between α estimates and optimal α scores were calculated (α_{diff}: $\alpha - \alpha_{optimal}$). A RM-ANOVA on α_{diff} values with factors information source, volatility and SERT genotype, was conducted, with a significant volatility by SERT interaction observed (F (2,567) = 4.998, p = 0.007, η^2p = 0.017). Post-hoc tests were not significant after correction for multiple

comparisons (all $p_{holm} > 0.05$). However, α estimates were closest to optimal estimates for the S/L group in volatile phases (Fig. 4.6).

Figure 4.6 Estimated α compared with α optimal scores

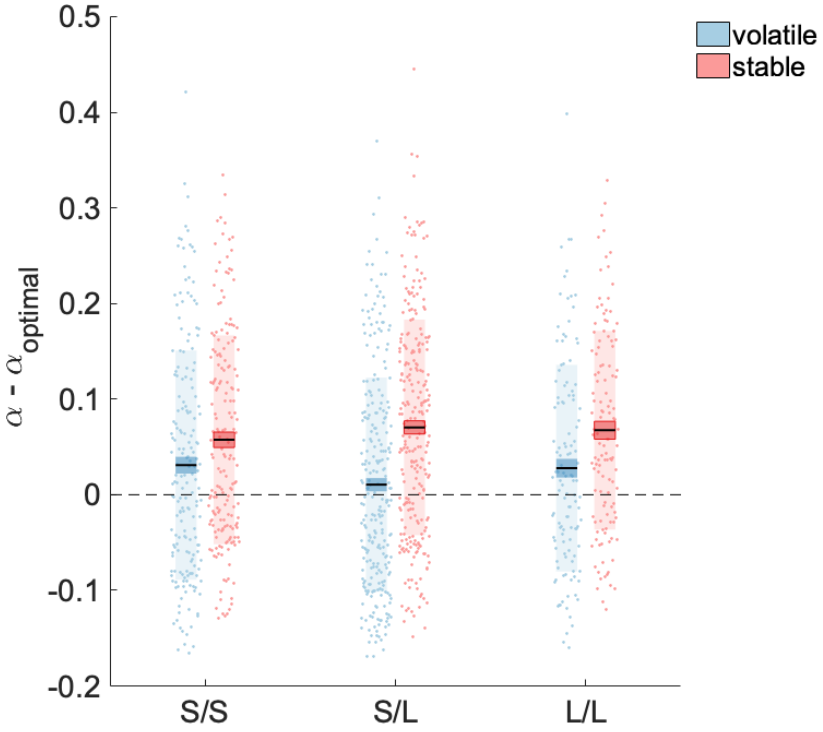


Figure 4.6. Estimated α compared with α optimal scores. α estimates were closet to optimal for S/L carriers in volatile phases. Data points indicate α estimates for individual participants (n = 570), dotted line indicates optimal scores, boxes = standard error of the mean, shaded region = standard deviation, * indicates statistical significance ($p < 0.05$).

When combining analysis of accuracy with our optimality analysis, these results suggest that S/L genotype carriers show reduced adaptation to volatility, driven by an optimal reduction in α in volatile phases. No differences were observed as a function of information type, however, suggesting similar volatility-dependent processes across both social and individual learning (common learning processes). However, the S/L group, rather than showing differences in learning rate depending on the social nature of the information source, were instead putting a greater weight on social information during decision-making. Evidence to support this comes from our analysis of ζ, representing the weight given to the social

information (see also an analysis of ideal learner scores for social information - Appendix 3.3). Finally, no effect of DAT genotype was found on accuracy ($F(2,567) = 0.279$ $p = 0.757$, $\eta^2p < 0.001$).

It was surprising to see that, on average, learning rates were significantly higher during stable (α_{stable} ($\bar{x}(\sigma_{\bar{x}})$ = 0.231 (0.005)) compared to volatile phases ($\alpha_{volatile}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.221 (0.005), ($t(569) = 2.061$, $p = 0.040$, $d = 0.086$). Could this indicate that participants were learning in the stable periods and simply "gave up" in the volatile periods? Given that $\alpha_{volatile}$ ($t(569) = 46.268$, $p < 0.001$, $BF_{10} > 100$) and α_{stable} ($t(569) = 50.206$, $p < 0.001$, $BF_{10} > 100$, both collapsed across social and individual learning) were significantly greater than zero, this potential explanation is unlikely. Furthermore, a Bayesian paired t-test revealed that frequentist statistics were somewhat misleading with respect to higher learning rates in stable versus volatile periods. In contrast with the frequentist statistics, Bayesian statistics provided anecdotal evidence *against* the presence of a difference in learning rates ($BF_{10} = 0.387$).

4.4 Discussion

The present study investigated whether naturally occurring variation in genes involved in the regulation of monoamine neurotransmission modulated learning from individual and social information in the same way. Participants completed an online social learning task, requiring learning from their own individual feedback, and from an additional, social source of information. Results did not confirm our original hypotheses: it was not the case that serotonin-related genetic variation accounted for more variation in social versus individual learning. Neither did we find that dopamine-related genetic variation accounted for more variation in individual versus social learning. Instead, our results revealed that variation in participants' ability to adapt their rate of both social *and* individual learning to suit the current level of environmental volatility covaried with SERT genotype. In addition, we observed significant relationships between dopamine-related genetic variation and both environmental volatility and learning source. However, as both $\alpha_{individual}$ and α_{social} varied as a function of DAT genotype, these results do not support the hypothesis that dopamine-related genetic variation would account for more variation in individual versus social learning. Finally, and in contrast with our predictions, we observed a significant relationship

between SERT genotype and a parameter within our decision rule which influences the extent to which participants are biased towards social, as opposed to individual, information.

Our original hypothesis, that genetic variation in DAT would vary with learning from individual information specifically, was not confirmed. However, our pattern of results does suggest that genetic variance in dopamine-related genes modulates social and individual learning in a dissociable manner. Specifically, we observed a DAT by information type interaction, with the 9/9 carriers showing numerically higher $\alpha_{individual}$ and lower α_{social}, relative to the other genotypes. Thus, cautiously indicating that genetic variation affecting the dopamine system might dissociate social from individual learning. This pattern of data is consistent with the proposal that humans possess social-specific neurochemical learning mechanisms, specialised for social learning, in line with predictions from some theories of cultural evolution [190],[197],[331]. On a cautionary note, however, participant numbers in the rarer 9/9 DAT genotype group may have skewed these results. To determine whether these results provided sufficient evidence to support the existence of dissociable genetic contributions to social and individual learning, we considered additional analysis, to supplement significance testing for the presence of main and interaction effects. Thus, effect sizes and Bayesian evidence for the inclusion of effects were calculated for all effects of interest. Bayesian analysis provided moderate evidence *against* the inclusion of the observed interaction between DAT genotype and information. Furthermore, post-hoc tests did not reach significance. Consequently, our results which hint at a dissociation between social and individual learning must be interpreted with caution.

Regarding variation in serotonin-related genes, we found that adaptation of learning rate to environmental volatility varied as a function of SERT genotype, but not between social and individual learning, with Bayesian statistics providing evidence against a difference in social and individual learning. Thus, suggesting common volatility-dependent processes across both social and individual learning. To conclude, results do not provide support for the presence of social-specific neurochemical learning pathways that are especially serotonin-dependent.

Separate interactions were observed between both serotonin-related genes and volatility and dopamine-related genes and volatility. Such interactions align well with theories proposing a modulatory effect of neuromodulators on learning about the underlying environmental statistics, or meta-learning [67],[91]. However, while dopaminergic signalling has previously been implicated in both increasing learning rate [94],[290] and in adjusting learning rate to suit

the environmental context [93], the current results provide preliminary, novel, evidence of a link between variation in serotonergic systems and meta-learning in human participants. Specifically, we observed a significant, but unpredicted, interaction between SERT genotype and volatility. While serotonergic signalling and meta-learning have not been directly investigated in humans, serotonergic signalling has been implicated in many aspects of reward learning [235],[238],[332]–[334], with serotonergic neurons active in response to prediction error-like signals [233],[239]. In addition, serotonin plays a crucial role in the modulation of behaviour in changing environments [92],[237],[239],[335],[336]. In non-human research, a recent study with mice performing a dynamic foraging task has proposed a key role for serotoninergic signalling in meta-learning [337], and a separate study has identified a role for serotonin signalling in altering learning rates [92]. Thus, the observed relationship between variation in serotonin genes and meta-learning is in accord with previous research [241]. On a neural level, serotonin-mediated modulation of learning rates could potentially be mediated via modulation of neurotransmitter release from midbrain dopaminergic neurons [338]. Therefore, we provide preliminary evidence in favour of a role for serotonergic signalling in meta-learning in human participants.

Interestingly, we observed that the heterozygous S/L group differed from both the short (s) and long (l) homozygotes in our analysis of learning rates. While the s variant, resulting from a deletion in the promoter region of the gene, has been associated with reduced SERT expression (leading to reduced serotonin reuptake and, therefore, greater synaptic availability [328],[329]), the opposite has been observed for the l-allele, namely a relative increase in SERT expression [339]. Thus, differences could be predicted between individuals who carry at least one copy of the 5-HTTPLR short variant (S/S and S/L) and homozygous carriers of the long variant (L/L). Similarly, to DAT genotype, however, the directional effects of SERT gene variants are hard to determine. For example, variants that confer both decreased *and* increased function have been associated with autism [340]. Moreover, the effects of genetic polymorphisms on 5-HT signalling are complex. The specific polymorphism investigated here, within the promoter region of the SERT gene, is thought to account for only approximately 25-30% of changes in 5-HT reuptake [341]. The effect on complex behaviours is thus harder to detect, making it difficult to directly map a causal effect of genetic variation to cognitive processes such as reinforcement learning. Future work could investigate a wider range of polymorphisms within the monoamine transporter genes, and interactions between

them, or utilise a pharmacological manipulation of serotonergic signalling, for example, via depletion of the serotonin precursor, tryptophan.

Focusing on the interaction between DAT genotype and volatility, 9/9 carriers showed a higher learning rate in volatile compared to stable phases, or greater adaptation to environmental volatility, in contrast with other genotypes. While these results are in accordance with previous research demonstrating a role for dopamine in adaptation to volatility [93], Bayesian analysis provided moderate evidence *against* the inclusion of the observed interaction between DAT genotype and volatility and (similarly to the previously-mentioned DAT by information interaction) post hoc tests did not reach significance. Furthermore, no correlation between genotype and accuracy was observed. Moreover, the effects of DAT polymorphisms on DAT expression (and therefore on extracellular dopamine concentration) are difficult to interpret [342]. For example, although recent meta-analyses suggest stronger associations between the 9R allele and DAT expression [343], both the 9R allele [344] and the 10R allele [320],[345] have been associated with higher striatal DAT availability. In contrast with genetic analysis, pharmacological manipulations of dopaminergic signalling, using drugs known to have a measurable effect on dopamine levels, allow directional perturbation of dopamine signalling. We follow up on this in Chapter 6, wherein participants complete the SLT under placebo and under administration of a dopamine antagonist, in a double-blind, within-subjects design.

In contrast to previous work, where higher learning rates were observed in volatile phases [88], we observed higher learning rates in stable phases overall in the current study. Previous research has demonstrated that healthy adults adapt their learning rate to the current environmental volatility [70],[88], with a higher learning rate predicted to be more adaptive when the environment is volatile, or reward probabilities are changing rapidly [89],[98]. This pattern of results forces us to consider whether participants were indeed learning in volatile phases, or rather, disengaging from the task. In refute of this, a higher learning rate in stable as compared with volatile phases was observed in one SERT group in particular, the S/L group. Crucially, this group also displayed higher accuracy on the behavioural task. This observation, coupled with our optimal learner analysis, whereby α estimates were closest to optimal estimates for the S/L group in volatile phases, suggests that the observed reduction in learning rate during volatile phases had a beneficial impact on task performance, bringing learning rate estimates within the optimal range. Additionally, we did not find a significant decrease in accuracy (across all participants) during volatile phases – as would be expected if

participants were disengaging from the task. Finally, results are in line with recent work by Cook and colleagues, where no difference in learning rate was observed between volatile and stale phases under placebo, but solely under the dopamine indirect agonist methylphenidate [93].

Focusing on social information, we observed a bias, specifically for s-allele carriers (S/S and S/L groups), towards a greater use of social information during decision making. We found evidence that the s-carriers, rather than showing differences in learning rate depending on the social/individual nature of the information source, were instead putting greater weight on social information during decision-making. Evidence to support this comes from our analysis of ζ, representing the weight given to the social information during decision-making (and optimal learner weights -Appendix 3) with s-allele carriers showing increased reliance on social information. Taken together, our results suggest that genetic variation in serotonergic signalling is linked to variation in the weight given to, or the attention paid to social information during learning, i.e., s-allele carriers are more sensitive to changes in social information. Indeed, gene variants associated with higher synaptic 5-HT availability (such as the s-allele of SERT), have been linked with greater sensitivity to social cues, across many different aspects of social cognition [240],[330]. One study in particular, focusing on observational fear learning, demonstrated increased social fear learning in s- relative to l-allele carriers [255]. Additionally, in the current task, increased use of social information correlated significantly with increased task accuracy, unsurprising as optimum performance on this task requires integrating the utility of both social and individual reward information. Thus, increased weighting of social information provides a potential explanation for improved task accuracy within the S/L group. In summary, we did not find evidence of SERT-related genetic dissociations between social and individual learning *per se*, but rather, an increased weighting of social information for s-allele carriers. Results are in accord with previous research demonstrating an increased sensitivity to social cues in s-allele carriers.

Taken together, our results suggest roles for both dopamine- and serotonin-related signalling with regard to meta-learning, or adjustment of learning rate to environmental context, and a bias towards increased weighting of social information for s-allele carriers. In accordance with theories predicting dissociable neurochemical pathways for social and individual learning, we observed weak evidence for an interaction between dopamine-related genotype and learning source [186],[189]. However, a confound exists, whereby, in this task, the social information is an additional, secondary source of information when compared to the

individual reward information. For example, it is less salient, in the form of a small orange frame and appears temporally after the individual information (large blue/green shapes). In addition, the validity of the social frame must be inferred from the individual or experienced outcome values (i.e., if the outcome indicator is blue and the orange frame surrounds the blue shape, then the social information is correct). Thus, the social learning source is also a secondary information source relative to the individual information (primary learning source), meaning that learning from individual and social information cannot be fully dissociated using this paradigm. Thus, it cannot be determined whether the dopaminergic mechanisms underpinning learning dissociate along a primary-secondary and/or a social-individual axis. In support of this, Cook and colleagues [93], measured the effects of catecholamine agonism on learning, using an adapted version of the current task, that included an explicitly 'non-social' control condition. While half of the participants were informed that the secondary information source (a red frame) represented social advice, the remainder were told that it represented non-social, additional information (output from a rigged roulette wheel). Thus, enabling the authors to determine whether dissociations between social and individual learning were better explained in terms of the social versus non-social or the secondary versus primary nature of the information source. No effect of catecholamine manipulation was found on learning from the secondary information source, regardless of whether participants believed it was "social" or not. Thus, suggesting that catecholamine-mediated dissociations during learning are potentially between primary and secondary, rather than individual and social, learning. To explore this further, we created a version of the behavioural task where social is the primary, and individual the secondary, learning source, i.e., we orthogonalized primary versus secondary and social versus individual learning. Task development and piloting will be described in the following chapter.

In summary, results suggest a role for serotonergic signalling in meta-learning or adaptation of learning rate to environmental volatility, as well as an influence on the weighting of social information during decision-making. Genetic variation in the dopamine transporter gene hints at the presence of a neurochemical dissociation between social and individual learning. However, results would be strengthened by the use of a pharmacological intervention and a behavioural task where social/individual and primary/secondary learning are orthogonalized.

Chapter 5: Manipulating the learning hierarchy – a pilot analysis/task development study

In the previous chapter, I demonstrate a dissociation between social and individual learning at the level of naturally occurring variation in genes important in the regulation of dopamine signalling. However, the social/non-social nature of the information source during learning is confounded with whether the information source is the primary source of information, or an additional source, with its utility needing to be inferred from the primary reward feedback. This chapter describes the development and piloting of a version of a social learning task where the primary and secondary nature of the learning sources are switched, resulting in social information acting as the primary source of learning, and individual information the secondary, additional information source. This chapter is therefore written with a focus on describing task design and development. The overall aim of this chapter is to demonstrate that the social/individual and primary/secondary nature of information can be orthogonalized. I investigate whether manipulation of several aspects of the task structure results in the social information becoming the primary information source during learning. These manipulations would then allow testing of the prediction that manipulations in dopamine signalling affect learning from the primary source, regardless of the social nature of the information source. To this end, the manipulated task version outlined in this chapter, and the standard version, are employed in a between-subjects pharmacological intervention (Chapter 6).

Supplementary materials for this chapter are in Appendix 4.

5.1 Introduction

It is unclear as to whether social learning is underpinned by uniquely social mechanisms or domain-general associative learning processes. In attempting to disentangle social and individual learning, cognitive neuroscientific studies have resulted in mixed evidence. Some studies investigating social learning have found that social learning is associated with the same dopamine-mediated prediction error signalling as non-social, or individual learning. These studies, however, usually feature the social information as the primary source of information; participants are encouraged to learn primarily from the social information, which is often the sole information source [221]–[223],[346]. In other, conflicting studies, where social and individual learning have been directly compared, dissociations have been found. However, in studies where dissociations have been found the social information is typically a secondary and indirect source of learning, when compared to the primary or directly experienced individual reward information [93],[95],[258],[347]. In these paradigms, the social information is usually less salient, with its utility needing to be inferred from the individual reward feedback. Consequently, when it is the primary learning focus, social learning may be underpinned by the same dopamine-rich mechanisms as individual learning, but not when it comprises a secondary, additional element. Therefore, observed dissociations between social and individual learning may not reflect social versus individual dissociations but rather, dissociations between primary and secondary learning status.

For example, a commonly used index of social learning is a social learning task (SLT), originally developed by Behrens and colleagues [95]. This task allows learning rates and other indices of learning to be estimated from social and individual reward information at the same time but in a dissociable manner. During the task, participants are required to choose between a blue or green shape to gain points, receiving feedback in the form of a blue/green indicator, which directly informs participants about the utility of their choice (blue/green). Participants can also learn from an additional source of information, a red frame surrounding one of the choices (the orange frame in Chapter 4), which represents the 'social' information. This, however, is a secondary source of information to the primary or experienced information, as the utility of the frame must be inferred from the experienced outcome values (i.e., if the indicator is blue and the red frame surrounds the blue shape, then the red frame is correct). In addition, the red frame appears temporally after the blue/green indicators and is less visually salient. Thus, the social learning source is also a secondary information source relative to

individual information and learning from individual and social information cannot be fully orthogonalized using extant paradigms.

Cook and colleagues provided preliminary evidence against a neurochemical dissociation between social and individual learning mechanisms. Using the social learning task (SLT) adapted from Behrens et al., [95], they demonstrated that manipulation of catecholamine signalling via methylphenidate (MPH) affected learning from a primary information source only. The task was adapted whereby half of the participants were informed that the secondary information source (red frame) represented social advice, the remainder were told that the red frame represented output from a rigged roulette wheel, enabling the authors to determine if dissociations between social and individual learning were better explained in terms of the social versus non-social or the secondary versus primary nature of the information source. MPH enhanced subjects' ability to optimize their learning rate for the current level of environmental volatility, specifically while learning from the primary (individual) reward, with no effect of MPH on learning from the secondary information source, regardless of whether participants believed it was "social" or not. These results suggested that the differing effect of catecholamine perturbation dissociated between learning from primary and secondary sources, rather than between individual and social sources. However, in this study, the social source was an additional, secondary, source of information, meaning that social nature (social versus individual) and rank (primary versus secondary status) were not fully orthogonalised. To provide positive evidence that dissociations lie between learning from primary and secondary, rather than between social and individual information, it is therefore necessary to orthogonalize the social/individual and primary/secondary nature of the information sources. This will allow investigation of the effects of these factors independently of one another and help in determining which factor accounts for previously observed dissociations. We, therefore, aimed to develop a version of this task whereby the primary and secondary nature of the learning sources were switched, resulting in social information acting as the primary source of learning, and individual information as the secondary, additional information source.

Within the SLT, there are three main differences that might be contributing to the primary versus secondary distinction between the blue/green boxes and the red frame: 1) saliency, 2) direct association with reward, 3) temporal order. The social information is visually less salient, appearing as a thin frame, and temporally appears after the individual information. Finally, the reward information is linked directly to the individual information, meaning that

the utility (correct/incorrect) of the social information must be inferred, in a secondary manner, from the individual reward feedback (i.e., which stimulus is correct). There exist some well-known phenomena in the associative learning literature, that suggest that these factors would result in stronger associations between the blue/green stimuli and reward than between the social stimulus and reward. For example, blocking [209],[348] occurs when a stimulus becomes associated with reward, blocking any further associations between that same reward and other stimuli; i.e., learning of a particular stimulus-reward association blocks further learning [24]. In the case of the current paradigm, as the blue/green stimuli are first associated with reward, they may block learning to the red frame (social information). Additionally, there is evidence that temporal primacy can reduce the effects of blocking [349], suggesting that the temporal order of stimuli presentation is highly important during learning [25]. In the current paradigm, the red frame appears temporally after the blue/green stimuli, meaning that any blocking effects of the blue/green stimuli are not altered. Another mechanism that may be occurring in the SLT is overshadowing, referring to a phenomenon whereby when two stimuli are present in compound and only the more salient (sometimes referred to as "intense") stimuli is associated with reward [350], overshadowing any learning to the less salient stimulus by taking up attentional resources. This, alongside blocking, is often referred to as cue competition, where one cue or stimulus prevents learning of another [24],[350],[351]. Indeed, associative learning between a stimulus and outcome can be modulated by the attention given to a particular stimulus [193],[201], with attention affected by a number of different factors, including the visual saliency of stimuli [352]. In addition, whether or not the stimulus is directly relevant to task performance, i.e., whether it is directly associated with a reward, also modulates the salience of the stimulus and attention given to the stimulus during learning [353]–[355], in line with an incentive motivation view of attentional orienting [356].

In order to make the social information into the primary information source, we adapted several aspects of the SLT structure, so that the social information was more salient, appeared first, and was linked directly to reward feedback. In this modified version of the task (referred to hereafter as the 'social-primary' task condition) the social information appears in the form of a large, solid red shape rather than a thin, red frame and is presented to participants first, before the blue/green frames (individual information) appear. Crucially, reward feedback is altered, such that it now informs participants as to whether the social group are correct or incorrect, rather than which shape (blue/green) is correct. All other aspects of the task,

including probabilistic schedules, are identical to the standard task version (referred to as the 'individual-primary' task from hereon).

To assess whether task manipulations had made the social information into the primary learning source, we aimed to quantify the weight given to both social and individual information during learning. Participants (n = 77) completed the social-primary version of the SLT. Both behavioural and computational measures of learning were collected. We compared these measures with data from previous work (n=102) [93] using the standard version of the task (individual-primary group). We predicted an increase in the weight or reliance given to social information and/or higher learning rates for social information within the social-primary versus the individual-primary groups.

5.2 Materials and Methods

5.2.1 Participants

Data was collected from 77 volunteers, recruited via the research participation scheme (RPS) system at the University of Birmingham (aged 18 – 37 years, mean (SD) = 20.8 (4.1); 58 women). All participants gave informed consent to participate. The study was approved by the University of Birmingham local ethics committee (ERN_16-021AP5). Participants received monetary compensation on completion of both testing sessions, at a rate of £7 or 1 research credit per hour.

Effect size calculation

Based on estimates of effect size from previous work [93], a sample of 71 participants was required, to obtain 90% power at 0.05 alpha, to detect the effect of information type (individual, social) on beta values.

5.2.2 General procedure

Participants took part in the experiment over two separate test days and completed a wider battery of tasks, including measures of interoception, emotion recognition and production, and movement tracking. The behavioural task reported here (social learning task (SLT); Appendix 4.1) was always completed on test day 1 and lasted approximately 35 minutes.

Participants were seated approximately 30cm from a computer screen. Stimuli were displayed using PsychToolBox and the task was programmed using MATLAB R2017b (The MathWorks, Natick, MA). Before the main task, participants completed a step-by-step on-screen practice task (10 trials) in which they learnt to choose between the two options to obtain a reward and learned that the "advice" represented by the frame(s) could help in making the correct choice in some phases. To ensure that participants were making a conceptual distinction between the social and individual learning sources, we required participants to complete a short pre-task quiz (Appendix 4.2), testing their knowledge, after the practice task. Participants were required to repeat the practice round until they achieved 100% correct score in the quiz, meaning that all participants understood the structure of the task and that the red shape represented *social* information. Furthermore, after the experiment, participants completed a feedback questionnaire (Appendix 4.3). Answers confirmed that participants understood the difference between, and paid attention to both, individual and social sources of information. Participants were informed as to whether they had earned a £5 bonus after the session. However, due to ethical considerations, all participants received the bonus.

Social-primary social learning task

Participants completed 120 trials of the SLT [309] (Fig. 5.1A). Participants were required to choose between two shapes (blue or green) to gain points. On each trial, participants were presented with two grey placeholders. One placeholder was filled with a red box, which participants were informed represented the most popular choice made by a group of participants who had previously completed the task, i.e., it indicated the 'social' choice or option. Blue/green frames then appeared around the placeholders. Participants were then asked to choose between the two colours. Outcome or social reward information appeared in the form of a tick or a cross, which primarily informed participants about the utility of the group's decision. The indicator thus primarily informed participants about whether the social group had been rewarded (and thus going with them would have resulted in points scoring but going against them would not) on the current trial. Whether the blue (or green) frame surrounded the correct or incorrect option could, secondarily, be inferred from the indicator. Participants were informed that the group's "advice" (the red stimulus) would fluctuate between being predominantly correct and predominantly incorrect, and that the task followed 'phases' wherein sometimes the blue, but at other times the green stimulus, was more likely to be associated with reward (Fig.5.1B). Both the probability of blue being correct and the

probability of the red shape indicating the correct answer varied according to four different probabilistic schedules, including stable and volatile blocks.

Figure 5.1 Behavioural task manipulations - social-primary task

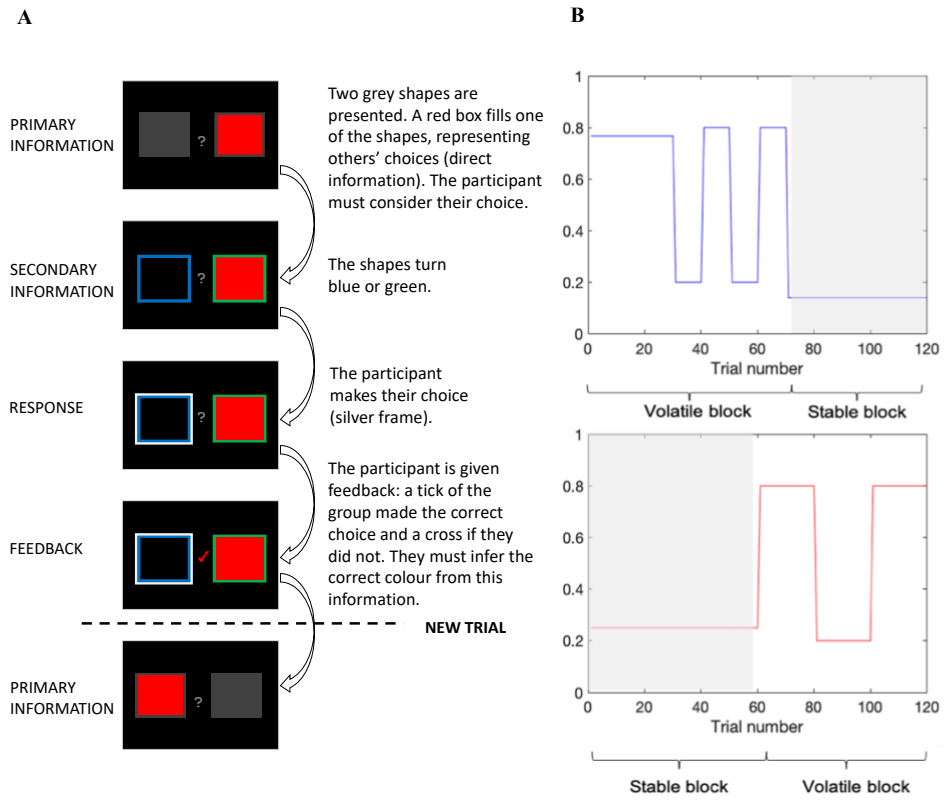


Figure 5.1. Behavioural task. A. Task structure. Participants selected between going with, or against a red box, which represented the social information. On each trial, the red box was displayed. After 1-4s, blue and green frames appeared. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a tick or a cross. This feedback informed participants if going with the group was correct or incorrect, from this feedback participants could infer whether the blue or green frame was correct. B. Stable and volatile phases. The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of blue being correct and the probability of the red shape indicating the correct answer.

5.2.3 Behavioural Measures

Accuracy & Reaction time

Accuracy was defined as the proportion of correct responses. Accuracy scores were summed and averaged, with Shapiro-Wilk (SW) testing failing to reject the null hypothesis that scores

followed a normal distribution ($p > 0.01$). Reaction time (RT) was calculated as time from stimulus presentation to response in milliseconds (ms). Scores did not meet normality assumptions (SW test, p values < 0.01); RT values were therefore square root transformed prior to analysis. After transformation, we failed to reject the null hypothesis that scores followed a normal distribution (SW test, p values > 0.01).

Optimal learner model

The influence of each information source (primary and secondary) on choices was quantified by regressing two "optimal learners" against subjects' choices. Learner models comprised an optimal "individual learner model" [88] simulating an optimal learner who learns solely from individual information (the blue and green stimuli) and a "social learner model" which simulated an optimal learner who learns solely from the social information (red stimuli). Both models were (separately) regressed against each individual participant's choice data using binomial logistic regression, with model predictions from the primary and secondary models as continuous predictor variables and participant response as the dependent variable (0/1). For each participant, this produced two parameter estimates, or standardised beta weights, each representing the degree to which individual experience and social information explained choices. For example, a participant whose choices were more strongly influenced by the social information than the individual information would have a high social $\beta_{optimal}$ value, and a low individual $\beta_{optimal}$ value (see Appendix 2.2 for full details of the optimal learner model).

Win-stay lose-shift beta regression (WSLS β)

WSLS behaviour was quantified by regressing participant choices (choose blue/green; follow/deviate from social advice) against a WSLS strategy. The analysis is identical to that used by Cook and colleagues [93] and is reproduced here for completion. A separate regressor was created for each participant for individual and social information, representing the choice that they would have made if following an ideal WSLS strategy. For individual learning, if on the previous trial (t-1) the participant had chosen blue and won, the regressor predicted that the participant should repeat this choice and choose blue on trial t, i.e., win and stay (coded as 1). If, however, the participant had chosen blue on trial t-1 and lost, the regressor would predict a shift to a green choice, or a lose-shift (coded as 0). Similarly, for learning from social information, if the participant had previously followed the social advice and won, the regressor would predict that they should stay with the advice; if they lost, the regressor would predict that the participant should *shift* such that they do not follow the frame's advice on the

current trial. WSLS regressors were regressed against each participant's choice data, separately for individual and social information and volatile and stable phases, resulting in four beta values for each participant. Beta values did not meet assumptions for normality (SW test, p values < 0.01); values were log-transformed prior to analysis.

Computational modelling framework

Further to behavioural analysis of participant choices, participant response was modelled using an adapted Rescorla-Wagner learning model [24]. The model relies on the assumption that updates to choice behaviour are based on prediction errors, i.e., the difference between an expected and the actual outcome. Participants were assumed to update their beliefs about outcomes based on sensory feedback (perceptual model), and to use this feedback to make decisions about the next action (response model) (see Appendix 2.3 for full model details). Model fitting was performed using scripts adapted from the TAPAS toolbox [230] (scripts available at OSF link https://tinyurl.com/b3c7d2zb). Parameters were fitted separately for each participant's choice data. Learning rate (α) was estimated for each participant, separately for primary and secondary information and volatile and stable phases, resulting in four estimated learning rates per participant. β and ζ values were also estimated for each participant.

Data pre-processing

Datasets were excluded based on the following: accuracy < 50% under placebo, conformed to the group choice on > 80% trials, chose the same side (left/right) or colour on > 80% trials, incomplete datasets (less than 120 trials completed). Five subjects were excluded, resulting in a final sample of n = 70.

5.3 Results

The main aim of the task manipulation was to make the social information into the primary learning source. We compared behavioural and computational measures of learning from the current pilot study (social-primary group) with a sample from previous work (n = 102) using the original version of the task (individual-primary group). All analyses, therefore, included group (social-primary, individual-primary) as a between-subjects (BS) factor, to assess the effect of the task manipulation on different variables. Participants from the individual-

primary group were excluded based on the exclusion criteria described above, resulting in a final sample of n = 98. Due to time constraints, groups were not matched on demographic variables (Table 5.1).

Table 5.1 Demographic information

n = 168	Individual-primary group (n = 98)	Social-primary group (n = 70)			
	Mean (SD)	Mean (SD))	t (1,166)	X^2 (1, N = 166)	p
Gender (*n* males: *n* females)	47:51	17:53		9.704	0.002
Age	21.5 (2.3)	20.8 (4.1)	1.409		0.161

Note: SD refers to standard deviation. Gender differed between the groups. Age did not significantly differ between groups.

Accuracy and reaction time

We first compared accuracy across task groups (individual-primary, social-primary). A repeated measures ANOVA (RM-ANOVA) was carried out, with accuracy as the dependent variable (DV), volatility (volatile, stable) as a within-subjects (WS) predictor variable and schedule (1-4) and task group as between-subjects predictors. Accuracy varied significantly as a function of group (F (1,160) = 4.469, p = 0.036, η_p^2 = 0.027), with higher accuracy for the individual-primary group (mean (standard error) ($\bar{x}(\sigma_{\bar{x}})$ = 0.630 (0.006) compared to the social-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.612 (0.006)) (Fig. 5.2A). Neither schedule (F (3, 160) = 1.923, p = 0.128, η_p^2 = 0.035) nor volatility (F (1, 160) = 0.037, p = 0.847, η_p^2 < 0.001) had a significant effect on accuracy scores and no other interactions with group were observed (all

$p > 0.05$). However, a significant volatility by schedule interaction was observed ($F(3, 160) = 15.227$, $p < 0.001$, $\eta_p^2 = 0.409$). For participants in schedules 1 and 2, no difference in accuracy was observed between stable and volatile phases ($p_{holm} > 0.05$). However, for schedules 3 and 4, accuracy significantly differed as a function of volatility; for schedule 3, higher accuracy was observed in volatile ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.678 (0.012) compared to stable phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.562 (0.011); $t(160) = 7.925$, $p_{holm} < 0.001$). For schedule 4, however, higher accuracy was observed in stable ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.678 (0.013) compared to volatile phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.563 (0.009); $t(160) = 7.533$, $p_{holm} < 0.001$). However, participants were counterbalanced to different randomisation schedules, with the proportion of participants assigned to each schedule not differing between groups (X^2 (1, $N = 168$) = 0.778, $p = 0.855$). The observed interaction therefore does not explain difference in accuracy between groups. Finally, log-transformed mean reaction times (RT) (ms) were submitted to an ANOVA, with group and schedule as BS factors. RT was significantly greater in the social-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.000 (0.022)) compared with the individual-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.926 (0.019); $F(1,160) = 6.361$, $p = 0.013$, $\eta_p^2 = 0.038$) (Fig. 5.2B). This pattern of results, namely lower accuracy, and slower reaction time, suggests that the social-primary condition was more difficult compared to the individual-primary condition. However, as results were taken from two separate experiments and groups were not matched, this cannot be confidently concluded from the current data.

Figure 5.2 Accuracy and reaction time

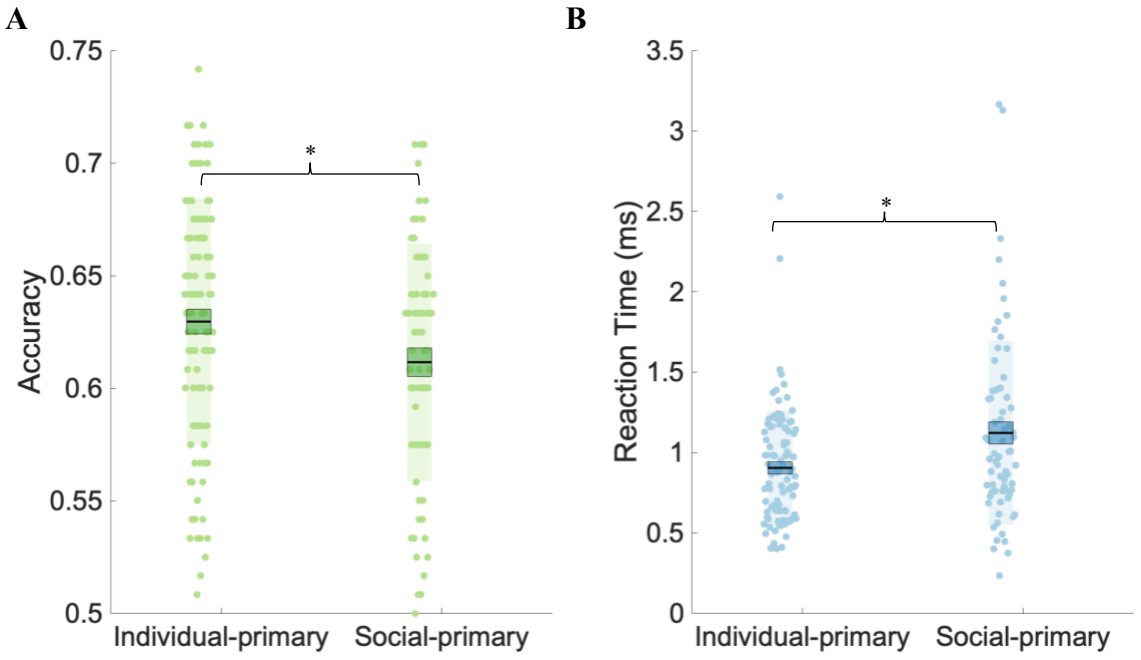


Figure 5.2. A. Mean accuracy across task groups. Accuracy was significantly higher for the individual-primary group. B. Mean reaction time across task groups. Reaction time was significantly slower for the social-primary group. Data points indicate mean accuracy for individual participants (total n = 168), bold point indicates the mean, bold line indicates standard error of the mean (1 SEM), * indicates statistical significance ($p < 0.05$).

Optimal learner model

We predicted an increase in the influence of social information on learning in the social-primary group, indexed by $\beta_{optimal}$ scores. $\beta_{optimal}$ values were submitted to a RM-ANOVA with factors information source (individual, social) and group (social-primary, individual-primary). A main effect of group was observed ($F(1,166) = 8.980$, $p < 0.001$, $\eta_p^2 = 0.051$), with (transformed) $\beta_{optimal}$ values (averaged across individual and social) significantly higher for the individual-primary group ($\bar{x}(\sigma_{\bar{x}}) = 1.334$ (0.016)), compared with the social-primary group ($\bar{x}(\sigma_{\bar{x}}) = 1.267$ (0.016) (Fig. 5.3). No main effect of information (individual/social) was observed ($F(1,166) = 0.976$, $p = 0.323$, $\eta_p^2 = 0.006$)). In contrast to predictions of increased $\beta_{optimal}$ for social information in the social-primary group, we did not observe a significant interaction between information and group ($F(1,166) = 2.438$, $p = 0.120$, $\eta_p^2 = 0.014$), with $\beta_{optimal}$ for social information not differing significantly between groups ($p_{holm} = 1.000$). However, in the social-primary group, $\beta_{optimal}$ for individual information ($\bar{x}(\sigma_{\bar{x}}) = 1.251$ (0.029) was significantly lower compared with the individual-primary group ($\bar{x}(\sigma_{\bar{x}}) =$

1.375 (0.033); t(166) = 2.906, p_{holm} = 0.024). These data suggest, that, rather than significantly increasing the influence of social information on learning, the task manipulation had reduced the influence of individual information on learning, with $\beta_{optimal}$ values comparable for individual and social information in the social-primary group. Crucially, $\beta_{optimal}$ weights in the social-primary group were significantly greater than zero for both information sources (individual: t (167) = 12.743, p < 0.001; social: t (167) = 13.539, p < 0.001), demonstrating that use of optimal models of information explained a significant amount of variance in the use of individual and social learning sources.

Figure 5.3 βoptimal for individual and social information across task groups.

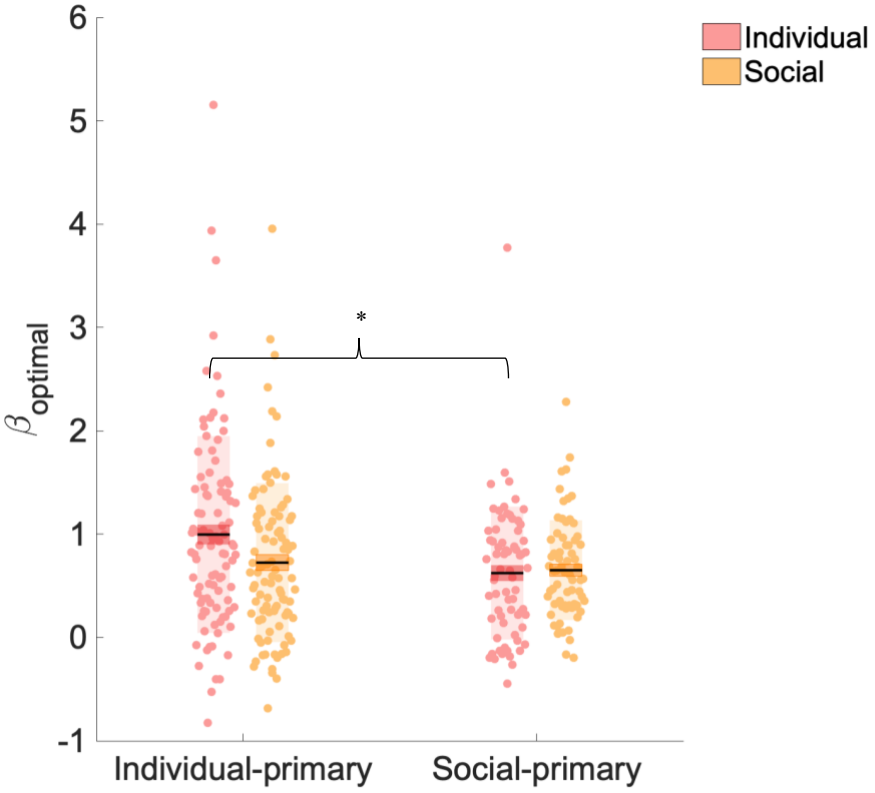


Figure 5.3. Beta weights ($\beta_{optimal}$) for individual and social information across task groups. Data points indicate estimated β for individual participants (n = 168), bold point indicates the mean, bold line indicates standard error of the mean (1 SEM), * indicates statistical significance (p_{holm} < 0.05). $\beta_{optimal}$ for individual information was significantly lower for participants in the social-primary group compared with the individual-primary group. $\beta_{optimal}$ for social information did not significantly differ between groups.

Analysis of win-stay, lose-shift behaviour

Next, WSLSβ values were compared across groups. A RM-ANOVA, with information (individual, social) and volatility (volatile, stable) as WS factors, and group (individual-primary, social-primary) as the BS factor, revealed a significant main effect of information type ($F(1,166) = 34.969$, $p < 0.001$, $\eta_p^2 = 0.174$) on (log-transformed) WSLSβ scores, with more WSLS behaviour for individual information. In addition, a significant main effect of volatility was observed ($F(1,166) = 69.409$, $p < 0.001$, $\eta_p^2 = 0.295$), with higher WSLSβ scores during volatile phases. Importantly, a significant group by information interaction was observed ($F(1,166) = 4.078$, $p = 0.045$, $\eta_p^2 = 0.024$, $BF_{incl} = 2.688$). Although post-hoc tests did not reach significance, a trend was observed whereby WSLSβ scores for individual information were (non-significantly) reduced in the social-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.119 (0.011)) compared with the individual-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.138 (0.007), t = -1.841, $p_{holm} = 0.133$). In contrast, WSLSβ scores for social information showed a (non-significant) increase in the social-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.085 (0.007)) compared with the individual-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$)= 0.070 (0.006), t = 1.435, $p_{holm} = 0.152$) (Fig. 5.4).

Figure 5.4 WSLSβ for individual and social information

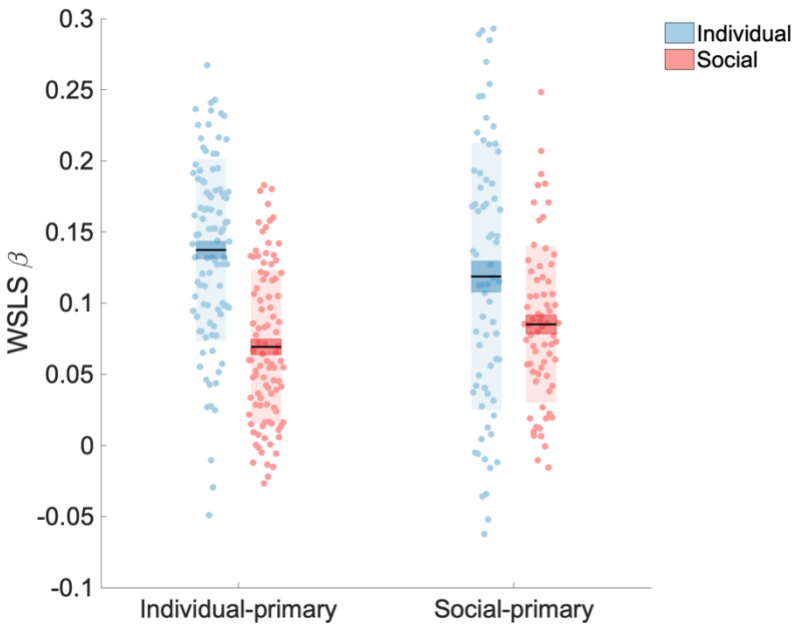


Figure 5.4. WSLSβ for individual and social information. Data points indicate (log-transformed) WSLSβ for individual participants (n = 168), bold point indicates the mean, bold line indicates standard error of the mean (1 SEM).

Results suggest a trend towards increased WSLSβ scores for social information in the social-primary group, coupled with a decrease in WSLSβ scores for individual information, with Bayesian analysis providing anecdotal-moderate evidence for this interaction effect.

Computational modelling parameters

Next, computational modelling parameters α, β and ζ were compared across task groups. We predicted that, if social information was the primary source of information for participants in the social-primary group, an increase in learning rates for social information would be observed, relative to learning rates for individual information. Specifically, we predicted an interaction between group and information type. A RM-ANOVA with factors information, volatility, and group, and α as the DV revealed main effects of information ($F(1,166) = 63.126$, $p < 0.001$, $\eta_p^2 = 0.276$), reflecting that $\alpha_{individual}$ was higher than α_{social} in both groups. No other main/interaction effects were observed. Surprisingly, and in contrast with our analysis of WSLSβ values, no significant group by information interaction was observed ($F(1,166) = 0.518$, $p = 0.473$, $\eta_p^2 = 0.003$), with Bayesian analyses providing moderate evidence against the presence of an interaction effect ($BF_{excl} = 5.108$). However, WSLSβ and α scores were significantly correlated for both individual (volatile: $r = 0.783$, $p < 0.001$; stable: $r = 0.822$, $p < 0.001$) and social information (volatile: $r = 0.666$, $p < 0.001$; stable: $r = 0.573$, $p < 0.001$). Despite the lack of a significant interaction, a trend towards a numerical increase in α_{social} was observed in the social-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.213 (0.014)) compared with the individual-primary group ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.193 (0.011)), $t(166) = 1.134$, $p = 0.258$, $d = 0.177$).

We then compared ζ across task groups. ζ represents the relative weighting of individual and social sources of information, with higher values indicating a bias towards the over-weighting of social relative to individual. An increase in ζ was predicted for participants in the social-primary, relative to the individual-primary group. Although there was no significant effect of group on ζ scores ($F(1,166) = 1.616$, $p = 0.205$, $\eta_p^2 = 0.010$), ζ scores showed a numerical increase in our social-primary ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.549 (0.030), relative to individual-primary, groups ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.500 (0.025)). However, a Bayesian independent samples t-test provided anecdotal evidence against higher ζ scores for the social-primary group ($BF_{01} = 1.58$). Finally, there was no effect of group on β values ($F(1,166) = 0.049$, $p = 0.825$, $\eta_p^2 < 0.001$).

In summary, although inconclusive, results suggest a trend towards an increase in the weighting of social information for participants in the social-primary group. Results do not, however, support an increase in learning rates with regard to social information for the social-primary group.

5.4 Discussion

The existence of neurochemical pathways specific to social learning is widely debated, with contrasting findings in the current literature. However, there is evidence to suggest that learning can be neurochemically dissociated at a primary versus secondary, rather than a social versus individual level [93]. As current paradigms lack the ability to separate these factors, we here aimed to develop a version of this paradigm that would enable us to orthogonalize social versus individual and primary versus secondary, to use in the context of a pharmacological manipulation in Chapter 6. Different aspects of the task structure, namely the salience of the social information, the temporal order of stimuli presentation and the nature of the reward feedback were manipulated, in an attempt to make the social information into the primary source of learning. We conducted a pilot study with healthy individuals (n = 72) who completed the adapted SLT (social-primary group). These results were compared with results from a separate sample (n = 98) who had previously completed the standard SLT (individual-primary group, data from Cook et al., 2019 [93]), to test if task manipulations had made the social information into the primary source of learning. Our analyses suggest that our task manipulation affected several indices of learning, with results in the expected direction. However, both frequentist and Bayesian analyses suggest that stronger manipulations are required, in order to fully orthogonalize these factors.

First, participants in the social-primary group showed significantly reduced individual $\beta_{optimal}$ weights, relative to the individual-primary group. $\beta_{optimal}$ weights represent the extent to which participants are following an optimal strategy for learning solely from individual or social information during learning, thus providing an index of reliance on each information source. Results show a reduced reliance on individual information in our social-primary task group, relative to in the individual-primary group, although Bayesian evidence was inconclusive. Next, our analysis of WSLSβ scores, representing the extent to which participants rely on performance on the previous trial, and commonly used as a proxy for

learning rate, showed a significant interaction between group and information type. Although post hoc tests were non-significant, results showed an increase in the use of a WSLS strategy for social information within the social-primary compared with the individual-primary group, coupled with a corresponding decrease in WSLS strategy for individual information. Thus, suggesting an increase in learning from social information in the social-primary group, with Bayesian analysis supporting these results. Analysis of the decision-rule parameter estimate ζ, representing the weighting of social, relative to individual information, revealed a trend towards increased weighting of social information in the social-primary group, although Bayesian analyses provided anecdotal evidence against a difference between groups. Finally, although results from our analyses of parameter estimates from a computational model of learning were non-significant, with Bayesian analyses providing evidence against differences in learning rates as a function of group and information source, the data showed a trend in the same direction as our behavioural measures of learning. Taken together, the data reported here suggest that participants in the social-primary group were putting more weight on the social information, in comparison with participants in the individual-primary group.

While the pattern of results trend in the predicted direction and provide tentative evidence that participants are putting more weight on social information during learning, several cautionary points should be noted. First, our analysis of optimal beta weights showed that, while the weight given to individual information significantly decreased for the social-primary group, the corresponding increase in the weighting of social information, although showing a trend in the predicted direction, was not significant. Second, although our behavioural proxy measure of learning rate, indexed by win-stay, lose-shift behaviour showed a trend towards an increase in learning from social information in the social-primary group, post hoc tests were not significant. Finally, we did not observe a corresponding effect in computational measures of learning rate.

It might therefore be the case that, while participants in the social-primary group are giving more weight to the social information, compared with those in the individual-primary group, further manipulations would strengthen this. For example, participants are presented with two shapes (blue/green) to choose between, and only one source of social information (red shape). While this cannot be altered, perhaps more importantly, participants' choice during the current pilot study was framed as choosing between two shapes, similarly to the individual-primary task. Specifically, participants were told that they had to choose between two shapes, which might encourage participants to think primarily of the blue and green shapes. The

possibility that this factor prevented the social information from being the primary source of information must be considered. Indeed, there exists a wide body of literature demonstrating that instructions can actively influence both behavioural and neural learning-related responses [227],[357]–[360], suggesting that prior experimental instructions affect performance on this task. Building on this idea, task instructions provided to participants were updated for use in the paradigm employed in Chapter 6. Participants were here primarily asked to choose between going with, or against, the group's choice, i.e., with or against the social information (Appendix 4.4).

For future task development and for adapting current paradigms to include appropriate non-social control conditions, it is important to disentangle which manipulations in particular are the most important in making an information source into the primary source of learning. Here, temporal order, salience and reward feedback were altered. Humans can learn from and integrate multiple sources of information to guide behaviour [211],[217],[361],[362]. While it is likely that all manipulations of the task have contributed to increasing the salience of the social information, linking the reward feedback directly to the social information source may have made the greatest contribution. Reward has been found to influence the amount of attention allocated to stimuli during decision-making tasks [353],[354],[363] and the source of information that results in the highest reward is thought to inform learning to the greatest extent [364]. Therefore, if there are reward-related consequences linked to a particular stimulus (for example, choosing the blue option and gaining feedback on whether blue was correct or incorrect), the stimulus becomes more salient and attracts more attention through top-down processes [365],[366]. Future studies could test if linking the reward feedback directly to the social information is sufficient to make the social information into the primary source, when learning from multiple sources of information.

The adapted task described here has an advantage over previous work, in that it provides a way to orthogonalize the social/individual and primary/secondary nature of information sources in the same paradigm. For example, in previous work, where the social information was the only, and therefore the primary source of information, social learning was found to activate the same brain areas as in individual learning, namely dopamine-rich areas of the striatum [221]–[223],[225]. In contrast, other work has shown dissociations between social and individual learning. However, here the social information is typically an additional, indirect source of information during learning [93],[95],[258],[309]. Thus, this paradigm could be used to resolve contradictory findings in the literature.

In the following chapter, we test the prediction that, using our adapted version of the task, ('social-primary' task, with social information comprising the primary, and individual information the secondary source), as well as the unaltered task ('individual-primary' task), pharmacological manipulation of dopamine signalling will affect learning only from the primary source of information. For example, in the social-primary task, a significant effect of dopaminergic manipulation will be observed on social, but not individual, learning, with the opposite true for the individual-primary task. A between-subjects design will be utilised, allowing the social nature of the primary information source to vary across task groups, with participants randomly assigned to a task group. Groups will, however, be matched on demographic measures, a key limitation in the current pilot study, as, due to time constraints, results were compared across samples from two separate experiments, and were not matched on demographic variables, including gender. Participants will complete the task twice; once under placebo and one under a dopamine antagonist. We predict a decrease in learning from the primary information, with no significant effect on secondary, inferred-value learning. This pattern of results would be consistent with a domain-general view of social learning at a neurochemical level, with regard to the dopaminergic system. They would also add support for research showing that, rather than being specialised for social processes, brain networks are adaptable, and can process both social and individual information, flexibly switching between them depending on which is primary or more relevant for the current task [265],[266],[367],[368].

To conclude, this chapter describes the underlying rationale and preliminary pilot data from the development of a new task version of a widely used social learning task. Pilot testing results supported the underlying aim; to make the social information into a more primary source of information during learning. This adapted version of the SLT has the potential to allow the orthogonalization of primary/secondary and social/individual information simultaneously, and manipulations described here could be utilised in other paradigms comparing social and individual learning. Future work will involve replication of these findings in a sample with matched task groups, and an investigation of the effects of pharmacological manipulation of dopamine signalling on learning from social/individual and primary/secondary information.

Chapter 6: Dopaminergic challenge dissociates learning from primary versus secondary sources of information

This chapter presents a published study investigating manipulation of dopamine signalling on learning from social and individual information, employing a design whereby the social nature and the status of information sources are orthogonalized.

Supplementary materials for this chapter are in Appendix 5.

Publication 2:

Dopaminergic challenge dissociates learning from primary versus secondary sources of information

Alicia J. Rybicki, Sophie L. Sowden, Bianca A. Schuster, and Jennifer L. Cook

eLife 2022;11:e74893 DOI: 10.7554/eLife.74893

Summary

Some theories of human cultural evolution posit that humans have social-specific learning mechanisms that are adaptive specialisations moulded by natural selection to cope with the pressures of group living. However, the existence of neurochemical pathways that are specialised for learning from social information and from individual experience is widely debated. Cognitive neuroscientific studies present mixed evidence for social-specific learning mechanisms: some studies find dissociable neural correlates for social and individual learning whereas others find the same brain areas and, dopamine-mediated, computations involved in both. Here we demonstrate that, like individual learning, social learning is modulated by the dopamine D2 receptor antagonist haloperidol when social information is the primary learning source, but not when it comprises a secondary, additional element. Two groups (total N = 43) completed a decision-making task which required primary learning, from own experience, and secondary learning from an additional source. For one group the primary source was social, and secondary was individual; for the other group this was reversed. Haloperidol affected primary learning irrespective of social/individual nature, with no effect on learning from the secondary source. Thus, we illustrate that dopaminergic mechanisms underpinning learning can be dissociated along a primary-secondary but not a social-individual axis. These results resolve conflict in the literature and support an expanding field showing that, rather than being specialised for particular inputs, neurochemical pathways in the human brain can process both social and non-social cues and arbitrate between the two depending upon which cue is primarily relevant for the task at hand.

Keywords

social learning, dopamine, reward learning, reinforcement learning, haloperidol

Introduction

The complexity and sophistication of human learning is increasingly appreciated. Enduring theoretical models illustrate that learners utilise "prediction errors" to refine their predictions of future states (e.g. Rescorla-Wagner and temporal difference models) [24],[25],[211],[217]. An explosion of studies, however, illustrates that this simple mechanism lies at the heart of more complex and sophisticated systems that enable humans (and other species) to learn from, keep track of the utility of, and integrate information from, multiple learning sources [347],[359],[369] meaning that one can learn from many sources of information simultaneously [364]. Such complexity enables individuals to, for example, rank colleagues according to the utility of their advice and learn primarily from the top-ranked individual [190],[195],[197],[331] whilst also tracking the evolving utility of advice from others [95],[227]. Recent studies have further revealed that learning need not rely solely on directly experienced associations, since one can also learn via inference [370]–[376]. This growing appreciation of the complexity and sophistication of human learning may help to explain contradictory findings in various fields. Here we focus on the field of social learning.

The existence in the human brain of neural and/or neurochemical pathways that are specialised for learning from social information and from individual experience respectively is the topic of much debate [186],[201]. Indeed, the claim that humans have *social-specific learning mechanisms* that are adaptive specialisations moulded by natural selection to cope with the pressures of group living, lies at the heart of some theories of cultural evolution [190],[197],[199]. Since cultural evolution is argued to be specific to humans [194], establishing whether humans do indeed possess social-specific learning mechanisms has attracted many scholars with its promise of elucidating the key ingredient that "makes us human".

Cognitive neuroscience offers tools that are ideally suited to investigating whether the mechanisms underpinning social learning (learning from others), do indeed differ from those that govern learning from one's individual experience (individual learning). Cognitive neuroscientific studies, however, present mixed evidence for *social-specific* learning mechanisms. Some studies find dissociable neural correlates for social and individual learning [95],[256]–[258]. For example, a study by Behrens and colleagues [95] reported that whilst individual learning was associated with activity in dopamine-rich regions such as the striatum that are classically associated with reinforcement learning, social learning was associated with activity in a dissociable network that instead included the anterior cingulate cortex gyrus

(ACCg) and temporoparietal junction. Further supporting this dissociation, studies have revealed correlations between personality traits, such as social dominance [309] and dimensions of psychopathy [377] and social, but not individual, learning; as well as atypical social, but not individual, prediction error-related signals in the ACCg in autistic individuals [187]. Together these studies support the existence of *social-specific* learning mechanisms. In contrast, other studies have reported that the same computations, based on the calculation of prediction error, are involved in both social and individual learning [230], and that social learning is associated with activity in dopamine-rich brain regions typically linked to individual learning [221]–[225],[347]. Diaconescu and colleagues [222], for example, observed that social learning-related prediction errors covaried with naturally occurring genetic variation that affected the function of the dopamine system. Further supporting this overlap between social and individual learning, behavioural studies have observed that social and individual learning are subject to the same contextual influences. For example, Tarantola and colleagues [378] observed that prior preferences bias social learning, just as they do individual learning. Such findings promote the view that 'domain-general' learning mechanisms underpin social learning: we learn from other people in the same way that we learn from any other stimulus in our environment [186],[201]. That is, there are no *social-specific* learning mechanisms.

One potential resolution to this conflict in the literature hinges on i) an appreciation of the complexity and sophistication of human learning systems and ii) a difference in study design between tasks that have, and have not, found evidence of *social-specific mechanisms*. In studies, that have linked social learning with the dopamine-rich circuitry typically associated with individual learning (and which are therefore consistent with the domain-general view), participants have been encouraged to learn *primarily* from social information. Indeed, in many cases the social source has been the sole information source [222],[223],[225]. For example, in the paradigm employed by Diaconescu and colleagues [222],[230], participants were required to choose between a blue and green stimulus and were provided with social advice which was sometimes valid and sometimes misleading; on each trial, participants received information about the time-varying probability of reward associated with the blue and green stimuli, thus participants did not have to rely on their own individual experience of blue/green reward associations and could fully dedicate themselves to social learning. That is, participants did not learn from multiple sources (i.e., social information *and* individual experience); participants *only* engaged in social learning. In contrast, in studies where social learning has been associated with neural correlates outside of the dopamine-rich regions classically linked

to individual learning (and which are therefore consistent with the domain specific view), social information has typically comprised a secondary, additional source [95],[309]. Typically, the non-social (individual) information is presented first to participants, represented in a highly salient form, and is directly related to the feedback information. The social information, in contrast, is presented second, is typically less salient in form, and is not directly related to the feedback information. For example, in the Behrens et al. study [95] (and in our own work employing this paradigm [93],[309]) participants were required to choose between two, highly salient, blue and green boxes to accumulate points. The boxes were the first stimuli that participants saw on each trial. Outcome information came in the form of a blue or green indicator thus *primarily* informing participants about whether they had made the correct choice on the current trial (i.e., if the outcome indicator was blue, then the blue box was correct). In addition, each trial also featured a thin red frame, which represented social information, surrounding one of the two boxes. The red frame was the second stimulus that participants saw on each trial and indirectly informed participants about the veracity of the frame: if the outcome was blue AND the frame surrounded the blue box, then the frame was correct. In such paradigms, participants must learn from multiple sources of information with one source taking primary status over the other. Consequently, in studies that have successfully dissociated social and individual learning the two forms of learning differ both in terms of social nature (social or non-social) and rank (primary versus secondary status). Thus, it is unclear which of these two factors accounts for the dissociation.

The current study tests whether social and individual learning share common neurochemical mechanisms when they are matched in terms of (primary versus secondary) status. Given its acclaimed role in learning [71],[379], we focus specifically on the role of the neuromodulator dopamine. Drawing upon recent studies illustrating the complexity and sophistication of human learning [361],[375],[380] we hypothesise that pharmacological modulation of the human dopamine system will dissociate learning from two sources of information along a primary versus secondary, but not along a social versus individual axis. In other words, we hypothesise that social learning relies upon the dopamine-rich mechanisms that also underpin individual learning when social information is the primary source, but not when it comprises a secondary, additional element. Such a finding would offer a potential resolution to the aforementioned debate concerning the existence of *social-specific* learning mechanisms.

Preliminary support for our hypothesis comes from three lines of work. First, studies have convincingly argued for flexibility within learning systems. For example, in a study by Daw and colleagues [364], participants tracked the utility of four uncorrelated bandits, with particular brain regions - such as the ventromedial prefrontal cortex - consistently representing the value of the top-ranked bandit, even though the identity of this bandit changed over time. Second, studies are increasingly illustrating the flexibility of social brain networks [367],[368]. The medial prefrontal cortex (mPFC), for example, is not - as was once thought - specialised for representing the self; if the concept of 'other' is primarily relevant for the task at hand, then the mPFC will prioritise representation of other over self [265],[266]. Finally, in a recent study [93], we provided preliminary evidence of a catecholaminergic (i.e. dopaminergic and noradrenergic) dissociation between learning from primary and secondary, but not social and individual, sources of information. In this work (Cook et al., 2019) we employed a between-groups design, wherein both groups completed a version of the social learning task adapted from Behrens and colleagues (2008; described above). For one group the secondary source was social in nature (social group). For the non-social group, the secondary source comprised a system of rigged roulette wheels and was thus non-social in nature. We observed that, in comparison to placebo, the catecholaminergic transporter blocker methylphenidate only affected learning from the primary source - which, in this paradigm, always comprised participant's own individual experience. Methylphenidate did not affect learning from the secondary source, irrespective of its social or non-social nature. That is, we found positive evidence supporting a dissociation between primary and secondary learning but no evidence to support a distinction between learning from social and non-social sources. Nevertheless, since we did not observe an effect of methylphenidate on learning from the (social or non-social) secondary source of information this study was unable to provide positive evidence of shared mechanisms for learning from social and non-social sources. If it is truly the case that domain-general (neurochemical) mechanisms underpin social learning, it should follow that pharmacological manipulations that affect individual learning when individual information is the primary source also affect social learning when social information is the primary source.

The current (pre-registered) experiment tested this hypothesis by orthogonalizing social versus individual and primary versus secondary learning. We perturbed learning using the dopamine D2 receptor antagonist haloperidol, in a double-blind, counter-balanced, placebo-controlled design. To test whether pharmacological manipulation of dopamine dissociates

learning along a primary-secondary and/or a social-individual axis, we developed a novel between-groups manipulation wherein one group of participants learned primarily from social information and could supplement this learning with their own individual experience, and a second group learned primarily from individual experience and could supplement this learning with socially learned information. To foreshadow our results, we demonstrate that haloperidol specifically affects learning from the primary (not secondary) source of information. Bayesian statistics confirmed that the effects of haloperidol were comparable between the groups thus, haloperidol affected individual learning when individual information was the primary source and, to the same extent, social learning when social information was the primary source. Our data support an expanding field showing that, rather than being fixedly specialised for particular inputs, neurochemical pathways in the human brain can process both social and non-social cues and arbitrate between the two depending upon which cue is primarily relevant for the task at hand [265],[266],[367].

Results

Participants (n = 43; aged 19-38, mean (standard error) $\bar{x}(\sigma_{\bar{x}})$ = 25.950 (0.970); 24 males, 19 females; see Methods) completed an adapted version of the behavioural task originally developed by Behrens and colleagues [95]. Participants were randomly allocated to one of two groups. Participants in the individual-primary group (n = 21) completed the classic version of this task (Figure 1A [95]) in which they were required to make a choice between a blue and green box in order to win points. A red frame (the social information), which represented the most popular choice made by a group of four participants who had completed the task previously, surrounded either the blue or green box on each trial and participants could use this to help guide their choice. The actual probability of reward associated with the blue and green boxes and the probability that the red frame surrounded the correct box varied according to uncorrelated pseudo-randomised schedules (Appendix 2 – Fig. 1). For the individual-primary group, the individual information (blue and green stimuli) was primary, and the social information (red stimulus) was secondary on the basis that the blue/green stimuli appeared first on the screen, were highly salient (large boxes versus a thin frame) and were directly related to the feedback information. That is, after making their selection, participants saw a small blue or green box which *primarily* informed them whether a blue or

green choice had been rewarded on the current trial. From this information the participant could, *secondarily*, infer whether the social information (red frame) was correct or incorrect.

Our social-primary group (n = 22; groups matched on age, gender, body mass index (BMI) and verbal working memory span (Table 1)) completed an adapted version of this task (Figure 1B) wherein the social information (red stimulus) was primary, and the individual information (blue/green stimuli) was secondary. Participants first saw two placeholders; one empty and one containing a red box which indicated the social information. Subsequently, a thin green and a thin blue frame appeared around each placeholder. Participants were told that the red box represented the group's choice. They were then required to choose whether to go with the social group (red box) or not. After making their choice a tick or cross appeared which *primarily* informed participants whether going with the social information was the correct option. From this they could, *secondarily*, infer whether the blue or green frame was correct. Consequently, for the social-primary group the social information was primary on the basis that it appeared first on the screen, was highly salient (a large red box versus thin green/blue frames) and was directly related to the feedback information.

Participants in both the individual-primary and social-primary groups performed 120 trials of the task on each of two separate study days. To perturb learning, on one day participants took 2.5mg of haloperidol (HAL), previously shown to affect learning [214] via multiple routes including perturbation of phasic dopamine signalling [211],[379] facilitated by action at mesolimbic D2 receptors [381]–[383]. On the other day, they took a placebo (PLA) under double-blind conditions, with the order of the days counterbalanced. 43 participants took part in at least one study day, 33 participants completed both study days. Two participants performed at below chance level accuracy and were excluded from further analysis. We present an analysis of data from the 31 participants who completed both study days with above chance accuracy (Table 1) in the main text of this manuscript, which we complement with a full analysis of all 41 datasets in Appendix 4i.

Table 6-1 Participant information

	Individual-primary group (n = 15) Mean (SD)	Social-primary group (n = 16) Mean (SD)	t (1,29)	X^2 (1, N = 31)	p
Gender (*n* males: *n* females)	7:8	8:8		0.034	0.853
Age	25.600 (5.448)	25.625 (4.745)	0.014		0.989
VWM	80.333 (6.016)	76.354 (7.823)	1.580		0.125
BMI	24.016 (2.807)	22.625 (2.606)	1.431		0.114

Note: SD refers to standard deviation, VWM refers to verbal working memory span, BMI refers to body mass index. Age, gender, BMI and VWM did not significantly differ between the groups.

We used the following strategy to analyse our data. First, we sought to validate our manipulation by testing (under PLA) whether participants in both the individual-primary and social-primary groups learned in a more optimal fashion from the primary, versus secondary, source of information. Next, we tested our primary hypothesis that both social and individual learning would be modulated by haloperidol when they are the *primary* source of learning, but not when they comprise the *secondary* source. To do so we estimated learning rates for primary and secondary sources of information, for each group (social-primary, individual-primary), under HAL and PLA, by fitting an adapted Rescorla-Wagner learning model to choice data. To ascertain that our model accurately described choices we used simulations and parameter recovery. We used random-effects Bayesian model selection to compare our

model with alternative models. These analyses provided confidence that our model accurately described participants' behaviour. After testing our primary hypothesis, we explored the relationship between parameters from our computational model and performance. To accomplish this, we first used an optimal learner model, with the same architecture and priors as our adapted Rescorla-Wagner model, to assess the extent to which haloperidol made participants' learning rates more (or less) optimal. Finally, we regressed estimated model parameters against accuracy to gain insight into the extent to which variation in these parameters (and the effect of the drug thereupon) contributed to correct responses on the task.

Figure 6.1 Behavioural task

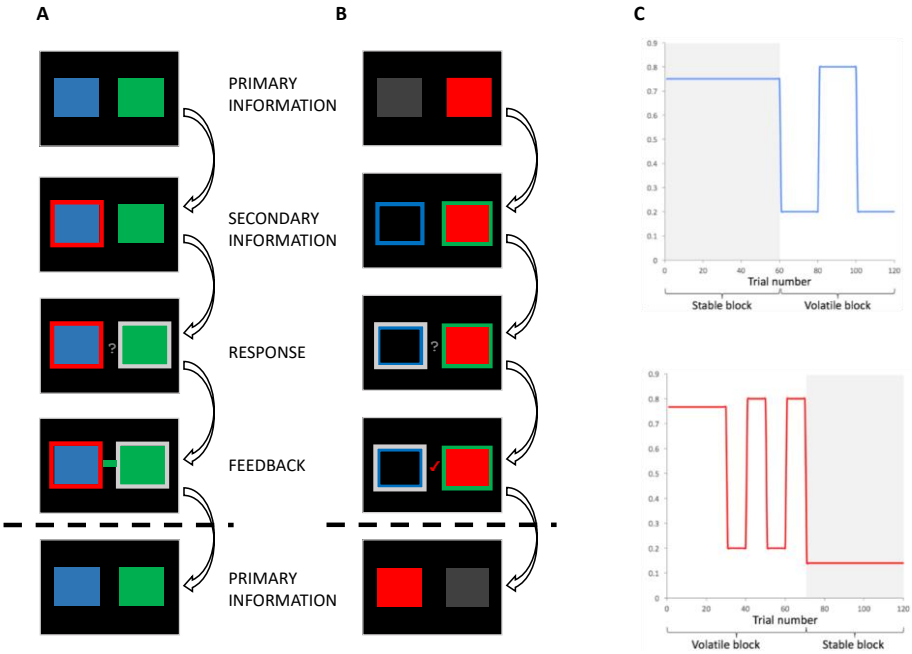


Figure 1. Behavioural task. A**. Individual-primary group**. Participants selected between a blue and a green box to gain points. On each trial, the blue and green boxes were presented first. After 1-4 seconds (s), one of the boxes was highlighted with a red frame, representing the social information. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a green or blue box in the middle of the screen. **B. Social-primary group.** Participants selected between going with, or against a red box, which represented the social information. On each trial, the red box was displayed. After 1-4s, blue and green frames appeared. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a tick or a cross. This feedback informed participants if going with the group was correct or incorrect, from this feedback participants could infer

whether the blue or green frame was correct. **C. Example of pseudo-randomised probabilistic schedule.** The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of the blue box/frame being correct (top) and the probability of the red (social) box/frame being correct (bottom).

Social information is the primary source of learning for participants in the social-primary group

Our novel manipulation orthogonalized primary versus secondary and social versus individual learning. To validate our manipulation, we tested whether participants in both the individual-primary and social-primary group learned in a more optimal fashion from the primary versus secondary source of information in our placebo condition. For this validation analysis we used a Bayesian learner model to create two optimal models (1) an optimal primary learner, and (2) an optimal secondary learner (Methods). Subsequently we regressed both models against participants' choice data, resulting in two $\beta_{optimal}$ values capturing the extent to which a participant made choices according to the optimal primary, and optimal secondary learner models respectively. $\beta_{optimal}$ values were submitted to a repeated-measures ANOVA with factors information source (primary, secondary) and group (social-primary, individual-primary), revealing main effects of information source ($F(1,29) = 6.594$, $p = 0.016$) and group ($F(1,29) = 10.423$, $p = 0.003$). $\beta_{optimal}$ values (averaged across individual-primary and social-primary groups) were significantly higher for the primary information ($\bar{x}(\sigma_{\bar{x}}) = 0.872\ (0.101)$), compared with secondary information source ($\bar{x}(\sigma_{\bar{x}}) = 0.438\ (0.101)$; $t(29) = 2.568$, $p_{holm} = 0.016$). $\beta_{optimal}$ values (averaged across primary and secondary conditions) were significantly higher for the social-primary group ($\bar{x}(\sigma_{\bar{x}}) = 0.833\ (0.078)$), compared with the individual-primary group ($\bar{x}(\sigma_{\bar{x}}) = 0.477\ (0.078)$; $t(29) = 3.228$, $p_{holm} = 0.003$) (Figure 2). Crucially, we did not observe a significant interaction between information and group ($F(1,29) = 0.067$, $p = 0.797$), meaning that participants' choices were more influenced by the primary information source, regardless of whether it was social or individual in nature. Furthermore, $\beta_{optimal}$ values for primary information alone did not significantly differ between groups ($t(29) = -1.982$, $p_{holm} = 0.257$). Note that, $\beta_{optimal}$ weights for both information sources were significantly greater than zero (primary: $t(30) = 7.534$, $p < 0.001$; secondary: $t(30) = 4.789$, $p < 0.001$) thus our optimal models of information use explained a significant amount of variance in the use of both primary and secondary learning sources. These data show that, irrespective of social (or individual) nature, participants

learned in a more optimal fashion from the primary (relative to secondary) learning source, which was first in the temporal order of events, highly salient and directly related to the reward feedback.

Figure 6.2 Beta weights (βoptimal)

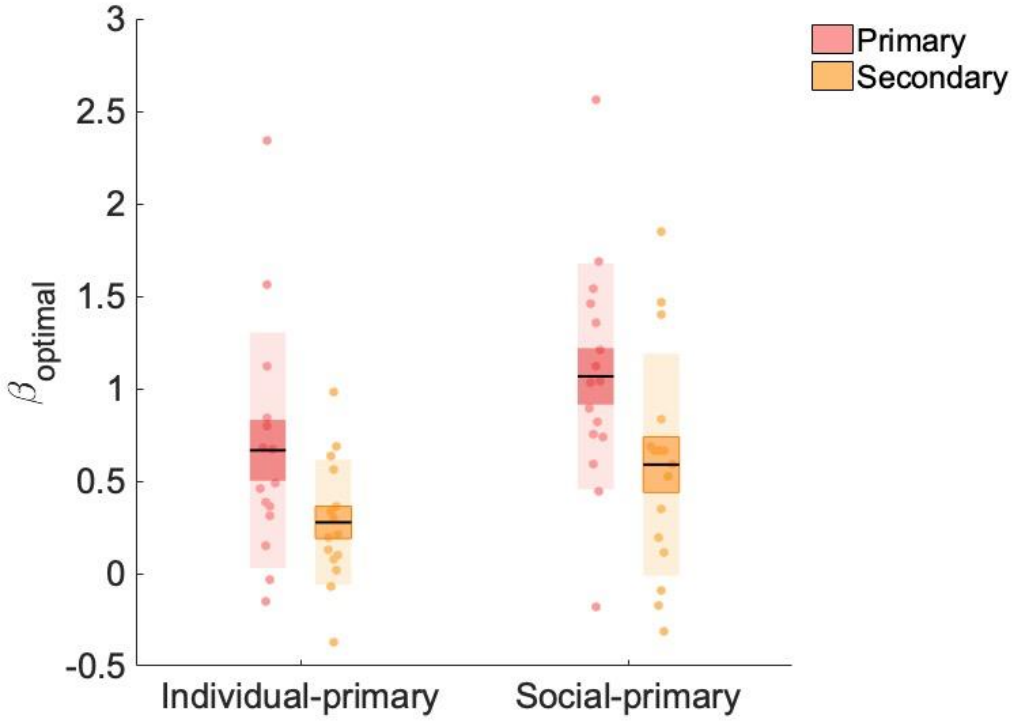


Figure 2. Beta weights (β_optimal) for primary and secondary information. $\beta_{optimal}$ values were significantly higher for the primary, compared to secondary, information source and for the social-primary, compared with the individual-primary, group. Data points indicate estimated $\beta_{optimal}$ weights for individual participants (n = 31, placebo data only), bold point indicates the mean, bold line indicates standard error of the mean (1 SEM).

Haloperidol reduces the rate of learning from primary sources

We hypothesed that both social and individual learning would be modulated by administration of the dopamine D2 receptor antagonist haloperidol when they were the *primary* source of learning, but not when they comprised the *secondary* source. To test this hypothesis we fitted an adapted Rescorla-Wagner (RW) learning model [24] to participants' choice data, enabling us to estimate various parameters that index learning from primary and secondary sources of information, for HAL and PLA conditions, for participants in the social-

primary and individual-primary groups. Our adapted RW model provided estimates, for each participant, of $\alpha, \beta,$ and ζ. The learning rate (α) controls the weighting of prediction errors on each trial. A high α favours recent over (outdated) historical outcomes, while a low α suggests a more equal weighting of recent and more distant trials. Since our pseudo-random schedules included stable phases (where the reward probability associated with a particular option was constant for > 30 trials), and volatile phases (where reward probabilities changed every 10-20 trials), α was estimated separately for volatile and stable phases (for both primary and secondary learning) to accord with previous research [88],[93],[125]. β captures the extent to which learned probabilities determine choice, with a larger β meaning that choices are more deterministic with regard to the learned probabilities. ζ represents the relative weighting of primary and secondary sources of information, with higher values indicating a bias towards the over-weighting of secondary relative to primary (see Methods and Appendix 3 for further details of the model, model fitting and model comparison).

We hypothesised an interaction between drug and (primary versus secondary) information source such that haloperidol would affect learning from the primary information source only, regardless of its social/individual nature. To test this hypothesis, we employed a linear mixed effects model with fixed factors information source (primary, secondary), drug (HAL, PLA), environmental volatility (volatile, stable) and group (social-primary, individual-primary) and dependent variable α (square-root transformed to meet assumptions of normality). We controlled for inter-individual differences by including random intercepts for subject. Including pseudo-randomisation schedule as a factor in all analyses did not change the pattern of results. The mixed model revealed a drug by information interaction (F (1, 203) = 6.852, p = 0.009, beta estimate ($\sigma_{\bar{x}}$) = 0.026 (0.010), t = 2.62, confidence interval [CI] = [0.010 – 0.050]) (Figure 3). There were no significant main effects of drug (F (1, 203) = 0.074, p = 0.786), group (F (1, 29) = 3.148, p = 0.087) or volatility (F (1, 203) = 1.470, p = 0.227) on α values, nor any other significant interactions involving drug (all p-values > 0.05, see Appendix 4v-vi for analysis including schedule, session and working memory). Planned contrasts showed that, whilst under PLA, α_{primary} ($\bar{x}(\sigma_{\bar{x}})$ = 0.451 (0.025), collapsed across volatility and group) was significantly greater than $\alpha_{\text{secondary}}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.370 (0.025); z(30) = 2.861, p = 0.004), this was not the case under HAL (α_{primary} $\bar{x}(\sigma_{\bar{x}})$ = 0.393 (0.025), $\alpha_{\text{secondary}}$ $\bar{x}(\sigma_{\bar{x}})$ = 0.417(0.025); z(30) = -0.843, p = 0.400). Furthermore, α_{primary} was decreased under HAL relative to PLA (z (30) = -2.050, p = 0.040). Although $\alpha_{\text{secondary}}$ was, in contrast, numerically increased under HAL ($\bar{x}(\sigma_{\bar{x}})$ = 0.417 (0.025) relative to PLA ($\bar{x}(\sigma_{\bar{x}})$ = 0.370

(0.025), this difference was not significant (z (30) = 1.654, p = 0.098). This drug x information interaction therefore illustrated that whilst haloperidol significantly reduced $\alpha_{primary}$ it had no significant effect on $\alpha_{secondary}$. Furthermore, under placebo there was a significant difference between $\alpha_{primary}$ and $\alpha_{secondary}$, which was nullified by haloperidol administration. Consequently, under placebo participants' rate of learning was typically higher for learning from the primary relative to the secondary source, however, under the D2 receptor antagonist haloperidol the rate of learning from the primary source was reduced and thus there was no significant difference in the rate of learning from primary and secondary sources.

Figure 6.3 Learning rate estimates

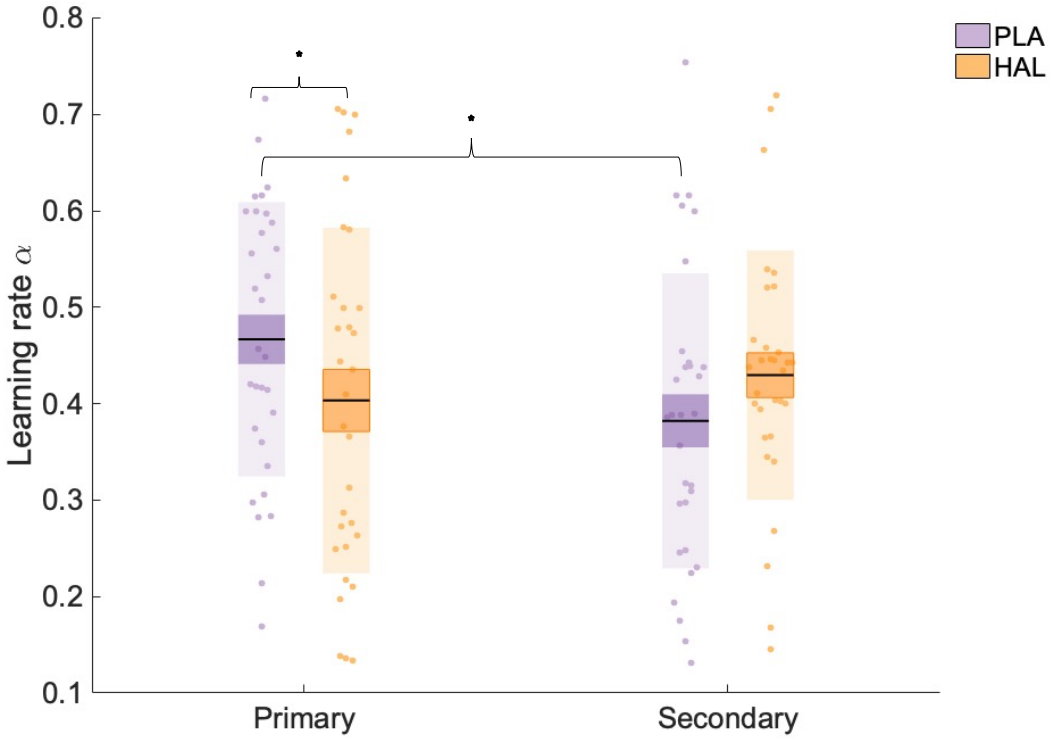


Figure 3. Learning rate (α) estimates for learning from primary and secondary information across all trials (averaged across volatile and stable phases). There was a significant interaction between information and drug, with α estimates significantly lower under haloperidol (orange), relative to placebo (purple), for primary information only. Data points indicate square-root transformed α estimates for individual participants (n = 31), boxes = standard error of the mean, shaded region = standard deviation, HAL = haloperidol, PLA = placebo, * indicates statistical significance (p < 0.05).

Linear mixed models, with fixed factors group and drug, and random intercepts for subject, were also used to explore drug effects on ζ values (representing the relative weighting of primary/secondary information) and β values. For ζ there were no significant main effects of drug (F (1, 29) = 1.941, p = 0.174, $\sigma_{\bar{x}}$ = -0.07 (0.050), t = -1.390, CI = [-0.170 – 0.003]) or group (F (1, 51) = 0.184, p = 0.669, $\sigma_{\bar{x}}$ = 0.020 (0.040), t = 0.430, CI = [-0.070 – 0.100]), nor drug by group interaction (F (1, 29) = 0.039, p = 0.845, $\sigma_{\bar{x}}$ = -0.001 (0.050), t = -0.200, CI = [-0.110 – 0.090]). Similarly, our analysis of β values revealed no main/interaction effect(s) of drug, group, or drug by group (all p > 0.05).

Haloperidol reduces the rate of learning from a primary source irrespective of its social or individual nature

Our primary hypothesis was that haloperidol would modulate the rate of learning from the primary source irrespective of its social or individual nature. This would be evidenced as an interaction between drug and (primary versus secondary) information source (see above) in the absence of an interaction between drug, information source and group (social-primary versus individual-primary). Crucially, we observed no significant interaction between drug, information source and group (F (1, 203) = 0.098, p = 0.754). To further assess whether drug effects on primary information differed as a function of group, results were also analysed within a Bayesian framework, using JASP software (JASP Team (2020)). A Bayes exclusion factor (BF_{excl}), representing the relative likelihood that a model without a drug x information x group interaction effect could best explain the observed data, was calculated [384]. Values of 3–10 are taken as moderate evidence in favour of the null hypotheses that there is no drug x information x group interaction [385] with values greater than 10 indicating strong evidence. The BF_{excl} value was equal to 7.516, providing moderate evidence in favour of the null hypothesis that there is no drug x information x group interaction. Consequently, results confirmed our hypothesis: haloperidol perturbed learning from the primary but not the secondary source, irrespective of social or individual nature.

Haloperidol brings $\alpha_{primary}$ estimates within the optimal range

To assess whether the effects of haloperidol on $\alpha_{primary}$ are harmful or beneficial with respect to performance we first explored drug effects on accuracy (see Appendix 4ii for a detailed analysis including randomisation schedule). There was no significant difference in accuracy

between haloperidol ($\bar{x}(\sigma_{\bar{x}})$ = 0.600 (0.013)), and placebo ($\bar{x}(\sigma_{\bar{x}})$ = 0.611 (0.010); F (1,29) = 0.904, p = 0.349, η_p^2 = 0.030) conditions.

The lack of a significant main effect of drug on accuracy was somewhat surprising given the significant (interaction) effect on learning rates, i.e., a decrease in $\alpha_{primary}$ under haloperidol relative to placebo. To investigate whether haloperidol resulted in learning rates that were less, or alternatively, more, optimal we compared our estimated α values with optimal α estimates. Since trial-wise outcomes were identical to those utilised by Cook et al. [93], optimal values are also identical and are described here for completeness. An optimal learner model, with the same architecture and priors as the model employed in the current task, was fit to 100 synthetic datasets, resulting in average optimal learning rates: $\alpha_{optimal\_primary\_stable}$ = 0.16, $\alpha_{optimal\_primary\_volatile}$ = 0.21, $\alpha_{optimal\_secondary\_stable}$ = 0.17, $\alpha_{optimal\_secondary\_volatile}$ = 0.19. Scores representing the difference between (untransformed) α estimates and optimal α scores were calculated ($\alpha_{diff} = \alpha - \alpha_{optimal}$). A linear mixed model analysis on α_{diff} values with factors group, drug, volatility and information source, and random intercepts for subject was conducted. A significant interaction between drug and information source was observed (F (1, 203) = 4.895, p = 0.028, $\sigma_{\bar{x}}$= 0.019 (0.010), t = 2.212, CI = [0.000 – 0.040]) (Figure 4). Planned contrasts showed that, for primary information, $\alpha_{diff_{primary}}$ was higher under PLA ($\bar{x}(\sigma_{\bar{x}})$ = 0.052 (0.023)) compared with HAL ($\bar{x}(\sigma_{\bar{x}})$ = 0.009 (0.028)); z(30) = 1.806, p = 0.071). In contrast, $\alpha_{diff_{secondary}}$ was lower under PLA ($\bar{x}(\sigma_{\bar{x}})$ = -0.011 (0.023)) compared with HAL ($\bar{x}(\sigma_{\bar{x}})$ = 0.021 (0.021)); z(30) = 1.323, p = 0.186). Learning rates for learning from the primary source were higher than optimal under placebo, with $\alpha_{diff_{primary}}$ significantly differing from 0 (one-sample t test; t(30) = 2.259, p = 0.031). Haloperidol reduced learning rates that corresponded to learning from the primary source, thus bringing them within the optimal range, with $\alpha_{diff_{primary}}$ not significantly differing from 0 under haloperidol (one-sample t test; t(30) = 0.319, p = 0.752). Consequently, under haloperidol relative to placebo, learning rates for learning from primary sources were *more optimal*.

Figure 6.4 Learning rate estimates compared with optimal learning rates

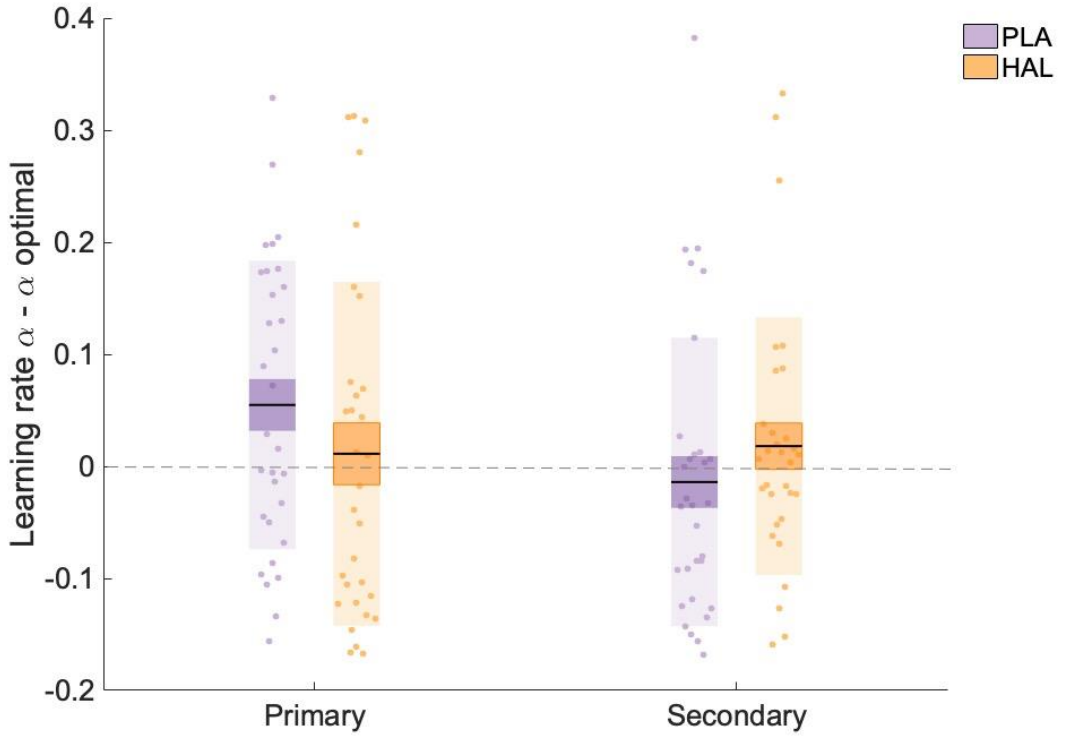


Figure 4. Learning rate estimates minus optimal learning rates. There was a significant interaction between information and drug, with $\alpha_{primary}$ scores significantly higher than optimal estimates under placebo but not under haloperidol. Data points indicate $\alpha - \alpha_{optimal}$ values for individual participants (n = 31) across all trials (averaged across volatile and stable phases), boxes = standard error of the mean, shaded region = standard deviation, HAL = haloperidol, PLA = placebo.

To explore whether α values were in some way related to accuracy scores we used two separate backwards regression models, for PLA and HAL conditions separately, with $\alpha_{primary}$ and $\alpha_{secondary}$ as predictors and accuracy as the dependent variable (see Appendix 4iii for details of a regression model with *all* model parameters). PLA accuracy was predicted by $\alpha_{secondary}$ though this model only approached significance (R = 0.121, F (1,29) = 3.981, p = 0.055). Under HAL however, accuracy was predicted by a model with $\alpha_{secondary}$ and $\alpha_{primary}$ (R = 0.450, F (2,28) = 3.560, p = 0.042), with $\alpha_{primary}$ a significant positive predictor of accuracy (β = 0.404, p = 0.028). Removing $\alpha_{secondary}$ as a predictor did not significantly improve the fit of this model (R^2change = 0.014, F change (1,29) = 0.495, p = 1.000). When combined with our optimality analysis, these results suggest that under placebo $\alpha_{primary}$ was outside of the optimal range of α values and thus accuracy was primarily driven by $\alpha_{secondary}$. However,

haloperidol reduced $\alpha_{primary,}$ bringing it within the optimal range. Thus, under haloperidol accuracy was driven by both $\alpha_{primary}$ and $\alpha_{secondary.}$

In sum, relative to placebo, the dopamine D2 receptor antagonist haloperidol significantly decreased learning rates relating to learning from primary, but not secondary sources of information, likely via mediation of phasic dopaminergic signalling (see Appendix 4iv). Interestingly, learning rates for learning from the primary source were higher than optimal under placebo and haloperidol brought them within the optimal range. Consequently, both primary and secondary learning contributed to accuracy under haloperidol but not under placebo. Importantly, the effects of haloperidol did not vary as a function of group allocation, which dictated whether the primary source was of social or individual nature. A Bayesian analysis confirmed that we had moderate evidence to support the conclusion that there was no interaction between drug, learning source and group. These data, thus, illustrate a dissociation along the primary-secondary but not social-individual axis.

Discussion

The current study tested the hypothesis that social and individual learning share common neurochemical mechanisms when they are matched in terms of (primary versus secondary) status. Specifically, we predicted that haloperidol would perturb learning from the primary but not the secondary source, irrespective of social or individual nature. Supporting our hypothesis, we observed an interaction between drug and information source (social versus individual) such that under haloperidol (compared to placebo) participants exhibited reduced learning rates with respect to learning from the primary, but not the secondary, source of information. Crucially, we did not observe an interaction between drug, information source and group (social-primary versus individual-primary). Bayesian statistics revealed that, given the observed data, a model that excludes this interaction is 7.5 times more likely than models which include the interaction.

An important question concerns whether the lack of a dopaminergic dissociation between social and individual learning could be explained by participants not fully appreciating the social nature of the red shape (the social information source). In opposition to this, we argue that since our participants could not commence the task until reaching 100% accuracy in a pre-task quiz, which questioned participants about the social nature of the red shape, we can

be confident that all participants knew that the red shape indicated information from previous participants. Participants also completed a post-task questionnaire (Appendix 5), which required them to reflect upon the extent to which their decisions were influenced by the social (red shape) and individual (blue/green shapes) information. If participants had not fully believed that the red shape represented social information, one might expect that they would indicate that they were *not* influenced by this source. In contrast, participants in both the individual-primary and social-primary groups believed that they were influenced by the red shape (as well as the blue/green stimuli). Furthermore, in our previous work, using the same social manipulation, we demonstrated that the personality trait social dominance significantly predicts social, but not individual, learning [309]. Thus, illustrating that participants treat the social information differently from the non-social information in this type of paradigm. Finally, based on previous studies, we argue that even with a more overtly social manipulation it is highly likely that social learning would still be perturbed by dopaminergic modulation when social information is the primary source. Indeed, in a study by Diaconescu et al. [222] social information was represented by a video of a person indicating one of the two options. Even with this overtly social stimulus, Diaconescu and colleagues still observed that social learning covaried with genetic polymorphisms that affect the functioning of the dopamine system.

The first part of our analysis illustrated that our manipulation produced the expected effect: when social information was first in the temporal order of events, highly salient and directly related to reward feedback participants learned in a more optimal fashion from this source of information. Such a result may be a surprise to some since one might think that, relative to learning from one's own experience, learning from others will always take a "backseat". Here we clearly demonstrate that, when cast as the primary task, participants can make good use of social information. This paradigm may comprise a step towards developing a system to support accelerated social learning. Future studies could, for instance, investigate whether similar manipulations can be used to improve learning *about* (as opposed to *from*) other individuals. Since temporal order, saliency and reward feedback were manipulated simultaneously we cannot determine which manipulation is the most influential. Future work may therefore also seek to manipulate these factors independently to establish the most effective method for promoting social learning.

Our results comprise an important contribution to the debate concerning the existence of social-specific learning mechanisms. We find that, like individual learning, social learning is

modulated by a dopaminergic manipulation when it is the primary source of information. This result marries well with previous studies that have linked social learning with dopamine-rich mechanisms when the social source has been the primary (or in many cases the sole) information source [222],[223],[225]. Our results are also consistent with studies that have associated social learning with different neural correlates, outside of the dopamine-rich regions classically linked to individual learning, when it is a *secondary* source of information [95],[257],[258]. Our data suggest that social and individual learning share common dopaminergic mechanisms when they are the primary learning source and that previous dissociations between these two learning types may be more appropriately thought of as dissociations between learning from a primary and secondary source. Extant studies e.g. [93] were not able to illustrate the importance of the primary versus secondary distinction because they did not fully orthogonalize primary versus secondary and social versus individual learning.

Though our results suggest shared neurochemical mechanisms for social and individual learning when they are matched in status, it is, nevertheless, essential to highlight that it does not follow that there are *no* dimensions along which social learning may be dissociated from individual learning. It is possible that although social and individual learning are affected by dopaminergic modulation - when they are the primary source -, there are differences in the *location* of neural activity that could be revealed by neuroimaging. For instance, although social and individual learning are both associated with activity within the striatum [229],[316], social-specific activation patterns have been observed in other brain regions, including the temporoparietal junction [95],[386] and the gyrus of the anterior cingulate cortex [95],[257],[258]. Consequently, it is possible that haloperidol has comparable effects on social and individual learning but that these effects (seen at an "algorithmic level of analysis" [189]) are associated with activity in different brain regions (i.e., dissociations at an "implementation level of analysis"[189]). For example, haloperidol may comparably affect the BOLD signal associated with social and individual prediction errors, but the effect may be localised to dissociable neural pathways. Such a location-based dissociation requires further empirical investigation as well as further consideration of the possible functional significance of such location-based differences, if they are indeed present when primary versus secondary status is accounted for. Nevertheless, whilst such location-based differences are *possible*, we argue that they are not *probable* since, given different distributions of dopamine neurons, receptors and reuptake mechanisms throughout the brain [323],[383],[387],[388], differences in location are relatively likely to result in differences in the magnitude of the effect of haloperidol [389],[390]. Additionally,

since we did not observe significant effects of haloperidol on learning from social or individual sources when they were secondary in status, it remains a logical possibility that social and individual learning can be neurochemically dissociated when they are the secondary source of information - though it is admittedly difficult to conceive of a parsimonious explanation for the existence of two neurochemical mechanisms for social and individual learning *from secondary sources*. Finally, it is possible that social and individual learning share common *dopaminergic* mechanisms when they are the primary source, but differentially recruit other neurochemical systems. For instance, some have argued that social learning may heavily rely upon serotonergic mechanisms [241],[254],[255]. The abovementioned avenues should be further explored, however, in the interim, it must be concluded that since existing studies have not controlled for primary versus secondary status, we do not currently have convincing evidence that social and individual learning can be dissociated in the human brain.

Notably, our results reveal a clear dissociation between learning from primary and secondary sources. For learning from primary sources haloperidol made learning rates more optimal, haloperidol did not have this effect on learning rates for secondary learning. Interestingly, a combined optimality analysis and regression model suggested that, under placebo, learning rates for learning from the primary source were "too high" and fell outside of the optimal range (for this specific task). Consequently, under placebo, variance in accuracy was primarily explained by learning rates for learning from the secondary source. However, haloperidol reduced learning rates for learning from the primary source, bringing them within the optimal range. Thus, under haloperidol, accuracy was driven by learning rates for learning from both the primary and secondary sources. An open question concerns whether haloperidol truly *optimises*, or simply *reduces* learning rate. Since the current paradigm was not designed to test this hypothesis a reduction in learning rate herein also corresponds to an optimisation of learning rate. To dissociate the two, one would need a paradigm that generates sufficient numbers of participants with learning rates (in the placebo condition) that are *sub-optimally low* such that one can observe whether, in these critical test cases, haloperidol *increases* (i.e., optimises) learning rate.

An intriguing question concerns the synaptic mechanisms by which haloperidol affects learning rates. Non-human animal studies, have shown that phasic signalling of dopaminergic neurons in the mesolimbic pathway encodes reward prediction error signals [211],[379]. Since haloperidol has high affinity for D2 receptors [383], which are densely distributed in the

mesolimbic pathway [381],[382], dopamine antagonists including haloperidol can affect phasic dopamine signals [215] - either via binding at postsynaptic D2 receptors (which blocks the effects of phasic dopamine bursts), or via presynaptic autoreceptors (which has downstream effects on the release and reuptake of dopamine and thus modulates bursting itself) [391]–[393]. That is, haloperidol may affect learning rate via blockade of the postsynaptic D2 receptors, which may mute the effects of phasic dopamine signalling (either directly or via reduction in the background tonic rate of activity which, in turn, reduces the amplitude of phasic responses [394],[395]), thus reducing the weight of prediction error signals on value updating (i.e., reducing the learning rate). Indeed a number of studies have shown that haloperidol can attenuate prediction error-related signals [214],[216],[220],[396]. For example, in the context of individual learning, Pessiglione et al. [214] demonstrated that haloperidol attenuated prediction error signals in the striatum, indexed via changes in blood oxygen levels (BOLD). In addition to effects on postsynaptic D2 receptors, haloperidol may modulate prediction errors via effects on presynaptic autoreceptors. Autoreceptor binding is suggested to increase phasic bursting [215],[397]–[399] thus enhancing the phasic signal that is indicative of positive prediction errors. A combination of pre- and post-synaptic effects could feasibly result in more optimal learning rates wherein dopamine signalling is muted via postsynaptic blockade thus muting (tonic background) "noise" (and signal) but where the phasic "signal" is enhanced via presynaptic effects, potentially resulting in an overall increased signal-to-noise ratio which may translate into more optimal learning rates.

Perhaps the most novel contribution of our work is that we here illustrate that, whilst dopaminergic modulation affects learning from the primary source, it does not significantly affect learning from the secondary source. Previous studies have illustrated that humans can learn - ostensibly simultaneously - from multiple sources of information and tend to organise this information in a hierarchical fashion such that the source which is currently of highest value has the greatest influence on a learner's behaviour [364]. Here we extend this work by showing that the primary source, at the top of the hierarchy, is more heavily influenced by modulation of the dopamine system, thus suggesting a graded involvement of the dopamine system according to a source's status in the "learning hierarchy". Extant studies [364] suggest that such learning hierarchies are flexible and can be rapidly remodelled according to a source's current value. The success of our orthogonalization of social versus individual and primary versus secondary learning depended on a within-subjects design, wherein the status (primary or secondary) of the learning source varied only between participants. Although our

study was therefore not optimised for studying the rapid remodelling of learning hierarchies, our results pave the way for future studies to investigate whether the impact of dopaminergic modulation of learning from a particular source quickly changes according to the source's current status in the learning hierarchy.

In sum, in previous paradigms that dissociate social and individual learning, the social information comprised a secondary or additional information source, differing from individual information both in terms of its social nature (social/individual) and status (secondary/primary). We here provide evidence that dissociable effects of dopaminergic manipulation on different learning types are better explained by primary versus secondary status, than by social versus individual nature. Specifically, we showed that, relative to placebo, haloperidol reduced learning rates relating to learning from the primary, but not secondary, source of information irrespective of social versus individual nature. Results illustrate that social and individual learning share a common dependence on dopaminergic mechanisms when they are the primary learning source.

Materials and Methods

Subjects

Subjects (n = 43, aged 19 to 42 years, mean (SD) = 26 (6.3); 19 female) were recruited from the University of Birmingham and surrounding areas in Birmingham city, via posters, email lists and social media. Four participants dropped out of the study after completing the first day. A further five participants could not complete the second test day, due to university-wide closures and a restriction of data collection. In total, 43 participants completed one session, with 33 participants completing both test days. However, Bayes exclusion factors were reported for interactions of interest, to avoid the possibility of type 2 error. The study was in line with the local ethical guidelines approved by the local ethics committee (ERN_18_1588) and in accordance with the Helsinki Declaration of 1975.

General procedure

The study protocol was pre-registered (see Open Science Framework (OSF) https://osf.io/drmjb for study design and *a priori* sample size calculations). All participants attended a preliminary health screening session with a qualified clinician, followed by two test sessions with an interval of one to a maximum of four weeks between testing session. The health screening session, lasting approximately one hour, started with informed consent, followed by a medical screening. Participants were excluded from further participation if they met any of the exclusion criteria. Participants then completed a battery of validated questionnaire measures (see Appendix 1 for inclusion/exclusion criteria, questionnaire measures, medical symptoms, and mood ratings). Both test days (1-4 weeks post health screening) followed the same procedure, starting with informed consent, followed by a medical screening. Participants were then administered capsules (by a member of staff not involved in data collection) containing either 2.5 mg haloperidol (HAL) or placebo (PLA), in a double-blind, placebo-controlled, cross-over design. Participants were told to abstain from alcohol and recreational drugs in the 24 hours prior to testing and from eating in the two hours prior to capsule intake.

1.5 hours after capsule intake, participants commenced a battery of behavioural tasks, including a probabilistic learning paradigm (Go-NoGo learning [215]) and a measure of verbal working memory [400]. The social learning task was started approximately 3 hours post-capsule administration, within the peak of HAL blood plasma concentration. HAL dosage and administration times were in line with similar studies which demonstrated both behavioural and psychological effects of haloperidol [215],[282]. Both test days lasted approximately 5.5 hours in total. Blood pressure, mood and medical symptoms were monitored throughout each day: before capsule intake, three times during the task battery and after finishing the task battery. On completion of the second session, participants reported on which day they thought they had taken the active drug or placebo. Participants received monetary compensation on completion of both testing sessions, at a rate of £10 per hour, with the opportunity to add an additional £5 based on their performance during the learning task.

Behavioural task

Participants completed a modified version of a social learning task [309], first developed by Behrens and colleagues [95]. The task was programmed using MATLAB R2017b (The MathWorks, Natick, MA). Participants were randomly allocated to one of two groups. For both groups, participants completed 120 trials on both test days. The task lasted approximately 35 minutes, including instructions. Before the main task, participants completed a step-by-step on-screen practice task (10 trials) in which they learnt to choose between the two options to obtain a reward and learned that the "advice" represented by the frame(s) could help in making the correct choice in some phases. In our previous work with the individual-primary condition alone, we demonstrated that social dominance significantly predicts social, but not individual, learning [309]. Thus, showing that participants maintain a conceptual distinction between the social and individual learning sources. In the current study we investigated whether participants, maintained this conceptual distinction by requiring participants to complete a short quiz (3 questions), testing their knowledge, after the practice task. Participants were required to repeat the practice round until they achieved 100% correct score in the quiz, meaning that all participants understood the structure of the task, and that the red shape represented *social* information. Furthermore, after the experiment, participants completed a feedback questionnaire (Appendix 5). Answers confirmed that participants understood the difference between, and paid attention to both, individual and social sources

of information. Participants were informed as to whether they had earned a £5 bonus after the second session. Due to ethical considerations, all participants received the bonus.

Individual-primary group

On each trial participants were required to choose between a blue or green box to gain points. Participants could also use an additional, secondary, source of information - a red frame surrounding either the blue or green box – to help make their decision. Participants were informed (see Appendix 5 for instruction scripts) that the frame represented the most popular choice made by a group of participants who had previously completed the task. They were also informed that the task followed 'phases' wherein sometimes the blue, but at other times the green choice, was more likely to result in reward and sometimes the social information predominantly indicated the correct box, but at other times it predominantly surrounded the incorrect box (Fig.1A). After making their choice participants received outcome information in the form of a blue or green indicator. The indicator primarily informed participants about whether the blue or green box had been rewarded on the current trial. Whether the social information surrounded the correct or incorrect box could, secondarily, be inferred from the indicator. For example, if the red frame indicated that the social group had chosen the blue shape, and the blue shape was shown to be correct, participants could infer that the social information had therefore been correct on that trial. Both the probability of reward associated with the blue/green stimuli and the utility of the social information, varied according to separate probabilistic schedules, with participants randomly assigned to one of four groups (Appendix 2). For both individual and social information, the probabilistic schedules featured stable phases, where the probability of reward was constant, and volatile phases, in which the probability switched every 10-20 trials. This feature of the task design was included to capture potential effects of dopaminergic modulation on adaptation to environmental volatility [93]. Participants were informed that correct choices would be rewarded, and thus to aim to accumulate points to obtain a reward at the end of the experiment. Although probabilistic schedules for Day 2 were the same as Day 1, there was variation in the trial-by-trial outcomes and advice. In addition, to prevent participants from transferring learned stimulus-reward associations from Day 1 to Day 2, different coloured stimuli were employed on the second session: participants viewed blue/green squares with advice represented as a

red frame on Day 1 and yellow/purple squares with advice represented as a blue frame on Day 2.

Social-primary group

For the social-primary group the social information source was the primary source of learning. On each trial participants were presented with two grey placeholders. One placeholder was filled with a red box, indicating the group's choice. Blue/green frames then appeared around the placeholders. As in the individual-primary group, participants were informed that the task followed 'phases' wherein sometimes going with, but at other times going against, the group's choice was more likely to result in reward and sometimes the blue frame predominantly indicated the correct box, whereas at other times the green frame predominantly indicated the correct box. After making their choice participants received outcome information in the form of a tick/cross indicator. The indicator primarily informed participants about whether the social group had been rewarded (and thus going with them would have resulted in points scoring but going against them would not) on the current trial. Whether the blue(green) frame surrounded the correct or incorrect option could, secondarily, be inferred from the indicator. As in the individual-primary task, both the probability of reward associated with the blue/green stimuli and the utility of the social information varied according to probabilistic schedules (Appendix 2). All other aspects of the task structure were the same as previously described in the individual-primary task group.

Data analysis

All analyses were conducted using MATLAB R2017b (The MathWorks, Natick, MA) and Bayesian analyses using JASP (JASP Team (2020). JASP (Version 0.14) [Computer software]). Linear mixed models were fitted to data using RStudio (RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA). In the instance of data not meeting assumptions of normality (as assessed by Kolmogorov–Smirnov testing), data were square-root-transformed. Learning rate α values were square-root transformed (see Table II for untransformed learning rates). We used the standard $p < .05$ criteria for determining if significant effects were observed, with a Holm correction applied for

unplanned multiple comparisons, to control for type I family-wise errors. In addition, effect sizes and beta weights for linear mixed model analysis are reported.

Data pre-processing

Datasets were excluded based on the following: accuracy < 50% under placebo, chose the same side (left/right) or colour on > 80% trials, incomplete datasets (less than 120 trials completed). Two subjects were excluded, resulted in a final sample of n = 31, with behavioural data for both testing days, and n = 41, with data for one day only (see Appendix 4i for analysis).

Optimal learner model

The influence of each information source (primary and secondary) on choices was quantified by regressing two "optimal learners" against subjects' choices. The first comprised an optimal "individual learner model", which was generated by using a Bayesian learner algorithm [88] to simulate an optimal learner who learns solely from individual information (the blue and green stimuli). The second comprised a "social learner model" which simulated an optimal learner who learns solely from the social information (red stimuli). The Bayesian learner algorithm [88] describes an optimal approach to tracking reward probabilities in a changing environment. It assumes an underlying probability of an outcome being correct and tracks this probability across time, as well as maintaining an estimate of the rate of change of probabilities, i.e., volatility. All probabilities are updated in a Markovian fashion, meaning there is no requirement to store the full history of decision outcomes or statistics of the environment [88]. Thus, on each trial, the individual learner model represented the reward probability associated with a blue choice, derived through learning, in an optimal fashion, exclusively from information about reward outcomes and ignoring the social information. The social learner model represented the probability, based on the (reward-weighted) social information, that the social information was correct. From the social learner model, on each trial, the reward probability of a blue choice was calculated, that would have been derived if a participant had been learning optimally, exclusively from the social information (i.e., ignoring individual reward outcomes). Subsequently both models were regressed separately against each individual participant's choice data using binomial logistic regression, with

model predictions from the primary and secondary models as continuous predictor variables and participant response as the dependent variable (0/1). For each participant, this produced two parameter estimates, or standardised beta weights, each representing the degree to which individual experience and social information explained choices. For example, a participant whose choices were more strongly influenced by the social information than the individual information, would have a high social $\beta_{optimal}$ value, and a low individual $\beta_{optimal}$ value.

Computational modelling framework

Participant responses were modelled using an adapted Rescorla-Wagner learning model [24]. The model relies on the assumption that updates to choice behaviour are based on prediction errors, i.e., the difference between an expected and the actual outcome. Participants were assumed to update their beliefs about outcomes based on sensory feedback (perceptual model), and to use this feedback to make decisions about the next action (response model). Model fitting was performed using scripts adapted from the TAPAS toolbox [230] (scripts available at OSF link https://tinyurl.com/b3c7d2zb). A systematic comparison of eight separate models (Appendix 3 for full details regarding model fitting and model comparison) showed that the exceedance probability of this particular model was ~1. This demonstrates (relative) evidence in favour of the conclusion that, the current model, with separate learning rates for primary and secondary information, and volatile and stable phases, provided the best fit to participant choice data and that the data likely originated from the same model for both HAL and PLA treatment conditions (Appendix 3 - Fig. 1). Further model validation, including simulation of data and parameter recovery, provided further support for the choice of computational model (Appendix 3).

Perceptual model

The Rescorla-Wagner predictors used in our learning models consisted of a modified version of a simple learning model, with one free parameter, the learning rate α, varying between 0 and 1.

$$V_{(i+1)} = V_i + \alpha(r_i - V_i)$$

According to this model the predicted value (V_i) is updated on each trial based on the prediction error (PE), or the difference between the actual and the expected reward ($r_i - V_i$), weighted by the learning rate α. α thus captures the extent to which the PE updates the estimated value on the next trial. In line with previous work [93], we used an extended version of this learning model, with separate α values for volatile and stable environmental phases. In a stable environment, learning rate will optimally be low, and reward outcomes over many trials will be taken into account. In a volatile environment, however, an increased learning rate is optimal, as more recent trials are used to update choice behaviour [88]. Furthermore, we simultaneously ran two Rescorla-Wagner predictors in order to estimate parameters relating to learning from primary and secondary information sources. Consequently, our model generated the predicted value of going with the primary source (going with the blue frame for the individual-primary group, going with the group for the social-primary group; $V\_primary_{(i+1)}$) and the predicted value of the secondary information (going with the group recommendation for the individual-primary group, going with the blue frame for the social-primary group; $V\_secondary_{(i+1)}$) and provided four α estimates: $\alpha_{\text{primary\_stable}}$, $\alpha_{\text{primary\_volatile}}$, $\alpha_{\text{secondary\_stable}}$, $\alpha_{\text{secondary\_volatile}}$.

Response model

Our response model assumed that participants integrated learning from both primary and secondary sources. The action selector predicts the probability that the primary information (blue choice/ group choice) will be rewarded on a given trial and was based on the softmax function (TAPAS toolbox), adapted by Diaconescu and colleagues [230]. This response model is adapted from that used by Cook and colleagues [93] and reproduced here with permission. The value of primary and secondary information was combined using the following:

$$V_{primary(i+1)} = \zeta\left(V_{secondary_{advice_{weighted(i+1)}}}\right) + (1-\zeta)(V_{primary(i+1)})$$

wherein ζ is a parameter that varies between individuals, and which controls the weighting of secondary relative to primary sources of information. $V_{secondary\_advice\_weighted(i+1)}$ comprises the advice provided by the secondary information (the red and blue frames, for individual-primary and social-primary groups respectively) weighted by the probability of advice accuracy ($V_{secondary(i+1)}$) in the context of making a choice to go with the primary information (the blue and red box for the individual-primary and social-primary groups respectively). That is:

$$V\_secondary\_advice\_weighted(i+1) = \left|\text{advice} - V_{secondary(i+1)}\right|$$

where advice from the red frame equals 0 for blue and 1 for green, and advice from the blue frame equals 0 for going with the red box and 1 for going against the red box. For example, for a participant in the social-primary group, if the blue frame advised them to go with the red box (the group choice) and the probability of advice accuracy was estimated at 80% ($V_{secondary(i+1)}$ = 0.80), the probability that the choice to go with the group will be rewarded, inferred from secondary learning, would be 0.8 ($V_{secondary\_advice\_weighted(i+1)}$ = |0−0.8|= 0.8). The probability that this integrated belief would determine participant choice was described by a unit square sigmoid function, describing how learned belief values are translated into choices.

$$P(y_{(i+1)} = 1 \,||V_{primary(i+1)}) = \frac{{V_{primary(i+1)}}^{\beta}}{{V_{primary(i+1)}}^{\beta} + \left(1-V_{primary(i+1)}\right)^{\beta}}$$

Here, responses are coded as $y_{(i+1)} = 1$ when selecting the primary option (going with the blue and red box for the individual-primary and social-primary groups respectively), and $y_{(i+1)} = 0$ when selecting the alternative (going with the green box and going against the red box for the individual-primary and social-primary groups respectively). The participant-specific free parameter β, the inverse of the decision temperature, describes the extent to which estimated value of choices determines actual participant choice: as β decreases, decision noise increases and decisions become more stochastic; as β increases, decisions become more deterministic towards the higher value option.

Significance tests for estimated model parameters

Parameters were fitted separately for each participant's choice data. Learning rate (α) was estimated for each participant, for primary and secondary learning, for volatile and stable phases, on both test days, resulting in 8 estimated learning rates per participant. β values were also estimated for each participant on both treatment days, resulting in two β values per participant. Effects-coded mixed model linear analyses were carried out, to allow for inclusion of subject as a random factor thus ensuring that between-participant variation in α could be controlled for. Fixed factors were drug (HAL, PLA), information type (primary, secondary), volatility (volatile, stable) and group (individual primary, social-primary), with the inclusion of random intercepts for participant: ~ group x information x drug x volatility + 1| subject.

Repeated-measures analysis of variance (RM-ANOVA) for linear mixed effects models was carried out using the Satterthwaite approximation for degrees of freedom, and the model was fit using maximum likelihood estimation, with a model including random intercepts, but not random slopes, providing the best fit to the data. All analyses were repeated with and without the inclusion of age, BMI and baseline working memory as covariates, with the pattern of results unchanged. Where appropriate, data were transformed to meet assumptions of normality for parametric testing.

Table 6-2 Untransformed estimated learning rates

		$\alpha_{primary\_volatile}$	$\alpha_{primary\_stable}$	$\alpha_{secondary\_volatile}$	$\alpha_{secondary\_stable}$
PLA	$\bar{x}(\sigma_{\bar{x}})$	0.184 (0.018)	0.290 (0.041)	0.187 (0.028)	0.151 (0.025)
	range	0.024-0.477	0.027-0.721	0.011-0.591	0.004-0.612
HAL	$\bar{x}(\sigma_{\bar{x}})$	0.169 (0.029)	0.218 (0.033)	0.200 (0.023)	0.202 (0.026)
	range	0.010-0.578	0.013-0.699	0.014-0.481	0.011-0.584

Note: $\bar{x}(\sigma_{\bar{x}})$ refers to mean (standard error of the mean), PLA refers to placebo, HAL refers to haloperidol.

Bayesian statistical testing

Bayesian statistical testing was implemented as a supplement to null hypothesis significance tests, to investigate if null results represent a true lack of a difference between the groups [384], using JASP software, based on the R package "BayesFactor" [401]. The JASP framework for repeated measures ANOVA was used [402], whereby exclusion Bayes factors were obtained for predictors of interest. The exclusion Bayes factor (BF_{excl}) for a given predictor or interaction quantifies the change in odds from the prior probability that the predictor is included in the regression model, to the probability of exclusion in the model after seeing the data (BF_{excl}). Bayes factors were computed by comparing all models with a predictor against all models without that predictor, i.e., comparing models that contain the effect of interest to equivalent models stripped of the effect. For example, an exclusion Bayes factor for an effect of 3 for a given predictor i can be interpreted as stating that, models which exclude the predictor i, are 3 times more likely to describe the observed data than models which include the predictor. In short, the exclusion Bayes factor is interpreted as the evidence given the observed data for excluding a certain predictor in the model and can be used as evidence to support null results. For all Bayesian analyses, the Bayes factor quantifies the relative

evidence for one theory or model over another. We followed the classification scheme used in JASP [385] to classify the strength of evidence given by the Bayes factors, with BF_{excl} between one and three considered as weak evidence, between three and ten as moderate evidence and greater than ten as strong evidence for the alternative hypothesis respectively.

Chapter 7: General Discussion

Bayesian inference and predictive processing have been proposed as a common mechanism that is atypical in autism spectrum disorder (ASD), accounting for many diverse characteristics of the autistic phenotype. This thesis focused on examining some key predictions from these accounts in the social and motor domains, before focusing on social learning in greater detail, investigating the extent to which social learning relies on domain-general neurochemical mechanisms. There were therefore two main themes to this research; the extent to which predictive mechanisms in social and motor learning are atypical in autistic adults (Chapters 2 and 3), and the neurochemical mechanisms underpinning social (and individual) learning in neurotypical adults (Chapters 4-6). As individual findings have been discussed in each chapter, I here briefly summarise and integrate all reported results in light of previous research. I will then discuss how the empirical findings in this thesis add to the overall literature on predictive processing in autism and to our understanding of the domain-specificity of these processes. As limitations specific to each chapter have been discussed already, this chapter will provide a more general overview of the strengths and limitations associated with each theme and outline future research directions to build on these results and explore outstanding questions.

7.1 Predictive learning in autistic adults

In empirical Chapters 2 and 3, I examined whether predictive learning mechanisms were atypical in autistic adults in the context of a probabilistic social reward learning and a motor sequence learning task. Predictive accounts of autism seek to explain diverse aspects of the autistic phenotype under a common 'predictive coding impairment', proposing a domain-general impairment in predictive processes in autism [11]–[13],[75],[78]. Under this framework, atypical prediction stems from an imbalance between the relative precision of prediction errors (PEs) and higher-level predictions and atypical adjustment of precision, with atypical prediction proposed as a mechanistic explanation for behavioural atypicalities across different domains. I will discuss the extent to which the empirical data in this thesis are consistent with hypotheses from these accounts.

7.1.1 Summary of results

7.1.1.1. Chapter 2 [403] compared the performance of autistic and non-autistic adults on a probabilistic motor sequence learning task. We examined key predictions from predictive coding accounts of autism. Specifically, that a relative imbalance of precision of prior predictions relative to incoming information results in an aberrantly high baseline level of surprise in autistic individuals, leading to an atypical behavioural response to surprising stimuli. In this task, participants were required to (implicitly) learn sequences of actions, wherein, on occasional "surprising" trials, an expected action had to be replaced with an unexpected action. We predicted that autistic individuals would overweight incoming, at the expense of prior, information, resulting in a decrease in building stable expectations of upcoming events (impaired learning of motor sequences) and a subsequent atypical response to statistically surprising events (reduced surprise where surprise would be expected). In contrast to our predictions, no significant differences were observed between autistic and non-autistic participants on either measure; autistic individuals demonstrated intact motor sequence learning and a typical response to surprising events. Furthermore, no correlations were observed between autistic traits and behavioural measures.

7.1.1.2. Chapter 3 examined social and individual learning, in autistic and non-autistic adults, using a probabilistic social learning task (SLT). This paradigm allowed comparison of performance in a stable environment, where the underlying cue-outcome probabilities were stable for at least 50 trials, and in a volatile environment, where probabilities changed rapidly (every 10-20 trials), separately for social and individual information. In line with accounts of atypical prediction in autism, specifically inflexibly high precision of PEs [12],[77], we predicted atypical adaptation to volatility during learning, with adjustment to volatility indexed by the difference in learning rates between volatile and stable phases. We also investigated whether social learning would be atypical in our autistic sample, without the presence of possible confounds, such as the requirement for processing of face/biological stimuli. In contrast to predictions, our results demonstrate no differences in social or individual learning between groups, and no differences in adjustment of learning to volatility. While we did observe a small and unpredicted difference in behaviour in autistic individuals with respect to winning on the previous trial in a volatile environment, Bayesian analysis did not provide support for this difference.

7.1.2 Interpretation of results

Taken together, with respect to the question of whether autistic adults exhibit differences in predictive learning processes, the current results do not add support. That is, empirical work in this thesis showed no differences in either social or motor learning when comparing autistic and non-autistic adults. These results contradict predictions from accounts proposing altered prediction-based learning in autism [50], altered relative precision and atypical adjustment of learning to the current state of environmental volatility [13],[77],[78]. However, while these accounts have been tested empirically across different domains in recent years, findings have remained mixed, with the presence of atypicalities in autistic individuals varying across studies. For example, while many studies find evidence in support [107],[119],[271],[289], our results, providing evidence that autistic adults do not differ from non-autistic adults, are not unprecedented and are in accordance with previous research reporting no differences between autistic and non-autistic individuals [106],[110],[125],[404]–[407]. In attempting to consider how the current empirical results integrate with current literature, I will identify the circumstances under which previous research supports or opposes predictions from these theories, highlighting differences between the measures used here and previous work.

First, the current results contrast with previous work reporting atypical responses to surprising events in autism, indexed by a decrease in the difference in response to unexpected and expected cues. For example, Lawson et al. [100] reported both reduced behavioural (indexed by reaction time (RT) and error rate) and pupillometric response to 'surprising trials', in autistic, compared with non-autistic adults. Similarly, reduced surprise when expectations were violated was reported in autistic children and young adolescents [168],[408]. Our results (Chapter 2) directly contrast with these findings; autistic adults showed typical use of priors and typical response to surprising cues. However, there are important methodological differences between our findings and previous work, namely the age of participants and the paradigm employed, with the aforementioned studies not requiring learning of a sequence. In addition, although Lawson and colleagues used a sample that was comparable to ours, with regard to age and symptom severity, they used a perceptual associative learning task, suggesting contrasting findings could be related to different networks of brain regions recruited across different tasks [291]–[293]. Finally, our results are in accord with findings examining predictive eye movements. Using a visual pursuit task, where participants were required to track targets that were transiently occluded, the authors reported

no differences in anticipatory eye movements between autistic and non-autistic children and adolescents [404]. In sum, although methodological differences make it difficult to directly compare between different studies, the current results contrast with previous reports of atypical use of priors and response to surprise. It is thus possible that predictive differences in autism cannot be fully extended to the motor domain and that some types of motor learning (e.g., motor sequence learning) are intact.

An important difference between the paradigm used in Chapter 2 and previous work concerns the manipulation of the volatility of the learning environment. For example, Lawson and colleagues proposed that atypical responses to surprising cue-cue parings in autistic adults in their study stemmed from over-estimation of volatility. The authors manipulated the volatility of the environment, with stable phases, where cue-cue associations rarely changed, and volatile phases, which featured frequent reversals in cue-cue associations. Group differences were found in higher-level beliefs about the volatility (the rate of change of cue-cue associations) of the environment, when modelled using a Hierarchical Gaussian Filter model (HGF) [65],[409], which allows estimation of uncertainty at different levels. It is thus possible that differences might have emerged in the paradigm we employed in Chapter 2 if the volatility of the learning environment had been manipulated. Indeed, although a study investigating predictive processes through sensorimotor control reported typical performance in autistic individuals [159], further work from the same research group found atypical performance in autistic individuals in a volatile environment [160]. While different paradigms were used in these studies (force-matching versus interception of a moving target), both indexed predictive sensorimotor control.

To test the proposal that the lack of group differences observed in Chapter 2 stemmed from lack of volatility manipulation, the volatility of the learning environment was manipulated in Chapter 3. However, in refute of this proposal, no differences in adjustment of learning rates were observed between groups. These results are in accord with previous work from Manning and colleagues which used a probabilistic associative learning task (adapted from Behrens et al., [88]) to investigate adjustment of learning rates to volatility in autistic children [125]. We replicated and extended this work, by examining probabilistic learning in autistic adults. Furthermore, learning demands were higher in our task, with the inclusion of an additional information source, the social cue. Thus, countering the proposal raised by Manning et al., that group differences in learning might emerge when using a more complex learning task. Taken together, results contrast with claims of atypical volatility processing in autism, at least

with regard to stimulus-reward associative learning. Furthermore, the empirical work in this thesis extends previous work, demonstrating typical adjustment of learning in autistic adults within a social context.

In Chapter 3, a group difference was observed in behaviour in volatile phases, independent of adjustment of learning rates. Specifically, on trials following a 'win', the non-autistic participants showed a greater tendency to stay with the rewarding response on the following trial, with more 'win-stay' behaviour in stable compared to volatile phases. Autistic individuals showed the opposite, with more win-staying in volatile phases. While this effect was unexpected and did not translate to any impairments in task performance, similar findings have been reported previously. For example, in a reward learning task, which included volatile, stable, and noisy conditions, less optimal behaviour in volatile phases was found to correlate with levels of autistic traits; however, this was unrelated to adjustment of learning rates to changes in volatility [99]. Similarly, less flexible behaviour on a probabilistic reversal learning task was observed in both adults and children [46]. Finally, studies have reported that autistic individuals more frequently revert to the previously learned response [45], and require more time to learn new responses after a reversal [44]. Atypical performance in volatile phases could thus reflect difficulties in learning new responses (and forgetting previously learned responses) when stimulus-outcome associations change after a reversal, rather than from atypical adjustment of learning rate stemming from high and inflexible PEs (i.e., from a reduced capacity to distinguish volatility from noise [12],[78]. Thus, differences in autistic behaviour in volatile environments cannot be conclusively attributed to an impairment in tracking volatility.

Next, autistic individuals are proposed to show atypicalities in the social domain, due to the inherently complex and unpredictable nature of social information [13],[78]. Indeed, a recent review examining predictive mechanisms in autism concluded that evidence supports atypical predictive mechanisms in autism overall, but that these atypicalities are likely to be increased in the social domain [410]. However, it must be highlighted that only six studies examining prediction in the social domain were included in this review; two examined social prediction through the perspective taking of another individual (with opposing results [187],[411]), one found differing neural, but not behavioural, response to social reward and one study found no differences between social and non-social conditions [121]. Two further studies focusing specifically on social learning found atypical response to social information. Robic and colleagues [101] reported impaired performance in volatile environments in autistic

individuals, with greater impairment when advice was social as opposed to non-social. However, the computational mechanisms underpinning this impairment are unclear, as, rather than estimating individual measures of learning, the authors compared the number of participants in each group who reached a success criterion (60% correct). Similarly, Sevgi et al. [188] reported that a higher level of autistic traits correlated with difficulties in integrating social information with individual information during learning. This research, however, examined autistic traits in a non-autistic population, and these results should be replicated in individuals with a clinical diagnosis of autism. In sum, there is still limited research examining predictive processing in the social domain. The current results add to this literature, providing evidence against atypical predictive learning from social information in autistic adults.

A plausible explanation for the null findings reported here is our use of non-overtly social cues when representing social information. Studies in support of atypical predictive processing in the social domain use images or videos of face stimuli [59],[101],[123],[188]. This is also the case in the motor domain, where the majority of research in this field examines action inference in autism. For example, studies finding differences in autistic individuals used biological stimuli, such as point-light displays or videos of actors making intentional movements [120],[173],[174]. In contrast, the paradigms used in this thesis, examining social learning and motor sequence learning, did not use overtly social stimuli. This raises the possibility that requirements for face perception or biological motion processing, which have been reported as atypical in autism [178],[304],[412],[413], could explain contrasting results. In conclusion, when social information is represented through non-overtly social cues, as in the paradigm used in this thesis, no differences are observed in social learning. This could constrain predictive coding accounts to learning from biological stimuli, although it is important to note that results to date are mixed, with some accounts supporting [103],[104] and some opposing [105],[106] atypical predictive mechanisms for biological stimuli in autism.

7.1.3 Strengths and limitations

Chapters 2 and 3 examined predictive processes in the context of motor sequence learning and probabilistic social learning. Both studies were (separately) carried out with individuals who had previously received a clinical diagnosis of autism spectrum disorder and a matched group of non-autistic control participants. Although different paradigms were used, both

indexed the PE-related processes hypothesised to be impaired under Bayesian/predictive processing accounts of autism.

Is it important to mention that drawing conclusions from estimated parameters from computational models of learning is not without problems, and, in order to draw strong conclusions from comparison of parameter estimates, a computational model is required to provide a good fit to choice data. However, to strengthen the conclusions drawn from analysis of estimated computational parameters, additional measures of learning were considered and null findings were supported by Bayesian statistical analyses [401],[414]. Furthermore, Bayesian model selection [415]was used to compare our model with alternative models, providing confidence that our model accurately described participants' behaviour.

Model comparison is a relative measure, however, raising the possibility that group differences could emerge when using a different model of learning. In studies finding evidence in support of atypical adjustment of learning, one commonality is the use of a hierarchical Gaussian filter (HGF) model [100],[188]. In the HGF model, volatility is estimated on trial-by-trial basis, and a parameter which captures beliefs about volatility is estimated. Neither the results reported in Chapter 3, nor the study by Manning and colleagues (which also reported typical predictive processes in ASD) used this analysis and thus do not estimate volatility on a trialwise basis. The choice of model across different studies could thus potentially explain contrasting findings. However, while possible, this explanation is unlikely as the change in learning rate between volatile and stable phases is also an index of volatility tracking, meaning that we would expect to see differences in changes in learning rate between stable and volatile blocks of trials, if differences existed.

One potential limitation (addressed in detail in Chapter 2), concerns the question of whether results from these studies can be generalized to the wider autistic population. Participants in both studies only included autistic adults without intellectual disability or impaired language skills. Furthermore, in the motor learning task reported in Chapter 2, there was a high ratio of female to male participants; in contrast to the general population [416]. Here, an opportunity sample was used due to practical considerations, and no differences were observed between male and female participants. It is, however, possible that sex differences in predictive processing mechanisms were present, but the studies in this thesis were not sufficiently powered to detect these effects. In refute of this, there is no a priori reason to believe this to be the case and the same pattern of results was observed when splitting by gender. Thus, even

if the ratio in our sample had included a higher proportion of males, it is unlikely that this would change our pattern of null results. Overall, the sample of participants who took part in the studies described in the current thesis was comparable to those from other research [100],[107],[108],[173], in sample size, age and diagnostic criteria.

Finally, the social learning task employed in Chapter 3 could be argued to be lacking in ecological validity. The social 'advice' in this paradigm was represented by a thin, red frame, highlighting a particular option, in contrast with previous work using similar paradigms, which utilised images/videos of faces to indicate social advice [95],[188],[230]. Thus, group differences might have emerged if the task had employed more overtly social stimuli. However, representing the social information as an abstract cue was a deliberate choice, to investigate social learning without the confound of face processing. In addition, all participants knew that the red shape indicated information from previous participants, as participants could not commence the task until reaching 100% accuracy on a pre-task quiz, with questions concerning the social nature of the red shape. Furthermore, individuals indicated (through a post-task questionnaire) both that they believed the red frame to represent social information, and that they relied on this information source when making decisions. Finally, the task used here has been demonstrated to be a good index of social learning, with social, but not individual, learning correlating with social traits [309]. Taken together, participants treat the red cue as social information in this paradigm, despite the use of non-overtly social stimuli.

7.1.4 Conclusions and future work

Taken together, the current results do not add support for atypical predictive learning in autism, with the empirical work reported in this thesis finding no differences in either social or motor learning when comparing autistic and non-autistic adults. Furthermore, results suggest that, when social information is predictable and represented with a simple cue, social learning is not impaired in autistic individuals. The evidence reported here adds to a growing body of work reporting *typical* predictive learning in autism, particularly outside the sensory-perceptual domain. The empirical findings in this thesis are therefore not consistent with the hypothesis that common, domain-general impairments in predictive processes underpin many diverse aspects of the autistic phenotype, and contrast proposals of amplified predictive deficits in the social domain.

These findings help to determine if predictive impairments are indeed a common computational cause, or whether they are limited to certain domains, i.e., to identify the specific circumstances wherein predictive learning is atypical in autism. This is important, both to increase our general understanding of underlying predictive mechanisms in the autistic brain, but also in highlighting areas for future investigation, i.e. to identify areas where training in predictive learning could bring tangible benefits to autistic individuals [10]. It is therefore important to determine if there are unique differences between learning types. Focusing on the social domain, in particular, an important question is whether training and aids to improve social cognitive mechanisms need to be specialised or whether domain-general learning aids would have the same effect. A key question thus concerns whether or not social learning is underpinned by domain-general associative learning principles. The current results showed that group membership did not dissociate social and individual learning, tentatively suggesting that social and individual learning rely on the same mechanisms. This is examined in more detail at a neurochemical level in subsequent chapters (Chapters 4 and 6).

The research presented here could be extended in a number of ways. For example, an investigation of predictive processes in motor sequence learning in a paradigm where volatility is manipulated (e.g. [290]). However, as autistic adults showed intact learning with regard to volatility in Chapter 3, it is possible, but improbable that differences would emerge with regard to motor learning. Furthermore, while both of the tasks employed here indexed prediction-related processes, they are testing slightly different hypotheses (i.e., reduced reliance on prior predictions in Chapter 2 versus atypical precision-weighting of PEs in Chapter 3). Indeed, this is the case in the wider literature, with methodology varying widely across studies. While differing paradigms are thought to index common predictive processes, it is unknown whether performance on one task correlates with that on another, and therefore, whether they can be used interchangeably in support (or indeed opposition) of the same hypotheses. Future work should assess this further. In addition, when comparing across different paradigms, possible confounds should be highlighted. For example, there is evidence that autistic predictive processing is intact when individuals are given explicit task instructions [404], when cues are attended versus unattended [417] and when stimuli are linked to reward [125], all highlighting a role for attention in modulating predictive processing in autism [418]. These potentially confounding factors should be given greater attention in future work, in order to allow stronger comparisons to be made across different studies.

In sum, with respect to one of the main hypotheses of this thesis, these results suggest that predictive mechanisms underpinning both social and motor sequence learning are intact in autistic adults. Results do not support impaired predictive coding as a core deficit that can be extended to explain social and motor learning atypicalities in autism. These results force us to think more critically about what overarching conclusions can be drawn from studies of predictive coding in autism and can help in refining and revising these theories.

7.2 Neurochemical mechanisms of social learning

This section examines evidence for, and against, the presence of social-specific neurochemical mechanisms for social learning, based on empirical findings in this thesis. The existence in the human brain of neural and neurochemical pathways that are specialised for learning from social and individual information is debated, with cognitive neuroscientific studies presenting mixed evidence. In line with theories of human cultural evolution, some cognitive neuroscience studies have found dissociable neural correlates for social and individual learning; evidence for social-specific learning mechanisms [95],[256]–[258]. In contrast, other studies have reported evidence showing that social learning is associated with the same brain areas [222]–[225] and, dopamine-mediated computations [230] as individual learning, i.e., domain-general neurochemical mechanisms for social learning. There is, however, a lack of research comparing the effect of variation in neurochemical signalling on social and individual reward learning in the same paradigm; crucial for analysing the differences between each type of learning simultaneously while allowing for control of variation in learning between individuals. Differences in social and individual learning as a function of variation in neurochemical signalling would provide evidence that social and individual learning are facilitated by different mechanisms, with similarities reinforcing the opposing argument, that both types of learning share underlying neurochemical mechanisms. In empirical Chapters 4-6, I examine evidence for the above. Specifically, with regards to both dopamine and serotonin signalling, I examine if there are dissociations between social and individual learning as a function of genetic variation in the serotonin transporter gene (SERT/SLC6A4) and the dopamine transporter genes (DAT/SLC6A3), and whether pharmacological manipulation of dopamine has a dissociable effect on social and individual learning. All results are integrated and interpreted in light of existing research. I then consider what the current results imply for the domain-specificity of social learning and other social

processes, both for autism and in wider research. General strengths and weaknesses of the methodological approaches used will be outlined and I will discuss how the empirical research reported in the current thesis could be extended in future work.

7.2.1 Summary of Results

7.2.1.1. In Chapter 4, a large-scale behavioural genetics approach was employed, to investigate whether there was evidence to support the existence of dissociable genetic contributions to social and individual learning. Focusing on genes related to the monoamine neurotransmitters dopamine and serotonin, we investigated whether genetic variation in dopamine- and serotonin-related single nucleotide polymorphisms had different effects on social versus individual learning. An online version of the probabilistic social learning task (SLT) employed in Chapter 3, was utilised to dissociate behavioural markers of learning from social and individual learning. Results demonstrated that variation in participants' ability to adapt both social and individual learning rates to volatility covaried with serotonin-related genotype, providing preliminary evidence for a (domain-general) role for serotonergic signalling in adjusting learning rate. In addition, serotonin-related genotype covaried with the extent to which participants were biased towards social information during decision making as indexed by the social weighting parameter, ζ. Finally, significant interactions were observed between dopamine-related genetic variation and both environmental volatility and learning source with regard to learning rates. These results suggested the presence of a neurochemical dissociation between social and individual learning, in line with theories proposing that social and individual learning can be dissociated. However, in the paradigm used in this chapter, and in previous work finding dissociations between social and individual learning [93],[95], the social information is an indirect source of information, secondary to the individual reward information. For example, the social information is less salient and represented temporally after the individual information, with its utility needing to be inferred from the primary reward feedback. The social/non-social nature of the information source during learning is therefore confounded with whether the information source is the primary source of information, or rather, an additional, secondary source. Thus, the dissociations between individual and social learning observed here could instead be dissociations between learning from a primary and secondary source of information.

7.2.1.2. Chapter 5 builds on the results reported in the previous chapter, describing the rationale behind the development of an adapted version of the social learning task (SLT). The aim was to develop a task version where the social nature (social/individual) of the information was not confounded with status (primary/secondary) during learning. In line with this, we aimed to create a version of the task where the primary and secondary nature of the learning sources were switched, with the social information the primary source, and the individual information, the secondary, additional information source. To do so, several aspects of the information sources were manipulated, including saliency, temporal order and link to reward. To test whether the task manipulation had modulated participants' behaviour in the expected direction (i.e., whether the social information was the primary source of learning), different indices of learning were measured in a large sample and compared with measures from previous work utilising the standard version of the SLT [93]. Results suggested that the task manipulation affected several different indices of learning, including optimal beta scores and win-stay, lose-shift behaviour, with more weight given to social information during learning in the manipulated version of the task. However, while results were in the expected direction, both frequentist and Bayesian analyses suggested that stronger manipulations were required. Subsequently, task instructions provided to participants were updated for use in the paradigm employed in Chapter 6; participants were here primarily asked to choose between going with, or against, the group's choice, i.e., with or against the social information. Thus, employment of this task version (social-primary task) and the original version (individual-primary) allowed orthogonalization of the status and social nature of learning sources.

7.2.1.3. In Chapter 6 [419], the social-primary task and the individual-primary task were employed in a between-subjects design, to investigate whether the dopamine-dependent reinforcement learning (RL) process could be dissociated in a social versus individual condition, independently of the "status" of the learning source. Thus, enabling testing of the hypothesis highlighted in both Chapter 4 and in previous work [93]; that learning types can be dissociated along a primary versus secondary rather than an individual versus social axis. In a double-blind, placebo-controlled design, over two separate days participants received a dopamine antagonist, or placebo, and completed the social learning task, requiring learning from social and individual sources. Participants were randomly assigned to one of two task groups; for the social-primary group, social information was the primary learning source, whereas for the individual-primary group, individual information was the primary learning

source, meaning that status (primary, secondary) and social nature (social, individual) were orthogonalized. Results showed that manipulation of dopamine signalling affected primary learning irrespective of the social/individual nature of the information source, with no effect on learning from the secondary source. That is, as with individual learning, social learning was modulated by manipulation of dopamine signalling when social information was the primary learning source (i.e., in the social-primary group), but not when it comprised a secondary, additional element. These results, showing that dopaminergic mechanisms underpinning learning can be dissociated along a primary-secondary but not a social-individual axis, comprise positive evidence for shared dopaminergic signalling mechanisms for social and individual learning. These results add support for the existence of domain-general mechanisms underlying social learning and suggest that dopaminergic pathways in the human brain can process both social and non-social cues and flexibly switch between the two.

7.2.2 Interpretation of results

Taken together, empirical results in this thesis do not add support for dissociable neurochemical mechanisms underpinning social and individual learning. First, although results from Chapter 4 indicated dissociations between dopaminergic signalling mechanisms underpinning social and individual learning, this cannot be taken as conclusive evidence in support of social-specific neurochemical learning mechanisms. Importantly, Bayesian analyses provided moderate evidence *against* the inclusion of an interaction between dopamine-related genotype and (social versus individual) learning source, and post-hoc tests were not significant, suggesting that these results should be interpreted with caution. Furthermore, dissociations between individual and social learning cannot be confidently ascribed to the social nature of the information source; in this task, the social nature of the information was confounded with its secondary status during learning.

This confound was highlighted by Cook and colleagues [93], who provided preliminary evidence against a neurochemical dissociation between social and individual learning mechanisms. The authors reported that methylphenidate (MPH), a catecholamine reuptake inhibitor, affected participants' ability to adjust learning rate in response to changes in environmental volatility, with this effect restricted to learning from the individual source of information. The task was adapted to include a non-social control condition; while half of the

participants were informed that the secondary cue represented social advice, the remainder were told that it represented non-social advice (output from a rigged roulette wheel). This enabled the authors to determine whether dissociations between social and individual learning were better explained in terms of social versus non-social or secondary versus primary nature of the information source. MPH affected learning specifically while learning from the primary (individual) reward, with no effect when learning from the secondary information source, regardless of whether participants believed it was "social" or not. These results suggested that the differing effect of catecholamine perturbation dissociated between learning from primary and secondary information, rather than between individual and social sources. However, this study could not provide positive evidence in support of domain-general mechanisms for social learning: as MPH did not affect secondary learning, social and non-social learning could have relied on different neurochemical mechanisms. This study therefore could not provide conclusive evidence for domain-general mechanisms for social learning.

The current results (Chapter 6) extend the work conducted by Cook et al. [93] in two ways. First, orthogonalization of the status and social nature of the learning source allowed investigation into the effects of these factors independently of one another, allowing determination of which factor accounted for dissociations. Results demonstrated common dopaminergic mechanisms for social and individual learning when they were the primary learning source, independently of social nature. Second, in the study conducted by Cook and colleagues [93], it is unclear if observed effects on learning rate were mediated through effects on dopaminergic or norepinephrinergic signalling, as MPH blocks the reuptake of both. In the current study, we used a specific dopamine receptor antagonist, haloperidol, meaning that effects could be attributed to dopamine-dependent processes. Furthermore, haloperidol acts with high affinity on D2 receptors [383], which are found in highest concentrations in the mesolimbic signalling pathway [381],[382], an area key for reward learning [71],[211]. Meaning that, although the effects of haloperidol differ across different regions [390], we can be relatively confident in localising the observed effects on learning to striatal regions. However, a combination of neuroimaging and a pharmacological intervention could strengthen these results, by allowing a more precise localisation of effects.

With regard to serotonergic signalling (Chapter 4), empirical findings in this thesis support a lack of dissociation between social and individual learning as a function of differences in serotonin-related genetic variation. Instead, there is evidence for an association between

serotonergic variation and adjustment of learning rate, suggesting a domain-general role for serotonergic signalling in adaptation of learning rate to volatility. These results are in line with a role for neuromodulators in meta-learning [67],[91]. Although serotonin has long been associated with learning and decision-making [241],[420], its exact role remains unclear, with proposals including opponency to dopamine signalling [236],[332], reward discounting [421] and reward learning [233],[235],[238]. The current results, suggesting a role in adjusting learning rate, are in agreement with previous reports of variation in serotonin signalling affecting reversal learning, alongside other indices of flexible behaviour [237],[239],[241],[335],[336]. Indeed, recent evidence from animal studies report a role for serotonergic signalling in adjusting learning to volatility [337], and altering learning rates, potentially through tracking uncertainty of the learning environment [92]. The empirical work in this thesis adds to this, providing preliminary evidence for a role for serotonin in adjusting learning rates in humans.

In Chapter 4, an unpredicted role for serotonin in decision-making was observed. Specifically, s-allele carriers showed a bias towards greater use of social information during decision making. The short (s) allele is one of two common alleles in the promotor region of the gene encoding the serotonin reuptake transporter (SERT), with the s-allele associated with increased extracellular serotonin, mediated by a decrease in transcriptional efficacy, and reduced SERT function [328],[329]. An association between s-allele carriers and social learning is in agreement with previous work in social fear learning, although previous work has not directly compared social and individual learning [255]. Moreover, it is important to note that the social nature and secondary status of the learning source are confounded, meaning that this interaction could instead reflect an association between increased serotonergic signalling and a reliance on the *secondary* source of information, during decision-making. Accordingly, this confound should be ruled out before the results from this chapter are used to support a specific association between serotonergic signalling and increased use of social information.

In addition to helping to resolve issues with regard to the existence of social-specific neurochemical learning mechanisms, the current work adds to a broader discussion concerning the existence of specific neural regions for social processes. In contrast to studies proposing regions or pathways specialised for social processes [184],[185],[257],[263],[312],[386], there is a growing consensus that many brain areas and networks previously thought to have a social function, are rather implicated in both social and non-social processes [422]. For example, there is evidence that neural pathways in the human brain vary based on the current task relevance of information. Nicolle and colleagues [265] reported that, rather than a self

versus other distinction, information in the medial prefrontal cortex (mPFC) is arranged according to task-relevance, with ventral regions of the mPFC tracking task-relevant information (e.g., information about the self during a self-relevant trial and information about another individual during an other-relevant trial). In contrast, dorsal regions of the mPFC keep track of task-irrelevant information (e.g., information about the self during an other-relevant trial and the other individual during a self-relevant trial) [265],[266]. Further research showed that, when tracking one's own, and another's beliefs, the distinction between PEs for the self and for others was flexible, and the manner in which PEs were identified as representing the self or other shared common computations with non-social processes (inter-temporal reasoning) [368]. Finally, the mPFC encodes value representations, independently of self-versus other distinctions [367],[423]. The results in this thesis are in agreement, suggesting neurochemical pathways (such as the mesolimbic dopaminergic signalling pathway), are not specialised for social or individual input, but rather, process both types of information flexibly, with manipulations of signalling having different effects depending on the status of the information source.

7.2.3 Strengths and limitations

Several limitations and strengths specific to each chapter have been discussed previously, however, some general limitations will be mentioned in this section. First, one limitation concerns the statistical analysis in Chapter 5, where indices of learning from different task versions were compared. Due to time constraints, datasets from different task versions were not collected at the same time, meaning that groups were not matched on demographic measures, and were conducted at varying times and locations; all factors which could have influenced the pattern of results. In addition, the participant sample in the social-direct group was relatively homogenous in terms of gender and age. Accordingly, when this task was employed in Chapter 6, groups were matched and randomly assigned to one of two tasks conditions, and, to improve generalisability of results, the participant sample included a wide age range and was balanced with regard to gender.

Second, the empirical results from Chapter 5, while suggesting that task manipulations resulted in an increase in the influence of social information during learning, were not conclusive; although the influence of social information on learning increased, it did not do so significantly, and, while manipulations affected certain indices of learning (win-stay, lose-

shift scores), learning rates were not affected. Accordingly, based on research showing that instructions can alter choice behaviour [357],[358], changes were made to the task instructions before this task was employed in Chapter 6 (Appendix 4.4). Indeed, in Chapter 6, to ensure that the task manipulation (orthogonalizing social/individual and primary/secondary information) had indeed modulated participants' behaviour, the influence of both information types on learning was compared, prior to our primary analysis of learning rate. Here, participants put more weight on the primary information source, regardless of social nature, during learning, thus validating our manipulation. Importantly, the rationale, methodology and analysis plan for the study reported in Chapter 6 were pre-registered.

A potential limitation of the current work concerns the level of analysis in Chapter 6. It has been argued that social-specific properties of processes can be observed on a number of different levels; the computational level, concerning the goal of the agent, the algorithmic level, representing the underlying computations, and the implementational level, for example, the neural region or circuit [189]. Here, we manipulated dopamine signalling (i.e., affecting the implementational level), and concluded that our manipulation had comparable effects at the algorithmic level; that is, manipulating dopamine signalling had comparable effects on learning rates related to social and individual learning. However, the current study only addresses social-specificity at the level of D2-mediated dopaminergic signalling, thus only testing one of many possible components at an implementational level. Therefore, the presence of dissociations between social and individual learning at a neural or circuit level cannot be fully ruled out, with potential differences including the location of neural activity. However, it is unclear as to how location-based differences would be selected for, in the absence of behavioural differences. Furthermore, if location-based differences do exist, these would be likely to result in differences in the magnitude of the effect of haloperidol on learning, as effects of haloperidol differ across the brain [389],[390]. In sum, it is unlikely that social learning relies on similar computational mechanisms as individual learning when it is the primary learning source but is localised to a different neural region or pathway. A further possibility is that other forms of disruption at the implementation level (e.g., pharmacological manipulation of *serotonin* signalling) could dissociate social from individual learning. However, results from Chapter 4 oppose this; no dissociations between social/individual learning rates were found as a function of variation in serotonergic signalling, meaning it is unlikely that implementational differences exist at the level of 5-HT signalling. A

pharmacological manipulation of the serotonin system would, however, provide stronger evidence.

A strength of the current empirical work concerns the computational modelling approach used in Chapters 4 and 6. In both chapters, random-effects Bayesian model selection was used to compare the model with alternative candidate models, with analyses supporting the use of the chosen model. Subsequently, to improve confidence in the ability of the model to accurately describe choice behaviour, data simulations and parameter recovery were carried out, with estimated and recovered parameters significantly correlated in both analyses. Furthermore, our main effect of interest (a significant effect of haloperidol on primary learning) was observable when analysing recovered parameters. Finally, simulated and actual choice data were significantly correlated. While confident that the chosen computational model was robust and could accurately describe behaviour, it is important to note that, as mentioned in the previous section, model selection is a relative technique. It is therefore possible that a different type of model, such as HGF, would have better explained choice data. However, with the current dataset, this model frequently failed to converge. Furthermore, since the task used in Chapters 5 and 6 featured a novel manipulation (enabling orthogonalization of primary/secondary and social/individual factors), it was unclear as to what the most appropriate priors for the modelling parameters were, if using a different model. In addition, as the study design and hypotheses were based on previous work [93]; the same model and model priors were employed here.

7.2.4 Conclusions and future work

Taken together, addressing the main research question of this section, the results in this thesis provide positive evidence in support of domain-general theories of social learning, and evidence against dissociable neurochemical mechanisms underpinning social and individual learning. Instead, results demonstrate that disrupting the dopamine system can have comparable effects on social and individual learning when they are the primary source. Existing studies did not control for primary versus secondary status, meaning that they could not provide convincing evidence for neurochemical dissociations between social and individual learning. The current work builds upon this, providing evidence that dissociations between learning types can be explained through dissociations between learning from a primary, versus secondary, rather than a social versus individual information source. These

results have important practical implications with respect to enhancing social and individual learning in a clinical setting and add to the wider debate concerning the social specificity of neural processes. However, as outlined in the previous paragraph, some outstanding questions remain.

First, while it is unlikely that social learning relies on similar computational mechanisms as individual learning but is localised to a different neural region or pathway, this should be ruled out. As the social learning task employed here is amenable to neuroimaging [95], a combined imaging and pharmacological approach would provide conclusive evidence in determining whether the location of neural activation varies depending on the primary/secondary status, rather than the social/ individual nature of information during learning. Second, while the task manipulation described in Chapter 5 resulted in the social information becoming the primary source of learning, this manipulation changed several aspects of the stimulus simultaneously: saliency, temporal order, and reward feedback. It is thus unclear as to which manipulation was the most successful. The utility of making social information into the primary learning source is highlighted in Chapter 6, whereby participants showed more optimal learning from the social information when it was the primary source. Thus, similar manipulations could potentially be used to improve learning about individuals, as well as from other individuals. To unpack which manipulation is the most important, future work could investigate this in a large sample, manipulating each aspect independently. Finally, future work could utilise this task in parallel with a pharmacological manipulation of serotonergic signalling.

Together the studies presented in this thesis implicate the dopaminergic and serotonergic neurotransmitter systems in both social and individual learning and contribute to the debate concerning the existence of social-specific learning mechanisms. These results support the view that there are domain-general neurochemical mechanisms supporting social learning and contrast the existence of social-specific mechanisms, particularly with regard to dopaminergic signalling.

7.3 General Conclusion

Taken together, the empirical findings in this thesis do not find evidence of atypical predictive coding in social and motor learning, in contrast to accounts proposing atypical

predictive processing as a core domain-general deficit in autism. These results add to a recent body of empirical research helping to constrain predictive coding accounts of autism. Furthermore, this thesis provides positive evidence for domain-general dopaminergic mechanisms for social learning, suggesting that the learning process is modulated as a function of domain-general factors such as primary versus secondary status and volatility, but not social versus individual nature. These results, finding evidence against specific neurochemical mechanisms for social learning, are in agreement with a lack of social-specific learning differences in autistic adults. Thus, highlighting the possibility that atypicalities in autism lie in processing overtly social cues, rather than in the social learning process itself. To conclude, the results reported in this thesis have important implications, both for predictive coding accounts of learning in autism and for helping to resolve a wider debate concerning the domain-specificity of social cognitive processes.

References

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. (Arlington, 2013).

2. Gowen, E. & Hamilton, A. Motor abilities in autism: A review using a computational context. *J. Autism Dev. Disord.* **43**, 323–344 (2013).

3. Coll, S. M., Foster, N. E. V., Meilleur, A., Brambati, S. M. & Hyde, K. L. Sensorimotor skills in autism spectrum disorder: A meta-analysis. *Res. Autism Spectr. Disord.* **76**, 101570 (2020).

4. Cook, J. L., Blakemore, S. J. & Press, C. Atypical basic movement kinematics in autism spectrum conditions. *Brain* **136**, 2816–2824 (2013).

5. Frith, U., Happe, F. & Siddons, F. Autism and Theory of Mind in Everyday Life. *Soc. Dev.* **3**, (1994).

6. Baron-Cohen, S. Autism and the origins of social neuroscience. *Cogn. Neurosci. Soc. Behav.* 239–255 (2005) doi:10.4324/9780203311875.

7. Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S. & Schultz, R. T. The social motivation theory of autism. *Trends Cogn. Sci.* **16**, 231–239 (2012).

8. Marco, E. J., Hinkley, L. B. N., Hill, S. S. & Nagarajan, S. S. Sensory Processing in Autism: A Review of Neurophysiologic Findings. *Pediatr. Res.* **69**, 1–7 (2011).

9. Mottron, L., Dawson, M., Soulières, I., Hubert, B. & Burack, J. Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *J. Autism Dev. Disord.* **36**, 27–43 (2006).

10. Haker, H., Schneebeli, M. & Stephan, K. E. Can Bayesian theories of autism spectrum disorder help improve clinical practice? *Front. Psychiatry* **7**, 0–17 (2016).

11. Sinha, P. *et al.* Autism as a disorder of prediction. *Proc. Natl. Acad. Sci. USA* **111**, 15220–15225 (2014).

12. Van de Cruys, S. *et al.* Precise minds in uncertain worlds: Predictive coding in autism. *Psychol. Rev.* **121**, 649–675 (2014).

13. Lawson, R. P., Rees, G. & Friston, K. J. An aberrant precision account of autism. *Front. Hum. Neurosci.* **8**, (2014).

14. Bushwick, N. L. Social learning and the etiology of autism. *New Ideas Psychol.* **19**, 49–75 (2001).

15. Cleeremans, A. & Dienes, Z. Computational models of implicit learning. in *The Cambridge handbook of computational psychology* (ed. Sun, R.) 396–421 (2008).

16. Lieberman, M. D. Intuition: A social cognitive neuroscience approach. *Psychol. Bull.* **126**, 109–136 (2000).

17. Happé, F. Why are savant skills and special talents associated with autism? *World Psychiatry* **17**, 280–281 (2018).

18. Crespi, B. J. Autism as a disorder of high intelligence. *Front. Neurosci.* **10**, 1–17 (2016).

19. Qian, N. & Lipkin, R. M. A Learning-Style Theory for Understanding Autistic Behaviors. *Front. Hum. Neurosci.* **5**, 1–17 (2011).

20. Plaisted, K., O'Riordan, M. & Baron-Cohen, S. Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *J. Child Psychol. Psychiatry Allied Discip.* **39**, 765–775 (1998).

21. Hannant, P., Tavassoli, T. & Cassidy, S. The role of sensorimotor difficulties in autism spectrum conditions. *Front. Neurol.* **7**, 124 (2016).

22. Dawson, M., Mottron, L. & Gernsbacher, M. Learning in Autism. In J.H. Byrne (Series Ed.) & H. Roediger (Vol. Ed.). *Learn. Mem. A Compr. Ref. Cogn. Psychol. Mem.* 759–772 (2008).

23. Cohen, M. X. Neurocomputational mechanisms of reinforcement-guided learning in humans: A review. *Cogn. Affect. Behav. Neurosci.* **8**, 113–125 (2008).

24. Rescorla, R. A. & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. in *Clasical conditioning II: current research and theory* 64–99 (Appleton Century Crofts, 1972).

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

26. Steinberg, E. E. *et al.* A Causal Link Between Prediction Errors, Dopamine Neurons and Learning. *Nat. Neurosci.* **16**, 1–19 (2013).

27. Friston, K. J. A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* **360**, 815–836 (2005).

28. Den Ouden, H., Kok, P. & de Lange, F. P. How prediction errors shape perception, attention, and motivation. *Front. Psychol.* **3**, 1–12 (2012).

29. Friston, K. J., Mattout, J. & Kilner, J. Action understanding and active inference. *Biol. Cybern.* **104**, 137–160 (2011).

30. Mussey, J. L., Travers, B. G., Klinger, L. G. & Klinger, M. R. Decision-Making Skills in ASD: Performance on the Iowa Gambling Task. *Autism Res.* **8**, 105–114 (2015).

31. Solomon, M., Smith, A. C., Frank, M. J., Ly, S. & Carter, C. S. Probabilistic reinforcement learning in adults with autism spectrum disorders. *Autism Res.* **4**, 109–120 (2011).

32. Solomon, M. *et al.* Feedback-driven trial-by-trial learning in autism spectrum disorders. *Am. J. Psychiatry* **172**, 173–181 (2015).

33. Harris, H. *et al.* Perceptual learning in autism: Over-specificity and possible remedies. *Nature Neuroscience* vol. 18 1574–1576 (2015).

34. Munson, J., Faja, S., Meltzoff, A., Abbott, R. & Dawson, G. Neurocognitive predictors of social and communicative developmental trajectories in preschoolers with autism spectrum disorders. *J. Int. Neuropsychol. Soc.* **14**, 956–966 (2008).

35. Lin, A., Rangel, A. & Adolphs, R. Impaired learning of social compared to monetary rewards in autism. *Front. Neurosci.* **6**, 1–7 (2012).

36. Clements, C. C. *et al.* Evaluation of the social motivation hypothesis of autism a systematic review and meta-analysis. *JAMA Psychiatry* **75**, 797–808 (2018).

37. Mosner, M. G. *et al.* Neural Mechanisms of Reward Prediction Error in Autism Spectrum Disorder. *Autism Res. Treat.* **2019**, 1–10 (2019).

38. Scott-Van Zeeland, A. A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A. & Bookheimer, S. Y. Reward processing in autism. *Autism Res.* **3**, 53–67 (2010).

39. Schipul, S. E. & Just, M. A. Diminished neural adaptation during implicit learning in autism. *Neuroimage* **125**, 332–341 (2016).

40. Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K. & Lum, J. A. G. Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Front. Psychol.* **7**, 1–18 (2016).

41. Foti, F., De Crescenzo, F., Vivanti, G., Menghini, D. & Vicari, S. Implicit learning in individuals with autism spectrum disorders: A meta-analysis. *Psychol. Med.* **45**, 897–910 (2015).

42. Cools, R., Clark, L., Owen, A. M. & Robbins, T. W. Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *J. Neurosci.* **22**, 4563–4567 (2002).

43. Brown, V. J. & Tait, D. S. Behavioral Flexibility: Attentional Shifting, Rule Switching, and Response Reversal. in *Encyclopedia of Psychopharmacology* (eds. Stolerman, I. P. & Price, L. H.) 1–7 (Springer Berlin Heidelberg, 2010). doi:10.1007/978-3-642-27772-6_340-2.

44. South, M., Newton, T. & Chamberlain, P. D. Delayed Reversal Learning and Association With Repetitive Behavior in Autism Spectrum Disorders. *Autism Research* vol. 5 398–406 (2012).

45. D'Cruz, A. M. *et al.* Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology* **27**, 152–160 (2013).

46. Crawley, D. *et al.* Modeling flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and less optimal learning in autism in each age group. *PLoS Biol.* **18**, 1–25 (2020).

47. Costescu, C. A., Vanderborght, B. & David, D. O. Reversal Learning Task in Children with Autism Spectrum Disorder: A Robot-Based Approach. *J. Autism Dev. Disord.* **45**, 3715–3725 (2015).

48. D'Cruz, A. M., Mosconi, M. W., Ragozzino, M. E., Cook, E. H. & Sweeney, J. A. Alterations in the functional neural circuitry supporting flexible choice behavior in autism spectrum disorders. *Transl. Psychiatry* **6**, e916 (2016).

49. Ridley, R. M. The psychology of perseverative and stereotyped behaviour. *Prog. Neurobiol.* **44**, 221–231 (1994).

50. Gomot, M. & Wicker, B. A challenging, unpredictable world for people with Autism Spectrum Disorder. *Int. J. Psychophysiol.* **83**, 240–247 (2012).

51. Landry, O. & Mitchell, P. An examination of perseverative errors and cognitive flexibility in autism. *PLoS One* **16**, 1–19 (2021).

52. Frith, U. Cognitive explanations of autism. *Acta Paediatr. Int. J. Paediatr. Suppl.* **85**, 63–68 (1996).

53. Pellicano, E. & Burr, D. When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends Cogn. Sci.* **16**, 504–510 (2012).

54. Brock, J. Alternative Bayesian accounts of autistic perception: Comment on Pellicano and Burr. *Trends Cogn. Sci.* **16**, 573–574 (2012).

55. Knill, D. C. & Pouget, A. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).

56. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).

57. Friston, K. J. Learning and inference in the brain. *Neural Networks* **16**, 1325–1352 (2003).

58. Lieder, I. *et al.* Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. *Nat. Neurosci.* **22**, 256–264 (2019).

59. Ewbank, M. P. *et al.* Repetition suppression and memory for faces is reduced in adults with autism spectrum conditions. *Cereb. Cortex* **27**, 92–103 (2017).

60. Ewing, L., Pellicano, E. & Rhodes, G. Atypical updating of face representations with experience in children with autism. *Dev. Sci.* **16**, 116–123 (2013).

61. Pellicano, E., Rhodes, G. & Calder, A. J. Reduced gaze aftereffects are related to difficulties categorising gaze direction in children with autism. *Neuropsychologia* **51**, 1504–1509 (2013).

62. Croydon, A., Karaminis, T., Neil, L., Burr, D. & Pellicano, E. The light-from-above prior is intact in autistic children. *J. Exp. Child Psychol.* **161**, 113–125 (2017).

63. Van de Cruys, S., Vanmarcke, S., Van de Put, I. & Wagemans, J. The Use of Prior Knowledge for Perceptual Inference Is Preserved in ASD. *Clin. Psychol. Sci.* **6**, 382–393 (2018).

64. Pell, P. J. *et al.* Intact priors for gaze direction in adults with high-functioning autism spectrum conditions. *Mol. Autism* **7**, 1–10 (2016).

65. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* **5**, 9 (2011).

66. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).

67. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).

68. Mathys, C. Uncertainty , precision , prediction errors – and their relevance to computational psychiatry Uncertainty : a shamelessly artificial example. (2015).

69. Moran, R. J. *et al.* Free energy, precision and learning: The role of cholinergic neuromodulation. *J. Neurosci.* **33**, 8227–8236 (2013).

70. Iglesias, S. *et al.* Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. *Neuron* **80**, 519–530 (2013).

71. Glimcher, P. W. & Bayer, H. M. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **103**, 2304–2312 (2005).

72. Sterzer, P. *et al.* The Predictive Coding Account of Psychosis. *Biol. Psychiatry* **84**, 634–643 (2018).

73. Stephan, K. E., Iglesias, S., Heinzle, J. & Diaconescu, A. O. Translational Perspectives for Computational Neuroimaging. *Neuron* **87**, 716–732 (2015).

74. Adams, R. A., Huys, Q. J. M. & Roiser, J. P. Computational Psychiatry: Towards a mathematically informed understanding of mental illness. *J. Neurol. Neurosurg. Psychiatry* **87**, 53–63 (2016).

75. van Boxtel, J. J. & Lu, H. A predictive coding perspective on autism spectrum disorders. *Front Psychol* **4**, 19 (2013).

76. Friston, K. J., Lawson, R. & Frith, C. D. On hyperpriors and hypopriors: Comment on Pellicano and Burr. *Trends Cogn. Sci.* **17**, 1 (2013).

77. Van de Cruys, S., Van der Hallen, R. & Wagemans, J. Disentangling signal and noise in autism spectrum disorder. *Brain Cogn.* **112**, 78–83 (2017).

78. Palmer, C. J., Lawson, R. P. & Hohwy, J. Bayesian approaches to autism: Towards volatility, action, and behavior. *Psychol. Bull.* **143**, 521–542 (2017).

79. Ruggeri, B., Sarkans, U., Schumann, G. & Persico, A. M. Biomarkers in autism spectrum disorder: The old and the new. *Psychopharmacology (Berl).* **231**, 1201–1216 (2014).

80. Gottesman, I. I. & Gould, T. D. The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions Irving. *Am. J. Psychiatry* **160**, 636–645 (2003).

81. Happé, F. & Frith, U. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* **36**, 5–25 (2006).

82. Booth, R. D. L. & Happé, F. G. E. Evidence of Reduced Global Processing in Autism Spectrum Disorder. *J. Autism Dev. Disord.* **48**, 1397–1408 (2018).

83. Boulter, C., Freeston, M., South, M. & Rodgers, J. Intolerance of Uncertainty as a Framework for Understanding Anxiety in Children and Adolescents with Autism Spectrum Disorders. *J Autism Dev Disord* **44**, 391–1402 (2014).

84. Jenkinson, R., Milne, E. & Thompson, A. The relationship between intolerance of uncertainty and anxiety in autism: A systematic literature review and meta-analysis. *Autism* vol. 24 1933–1944 (2020).

85. Dayan, P., Kakade, S. & Read Montague, P. Learning and selective attention. *Nat. Neurosci.* **3**, 1218–1223 (2000).

86. O'Reilly, J. X. Making predictions in a changing world-inference, uncertainty, and learning. *Front. Neurosci.* **7**, 105 (2013).

87. Soltani, A. & Izquierdo, A. Adaptive learning under expected and unexpected uncertainty. *Nat. Rev. Neurosci.* **20**, 635–644 (2019).

88. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).

89. Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* **30**, 12366–12378 (2010).

90. Schweighofer, N. & Doya, K. Meta-learning in reinforcement learning. *Neural Networks* **16**, 5–9 (2003).

91. Doya, K. Metalearning and neuromodulation. *Neural Networks* **15**, 495–506 (2002).

92. Iigaya, K., Fonseca, M. S., Murakami, M., Mainen, Z. F. & Dayan, P. An effect of serotonergic stimulation on learning rates for rewards apparent after long intertrial intervals. *Nat. Commun.* **9**, 2477 (2018).

93. Cook, J. L. *et al.* Catecholaminergic modulation of meta-learning. *Elife* **8**, 1–38 (2019).

94. Jepma, M. *et al.* Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. *PLoS Comput. Biol.* **12**, e1005171 (2016).

95. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).

96. Farashahi, S., Donahue, C. H., Hayden, B. Y., Lee, D. & Soltani, A. Flexible combination of reward information across primates. *Nat. Hum. Behav.* **3**, 1215–1224 (2019).

97. Fineberg, S. K. *et al.* Differential Valuation and Learning From Social and Nonsocial Cues in Borderline Personality Disorder. *Biol. Psychiatry* **84**, 838–845 (2018).

98. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).

99. Goris, J. *et al.* Autistic traits are related to worse performance in a volatile reward learning task despite adaptive learning rates. *Autism* **25**, 440–451 (2021).

100. Lawson, R. P., Mathys, C. & Rees, G. Adults with autism overestimate the volatility of the sensory environment. *Nat. Neurosci.* **20**, 1293–1299 (2017).

101. Robic, S. *et al.* Decision-Making in a Changing World: A Study in Autism Spectrum Disorders. *J. Autism Dev. Disord.* **45**, 1603–1613 (2015).

102. Jenkins, R., Beaver, J. D. & Calder, A. J. I thought you were looking at me: Direction-specific aftereffects in gaze perception. *Psychol. Sci.* **17**, 506–513 (2006).

103. Pellicano, E., Jeffery, L., Burr, D. & Rhodes, G. Abnormal adaptive face-coding mechanisms in children with Autism Spectrum Disorder. *Curr. Biol.* **17**, 1508–1512 (2007).

104. van Boxtel, J. J., Dapretto, M. & Lu, H. Intact recognition, but attenuated adaptation, for biological motion in youth with autism spectrum disorder. *Autism Res.* **9**, 1103–1113 (2016).

105. Karaminis, T., Arrighi, R., Forth, G., Burr, D. & Pellicano, E. Adaptation to the Speed of Biological Motion in Autism. *J. Autism Dev. Disord.* **50**, 373–385 (2020).

106. Cook, R., Brewer, R., Shah, P. & Bird, G. Intact facial adaptation in autistic adults. *Autism Res.* **7**, 481–490 (2014).

107. Lawson, R. P., Aylward, J., White, S. & Rees, G. A striking reduction of simple loudness adaptation in autism. *Sci. Rep.* **5**, 1–7 (2015).

108. Turi, M., Karaminis, T., Pellicano, E. & Burr, D. No rapid audiovisual recalibration in adults on the autism spectrum. *Sci. Rep.* **6**, 2–8 (2016).

109. Millin, R. *et al.* Reduced auditory cortical adaptation in autism spectrum disorder. *Elife* **7**, 1–15 (2018).

110. Tewolde, F. G., Bishop, D. V. M. & Manning, C. Visual Motion Prediction and Verbal False Memory Performance in Autistic Children. *Autism Res.* **11**, 509–518 (2018).

111. Gandhi, T. K. *et al.* Autonomic and Electrophysiological Evidence for Reduced Auditory Habituation in Autism. *J. Autism Dev. Disord.* **51**, 2218–2228 (2021).

112. Font-Alaminos, M. *et al.* Increased subcortical neural responses to repeating auditory stimulation in children with autism spectrum disorder. *Biol. Psychol.* **149**, 107807 (2020).

113. Kolesnik, A. *et al.* Increased cortical reactivity to repeated tones at 8 months in infants with later ASD. *Transl. Psychiatry* **9**, 16–18 (2019).

114. Tam, F. I. *et al.* Altered behavioral and amygdala habituation in high-functioning adults with

autism spectrum disorder: An fMRI study. *Sci. Rep.* **7**, 1–9 (2017).

115. Jamal, W., Cardinaux, A., Haskins, A. J., Kjelgaard, M. & Sinha, P. Reduced Sensory Habituation in Autism and Its Correlation with Behavioral Measures. *J. Autism Dev. Disord.* **51**, 3153–3164 (2021).

116. Vivanti, G. *et al.* Attention to novelty versus repetition: Contrasting habituation profiles in Autism and Williams syndrome. *Dev. Cogn. Neurosci.* **29**, 54–60 (2018).

117. Utzerath, C., Schmits, I. C., Buitelaar, J. & de Lange, F. P. Adolescents with autism show typical fMRI repetition suppression, but atypical surprise response. *Cortex* **109**, 25–34 (2018).

118. Leekam, S. R., Nieto, C., Libby, S. J., Wing, L. & Gould, J. Describing the sensory abnormalities of children and adults with autism. *J. Autism Dev. Disord.* **37**, 894–910 (2007).

119. Goris, J. *et al.* Sensory prediction errors are less modulated by global context in Autism Spectrum Disorder. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **3**, 667–674 (2018).

120. Amoruso, L. *et al.* Contextual priors do not modulate action prediction in children with autism. *Proc. R. Soc. B Biol. Sci.* **286**, (2019).

121. Greene, R. K. *et al.* Social and nonsocial visual prediction errors in autism spectrum disorder. *Autism Res.* **12**, 878–883 (2019).

122. Sapey-Triomphe, L. A., Weilnhammer, V. A. & Wagemans, J. Associative learning under uncertainty in adults with autism: Intact learning of the cue-outcome contingency, but slower updating of priors. *Autism* (2021) doi:10.1177/13623613211045026.

123. Kinard, J. L. *et al.* Neural Mechanisms of Social and Nonsocial Reward Prediction Errors in Adolescents with Autism Spectrum Disorder. *Autism Res.* **13**, 715–728 (2020).

124. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).

125. Manning, C., Kilner, J., Neil, L., Karaminis, T. & Pellicano, E. Children on the autism spectrum update their behaviour in response to a volatile environment. *Dev. Sci.* **20**, (2017).

126. Fournier, K. A., Hass, C. J., Naik, S. K., Lodha, N. & Cauraugh, J. H. Motor coordination in autism spectrum disorders: A synthesis and meta-analysis. *J. Autism Dev. Disord.* **40**, 1227–1240 (2010).

127. Donnellan, A. M., Hill, D. H. & Leary, M. R. Rethinking autism: Implications of sensory and movement differences for understanding and support. *Front. Integr. Neurosci.* **6**, 124 (2012).

128. Green, D. *et al.* Impairment in movement skills of children with autistic spectrum disorders. *Dev. Med. Child Neurol.* **51**, 311–316 (2009).

129. Pan, C. Y., Tsai, C. L. & Chu, C. H. Fundamental Movement skills in children diagnosed with autism spectrum disorders and attention deficit hyperactivity disorder. *J. Autism Dev. Disord.* **39**, 1694–1705 (2009).

130. Lai, M. C., Lombardo, M. V. & Baron-Cohen, S. Autism. *Lancet* **383**, 896–910 (2014).

131. Bhat, A. N. Is Motor Impairment in Autism Spectrum Disorder Distinct From Developmental Coordination Disorder? A Report From the SPARK Study. *Phys. Ther.* **100**, 633–644 (2020).

132. Provost, B., Lopez, B. R. & Heimerl, S. A comparison of motor delays in young children:

Autism spectrum disorder, developmental delay, and developmental concerns. *J. Autism Dev. Disord.* **37**, 321–328 (2007).

133. Gidley Larson, J. C. & Mostofsky, S. H. Motor deficits in autism. in *Autism: a neurological disorder of early brain development.* (eds. Tuchman, R. & Rapin, I.) (MacKeith Press, 2006).

134. Rinehart, N. J. *et al.* Gait function in high-functioning autism and Asperger's disorder: Evidence for basal-ganglia and cerebellar involvement? *Eur. Child Adolesc. Psychiatry* **15**, 256–264 (2006).

135. Jansiewicz, E. M. *et al.* Motor signs distinguish children with high functioning autism and Asperger's syndrome from controls. *J. Autism Dev. Disord.* **36**, 613–621 (2006).

136. Dziuk, M. A. *et al.* Dyspraxia in autism: Association with motor, social, and communicative deficits. *Dev. Med. Child Neurol.* **49**, 734–739 (2007).

137. Dowell, L. R., Mahone, E. M. & Mostofsky, S. H. Associations of postural knowledge and basic motor skill with Dyspraxia in Autism: Implication for abnormalities in distributed connectivity and motor learning. *Neuropsychology* **23**, 563–570 (2009).

138. Glazebrook, C. M., Elliott, D. & Szatmari, P. How do individuals with autism plan their movements? *J. Autism Dev. Disord.* **38**, 114–126 (2008).

139. Nazarali, N., Glazebrook, C. M. & Elliott, D. Movement planning and reprogramming in individuals with autism. *J. Autism Dev. Disord.* **39**, 1401–1411 (2009).

140. Jasmin, E. *et al.* Sensori-motor and daily living skills of preschool children with autism spectrum disorders. *J. Autism Dev. Disord.* **39**, 231–241 (2009).

141. Bo, J., Lee, C. M., Colbert, A. & Shen, B. Do children with autism spectrum disorders have motor learning difficulties? *Res. Autism Spectr. Disord.* **23**, 50–62 (2016).

142. Lloyd, M., MacDonald, M. & Lord, C. Motor skills of toddlers with autism spectrum disorders. *Autism* **17**, 133–146 (2013).

143. Belmonte, M. K. *et al.* Oral motor deficits in speech-impaired children with Autism. *Front. Integr. Neurosci.* **7**, 47 (2013).

144. Bhat, A. N., Landa, R. J. & Galloway, J. C. Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders. *Phys. Ther.* **91**, 1116–1129 (2011).

145. Cook, J. L. From movement kinematics to social cognition: The case of autism. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, (2016).

146. de Moraes, I. A. P. *et al.* Motor learning characterization in people with autism spectrum disorder: A systematic review. *Dement Neuropsychol* **11**, 276–286 (2017).

147. Gidley Larson, J. C. & Mostofsky, S. H. Evidence that the pattern of visuomotor sequence learning is altered in children with autism. *Autism Res.* **1**, 341–353 (2008).

148. Mostofsky, S. H., Goldberg, M. C., Landa, R. J. & Denckla, M. B. Evidence for a deficit in procedural learning in children and adolescents with autism: Implications for cerebellar contribution. *J. Int. Neuropsychol. Soc.* **6**, 752–759 (2000).

149. Grisoni, L. *et al.* Prediction and Mismatch Negativity Responses Reflect Impairments in Action Semantic Processing in Adults With Autism Spectrum Disorders. *Front. Hum. Neurosci.* **13**, 1–11 (2019).

150. Turi, M., Muratori, F., Tinelli, F., Morrone, M. C. & Burr, D. C. Autism is associated with reduced ability to interpret grasping actions of others. *Sci. Rep.* **7**, 1–8 (2017).

151. Ganglmayer, K., Schuwerk, T., Sodian, B. & Paulus, M. Do Children and Adults with Autism Spectrum Condition Anticipate Others' Actions as Goal-Directed? A Predictive Coding Perspective. *J. Autism Dev. Disord.* **50**, 2077–2089 (2020).

152. Amoruso, L., Finisguerra, A. & Urgesi, C. Autistic traits predict poor integration between top-down contextual expectations and movement kinematics during action observation. *Sci. Rep.* **8**, 1–10 (2018).

153. Friston, K. J., Daunizeau, J., Kilner, J. & Kiebel, S. J. Action and behavior: A free-energy formulation. *Biol. Cybern.* **102**, 227–260 (2010).

154. Friston, K. J. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).

155. Adams, R. A., Shipp, S. & Friston, K. J. Predictions not commands: Active inference in the motor system. *Brain Struct. Funct.* **218**, 611–643 (2013).

156. Körding, K. & Wolpert, D. M. Bayesian integration in sensorimotor learning. *Nature* **427**, 1–4 (2004).

157. Shipp, S., Adams, R. A. & Friston, K. J. Reflections on agranular architecture: Predictive coding in the motor cortex. *Trends Neurosci.* **36**, 706–716 (2013).

158. Finnemann, J. J. S., Plaisted-Grant, K., Moore, J., Teufel, C. & Fletcher, P. C. Low-level, prediction-based sensory and motor processes are unimpaired in Autism. *Neuropsychologia* **156**, 107835 (2021).

159. Arthur, T., Vine, S., Brosnan, M. & Buckingham, G. Predictive sensorimotor control in autism. *Brain* **143**, 3151–3163 (2020).

160. Arthur, T. *et al.* An examination of active inference in autistic adults using immersive virtual reality. *Sci. Rep.* 1–14 (2021) doi:10.1038/s41598-021-99864-y.

161. Nissen, M. J. & Bullemer, P. Attention requirements of learning: Evidence from performance measures. *Cogn. Psychol.* **19**, 1–32 (1987).

162. Nemeth, D. *et al.* Learning in autism: implicitly superb. *PLoS One* **5**, e11731 (2010).

163. Brown, J., Aczel, B., Jiménez, L., Kaufman, S. B. & Grant, K. P. Intact implicit learning in autism spectrum conditions. *Q. J. Exp. Psychol.* **63**, 1789–1812 (2010).

164. Travers, B. G., Klinger, M. R., Mussey, J. L. & Klinger, L. G. Motor-linked implicit learning in persons with autism spectrum disorders. *Autism Res.* **3**, 68–77 (2010).

165. Gordon, B. & Stark, S. Procedural learning of a visual sequence in individuals with Autism. *Focus Autism Other Dev. Disabl.* **22**, 14–22 (2007).

166. Müller, R. A., Cauich, C., Rubio, M. A., Mizuno, A. & Courchesne, E. Abnormal activity patterns in premotor cortex during sequence learning in autistic patients. *Biol. Psychiatry* **56**, 323–332 (2004).

167. Bestmann, S. *et al.* Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Curr. Biol.* **18**, 775–780 (2008).

168. Rinehart, N. J., Bradshaw, J. L., Brereton, A. V. & Tonge, B. J. Movement preparation in

high-functioning Autism and Asperger Disorder: A serial choice reaction time task Involving motor reprogramming. *J. Autism Dev. Disord.* **31**, 79–88 (2001).

169. Leekam, S. R. Social cognitive impairment and autism: What are we trying to explain? *Philos. Trans. R. Soc. B Biol. Sci.* **371**, (2016).

170. Baron-Cohen, S., Leslie, A. M. & Frith, U. Does the autistic child have a "theory of mind" ? *Cognition* **21**, 37–46 (1985).

171. Moutoussis, M., Fearon, P., El-Deredy, W., Dolan, R. J. & Friston, K. J. Bayesian inferences about the self (and others): A review. *Conscious. Cogn.* **25**, 67–76 (2014).

172. Kilner, J. M., Friston, K. J. & Frith, C. D. Predictive coding: An account of the mirror neuron system. *Cogn. Process.* **8**, 159–166 (2007).

173. Von der Lühe, T. *et al.* Interpersonal predictive coding, not action perception, is impaired in autism. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, (2016).

174. Chambon, V. *et al.* Reduced sensitivity to social priors during action prediction in adults with autism spectrum disorders. *Cognition* **160**, 17–26 (2017).

175. Palumbo, L., Burnett, H. G. & Jellema, T. Atypical emotional anticipation in high-functioning autism. *Mol. Autism* **6**, 1–17 (2015).

176. Lawson, R. P., Aylward, J., Roiser, J. P. & Rees, G. Adaptation of social and non-social cues to direction in adults with autism spectrum disorder and neurotypical adults with autistic traits. *Dev. Cogn. Neurosci.* **29**, 108–116 (2018).

177. Dawson, G. *et al.* Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Dev. Psychol.* **40**, 271–283 (2004).

178. Todorova, G. K., Hatton, R. E. & Pollick, F. E. Distinct Biological Motion Perception in Autism Spectrum Disorder: A Meta-Analysis. *Mol. Autism* **10**, (2019).

179. Heyes, C. M. Social learning in animals: categories and mechanisms. *Biol. Rev.* **69**, 207–231 (1994).

180. Gariépy, J. F. *et al.* Social learning in humans and other animals. *Front. Neurosci.* **8**, 58 (2014).

181. Galef, B. G. & Laland, K. N. Social learning in animals: Empirical studies and theoretical models. *Bioscience* **55**, 489–499 (2005).

182. Rendell, L. *et al.* Why copy others? *Science (80-.).* **328**, 208–213 (2010).

183. Thornton, A. & Clutton-Brock, T. Social learning and the development of individual and group behaviour in mammal societies. *Philos. Trans. R. Soc. B Biol. Sci.* **366**, 978–987 (2011).

184. Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E. & Roiser, J. P. Neurocomputational mechanisms of prosocial learning and links to empathy. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 9763–9768 (2016).

185. Apps, M. A. J., Lesage, E. & Ramnani, N. Vicarious reinforcement learning signáis when instructing others. *J. Neurosci.* **35**, 2904–2913 (2015).

186. Heyes, C. M. What's social about social learning? *J. Comp. Psychol.* **126**, 193–202 (2012).

187. Balsters, J. H. *et al.* Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain* **140**, 235–246 (2017).

188. Sevgi, M., Diaconescu, A. O., Henco, L., Tittgemeyer, M. & Schilbach, L. Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition. *Biol. Psychiatry* **87**, 185–193 (2020).

189. Lockwood, P. L., Apps, M. A. J. & Chang, S. W. C. Is There a 'Social' Brain? Implementations and Algorithms. *Trends Cogn. Sci.* **24**, 802–813 (2020).

190. Kendal, R. L. *et al.* Social learning strategies: Bridge-building between fields. *Trends Cogn. Sci.* **22**, 651–665 (2018).

191. Heyes, C. M. & Galef, B. G. *Social learning in animals: the roots of culture.* (Academic Press, 1996).

192. Hoppitt, W. & Laland, K. N. *Social learning: an introduction to mechanisms, methods, and models.* (Princeton University Press, 2013).

193. Pearce, J. M. Animal learning and cognition: an introduction. (2008).

194. Richerson, P. J. & Boyd, R. *Not By Genes Alone: How Culture Transformed Human Evolution*. (The University of Chicago Press, 2005).

195. Laland, K. N. Social learning strategies. *Learn. Behav.* **32**, 4–14 (2004).

196. Zentall, T. R. Perspectives on observational learning in animals. *J. Comp. Psychol.* **126**, 114–128 (2012).

197. Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W. & Laland, K. N. The evolutionary basis of human social learning. *Proc. R. Soc. B Biol. Sci.* **279**, 653–662 (2012).

198. McElreath, R. *et al.* Applying evolutionary models to the laboratory study of social learning. *Evol. Hum. Behav.* **26**, 483–508 (2005).

199. Templeton, J. J., Kamil, A. C. & Balda, R. P. Sociality and social learning in two species of corvids: The pinyon jay (Gymnorhinus cyanocephalus) and the Clark's nutcracker (Nucifraga columbiana). *J. Comp. Psychol.* **113**, (1999).

200. Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B. & Tomasello, M. Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science (80-.).* **317**, 1360–1366 (2007).

201. Heyes, C. M. & Pearce, J. M. Not-so-social learning strategies. *Proc. R. Soc. B Biol. Sci.* **282**, (2015).

202. Boogert, N. J., Giraldeau, L. A. & Lefebvre, L. Song complexity correlates with learning ability in zebra finch males. *Anim. Behav.* **76**, 1735–1741 (2008).

203. Bouchard, J., Goodyer, W. & Lefebvre, L. Social learning and innovation are positively correlated in pigeons (Columba livia). *Anim. Cogn.* **10**, 259–266 (2007).

204. Lefebvre, L. & Giraldeau, L. A. Is social learning an adaptive specialization? in *Social Learning in Animals: the Roots of Culture* (eds. Heyes, C. M. & Galef, B. G.) 107–128 (Academic Press, 1996). doi:10.1016/B978-012273965-1/50007-8.

205. Wilkinson, A., Kuenstner, K., Mueller, J. & Huber, L. Social learning in a non-social reptile (Geochelone carbonaria). *Biol. Lett.* **6**, 614–616 (2010).

206. McCune, K. B., Valente, J. J., Jablonski, P. G., Lee, S. & Ha, R. R. Social behavior mediates the use of social and personal information in wild jays. *Sci. Rep.* **12**, 1–11 (2022).

207. Heyes, C. M. Four Routes of Cognitive Evolution. *Psychol. Rev.* **110**, 713–727 (2003).

208. Leadbeater, E. What evolves in the evolution of social learning? *J. Zool.* **295**, 4–11 (2015).

209. Kamin, L. J. Predictability, surprise, attention, and conditioning. in *Punishment and Aversive Behavior.* (eds. Campbell, B. A. & Church, R. M.) 279–296 (Appleton-Century-Crofts, 1969).

210. Lind, J., Ghirlanda, S. & Enquist, M. Social learning through associative processes: a computational theory. *R. Soc. Open Sci.* **6**, (2019).

211. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science (80-.).* **275**, 1593–1599 (1997).

212. Montague, P. R., Dayan, P. & Sejnowski, T. J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).

213. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913 (1993).

214. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).

215. Frank, M. J. & O'Reilly, R. C. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* **120**, 497–517 (2006).

216. Menon, M. *et al.* Temporal Difference Modeling of the Blood-Oxygen Level Dependent Response During Aversive Conditioning in Humans: Effects of Dopaminergic Modulation. *Biol. Psychiatry* **62**, 765–772 (2007).

217. O'Doherty, J. P., Dayan, P., Friston, K. J., Critchley, H. & Dolan, R. J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329–337 (2003).

218. McClure, S. M., Berns, G. S. & Montague, P. R. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* **38**, 339–346 (2003).

219. D'Ardenne, K., McClure, S. M., Nystrom, L. & Cohen, J. BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area. *Science* **319**, 1264–1267 (2008).

220. Haarsma, J. *et al.* Precision weighting of cortical unsigned prediction error signals benefits learning, is mediated by dopamine, and is impaired in psychosis. *Mol. Psychiatry* **26**, 5320–5333 (2021).

221. Delgado, M. R., Frank, R. H. & Phelps, E. A. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* **8**, 1611–1618 (2005).

222. Diaconescu, A. O. *et al.* Hierarchical prediction errors in midbrain and septum during social learning. *Soc. Cogn. Affect. Neurosci.* **12**, 618–634 (2017).

223. Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A. & Fernández, G. Reinforcement learning signal predicts social conformity. *Neuron* **61**, 140–151 (2009).

224. Braams, B. R. *et al.* Reward-related neural responses are dependent on the beneficiary. *Soc. Cogn. Affect. Neurosci.* **9**, 1030–1037 (2014).

225. Campbell-Meiklejohn, D. K., Bach, D. R., Roepstorff, A., Dolan, R. J. & Frith, C. D. How the

opinion of others affects our valuation of objects. *Curr. Biol.* **20**, 1165–1170 (2010).

226. Joiner, J., Piva, M., Turrin, C. & Chang, S. W. C. Social learning through prediction error in the brain. *npj Sci. Learn.* **2**, 8 (2017).

227. Biele, G., Rieskamp, J., Krugel, L. K. & Heekeren, H. R. The neural basis of following advice. *PLoS Biol.* **9**, (2011).

228. Suzuki, S. *et al.* Learning to Simulate Others' Decisions. *Neuron* **74**, 1125–1137 (2012).

229. Burke, C. J., Tobler, P. N., Baddeley, M. & Schultz, W. Neural mechanisms of observational learning. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 14431–14436 (2010).

230. Diaconescu, A. O. *et al.* Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Comput. Biol.* **10**, e1003810 (2014).

231. De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T. & Love, B. C. Social information is integrated into value and confidence judgments according to its reliability. *J. Neurosci.* **37**, 6066–6074 (2017).

232. Rogers, R. D. The roles of dopamine and serotonin in decision making: Evidence from pharmacological experiments in humans. *Neuropsychopharmacology* **36**, 114–132 (2011).

233. Cohen, J. Y., Amoroso, M. W. & Uchida, N. Serotonergic neurons signal reward and punishment on multiple timescales. *Elife* **2015**, (2015).

234. Rogers, R. D. *et al.* Tryptophan depletion impairs stimulus-reward learning while methylphenidate disrupts attentional control in healthy young adults: Implications for the monoaminergic basis of impulsive behaviour. *Psychopharmacology (Berl).* **146**, 482–491 (1999).

235. Seymour, B., Daw, N. D., Roiser, J. P., Dayan, P. & Dolan, R. Serotonin selectively modulates reward value in human decision-making. *J. Neurosci.* **32**, 5833–5842 (2012).

236. den Ouden, H. *et al.* Dissociable Effects of Dopamine and Serotonin on Reversal Learning. *Neuron* **80**, 1090–1100 (2013).

237. Kanen, J. W. *et al.* Serotonin depletion impairs both Pavlovian and instrumental reversal learning in healthy humans. *Mol. Psychiatry* (2021) doi:10.1038/s41380-021-01240-9.

238. Fischer, A. G. & Ullsperger, M. An update on the role of serotonin and its interplay with dopamine for reward. *Front. Hum. Neurosci.* **11**, 484 (2017).

239. Matias, S., Lottem, E., Dugué, G. P. & Mainen, Z. F. Activity patterns of serotonin neurons underlying cognitive flexibility. *Elife* **6**, 1–24 (2017).

240. Kiser, D., SteemerS, B., Branchi, I. & Homberg, J. R. The reciprocal interaction between serotonin and social behaviour. *Neurosci. Biobehav. Rev.* **36**, 786–798 (2012).

241. Roberts, C., Sahakian, B. J. & Robbins, T. W. Psychological mechanisms and functions of 5-HT and SSRIs in potential therapeutic change: Lessons from the serotonergic modulation of action selection, learning, affect, and social cognition. *Neurosci. Biobehav. Rev.* **119**, 138–167 (2020).

242. Bethea, C. L. *et al.* Anxious Behavior and Fenfluramine-Induced Prolactin Secretion in Young Rhesus Macaques with Different Alleles of the Serotonin Reuptake Transporter Polymorphism (5HTTLPR). *Behav. Genet.* **34**, 295–307 (2004).

243. Homberg, J. R., Schiepers, O. J. G., Schoffelmeer, A. N. M., Cuppen, E. & Vanderschuren, L. J. M. J. Acute and constitutive increases in central serotonin levels reduce social play behaviour in peri-adolescent rats. *Psychopharmacology (Berl).* **195**, 175–182 (2007).

244. Suomi, S. J. Risk, resilience, and gene x environment interactions in rhesus monkeys. *Ann. N. Y. Acad. Sci.* **1094**, 52–62 (2006).

245. Watson, K. K., Ghodasra, J. H. & Platt, M. L. Serotonin transporter genotype modulates social reward and punishment in rhesus macaques. *PLoS One* **4**, (2009).

246. Higley, J. D. *et al.* CSF testosterone and 5-HIAA correlate with different types of aggressive behaviors. *Biol. Psychiatry* **40**, 1067–1082 (1996).

247. Raleigh, M. J., McGuire, M. T., Brammer, G. L., Pollack, D. B. & Yuwiler, A. Serotonergic mechanisms promote dominance acquisition in adult male vervet monkeys. *Brain Res.* **559**, 181–190 (1991).

248. Crockett, M. J. & Cools, R. Serotonin and aversive processing in affective and social decision-making. *Curr. Opin. Behav. Sci.* **5**, 64–70 (2015).

249. Beacher, F. D. C. C. *et al.* Acute tryptophan depletion attenuates conscious appraisal of social emotional signals in healthy female volunteers. *Psychopharmacology (Berl).* **213**, 603–613 (2011).

250. Crockett, M. J., Clark, L., Hauser, M. D. & Robbins, T. W. Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 17433–17438 (2010).

251. Wood, R. M., Rilling, J. K., Sanfey, A. G., Bhagwagar, Z. & Rogers, R. D. Effects of tryptophan depletion on the performance of an iterated Prisoner's Dilemma game in healthy adults. *Neuropsychopharmacology* **31**, 1075–1084 (2006).

252. Crockett, M. J. *et al.* Dissociable Effects of Serotonin and Dopamine on the Valuation of Harm in Moral Decision Making. *Curr. Biol.* **25**, 1852–1859 (2015).

253. Simonsen, A. *et al.* Serotoninergic effects on judgments and social learning of trustworthiness. *Psychopharmacology (Berl).* **231**, 2759–2769 (2014).

254. Frey, A. L. & McCabe, C. Effects of serotonin and dopamine depletion on neural prediction computations during social learning. *Neuropsychopharmacology* **45**, 1431–1437 (2020).

255. Crişan, L. G. *et al.* Genetic contributions of the serotonin transporter to social learning of fear and economic decision making. *Soc. Cogn. Affect. Neurosci.* **4**, 399–408 (2009).

256. Apps, M. A. J., Rushworth, M. F. S. & Chang, S. W. C. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron* **90**, 692–707 (2016).

257. Hill, M. R., Boorman, E. D. & Fried, I. Observational learning computations in neurons of the human anterior cingulate cortex. *Nat. Commun.* **7**, 12722 (2016).

258. Zhang, L. & Gläscher, J. A brain network supporting social influences in human decision-making. *Sci. Adv.* **6**, 1–20 (2020).

259. Carter, R. M., Bowling, D. L., Reeck, C. & Huettel, S. A. A Distinct Role of the Temporal-Parietal Junction in Predicting Socially Guided Decisions. **820**, 109–112 (2012).

260. Suzuki, S. & O'Doherty, J. P. Breaking human social decision making into multiple components and then putting them together again. *Cortex* **127**, 221–230 (2020).

261. Saxe, R. & Kanwisher, N. People thinking about thinking people. The role of the temporo-parietal junction in " theory of mind ". **19**, 1835–1842 (2003).

262. Ruff, C. C. & Fehr, E. The neurobiology of rewards and values in social decision makingThe neurobiology of rewards and values in social decision makingThe neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15**, 549–562 (2014).

263. Wittmann, M. K., Lockwood, P. L. & Rushworth, M. F. S. Neural mechanisms of social cognition in primates. *Annu. Rev. Neurosci.* **41**, 99–118 (2018).

264. Denny, B. T., Kober, H., Wager, T. D. & Ochsner, K. N. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* **24**, 1742–1752 (2012).

265. Nicolle, A. *et al.* An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron* **75**, 1114–1121 (2012).

266. Cook, J. L. Task-relevance dependent gradients in medial prefrontal and temporoparietal cortices suggest solutions to paradoxes concerning self/other control. *Neurosci. Biobehav. Rev.* **42**, 298–302 (2014).

267. Stins, J. F. & Emck, C. Balance performance in autism: A brief overview. *Front. Psychol.* **9**, 1–6 (2018).

268. Casartelli, L., Molteni, M. & Ronconi, L. So close yet so far: Motor anomalies impacting on social functioning in autism spectrum disorder. *Neurosci. Biobehav. Rev.* **63**, 98–105 (2016).

269. Sacrey, L. A. R., Germani, T., Bryson, S. E. & Zwaigenbaum, L. Reaching and grasping in autism spectrum disorder: A review of recent literature. *Front. Neurol.* **5**, 1–12 (2014).

270. Robertson, E. M. The serial reaction time task: Implicit motor skill learning? *J. Neurosci.* **27**, 10073–10075 (2007).

271. Karvelis, P., Seitz, A. R., Lawrie, S. M. & Seriès, P. Autistic traits, but not schizotypy, predict increased weighting of sensory information in bayesian visual integration. *Elife* **7**, (2018).

272. Fairweather, H. & Hutt, S. J. On the rate of gain of information in children. *J. Exp. Child Psychol.* **26**, 216–229 (1978).

273. Deary, I. J., Liewald, D. & Nissan, J. A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behav. Res. Methods* **43**, 258–268 (2011).

274. Deary, I. J., Der, G. & Ford, G. Reaction times and intelligence differences. A population-based cohort study. *Intelligence* **29**, 389–399 (2001).

275. Der, G. & Deary, I. J. IQ, reaction time and the differentiation hypothesis. *Intelligence* **31**, 491–503 (2003).

276. Zießler, M. The impact of motor responses on serial-pattern learning. *Psychol. Res.* **57**, 30–41 (1994).

277. Galea, J. M., Bestmann, S., Beigi, M., Jahanshahi, M. & Rothwell, J. C. Action reprogramming in Parkinson's disease: Response to prediction error is modulated by levels of dopamine. *J. Neurosci.* **32**, 542–550 (2012).

278. World Health Organisation. International classification of diseases (ICD). in (2012).

279. Lord, C. *et al. Autism Diagnostic Observation Schedule – Second Edition (ADOS-2). Manual (Part I): Modules 1–4.* (Western Psychological Services, 2012). doi:10.4135/9781483392271.n47.

280. Bagby, R. M., Taylor, G. J. & Parker, J. D. A. The twenty-item Toronto Alexithymia scale-II. Convergent, discriminant, and concurrent validity. *J. Psychosom. Res.* **38**, 33–40 (1994).

281. Dumont, R. & Willis, J. O. *Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II)*. (NCS Pearson, 2008). doi:10.1002/9780470373699.speced2202.

282. Bestmann, S., Ruge, D., Rothwell, J. & Galea, J. M. The role of dopamine in motor flexibility. *J. Cogn. Neurosci.* **27**, 365–376 (2014).

283. Townsend, J. T. & Ashby, F. G. Methods of modeling capacity in simple processing systems. in *Cognitive Theory Vol. III* 199–239 (Erlbaum Associates, 1978). doi:10.1163/_q3_SIM_00374.

284. Taylor, J. A. & Ivry, R. B. Flexible cognitive strategies during motor learning. *PLoS Comput. Biol.* **7**, (2011).

285. Taylor, J. A., Krakauer, J. W. & Ivry, R. B. Explicit and implicit contributions to learning in a sensorimotor adaptation task. *J. Neurosci.* **34**, 3023–3032 (2014).

286. Jiménez, L., Méndez, C. & Cleeremans, A. Comparing direct and indirect measures of sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **22**, 948–969 (1996).

287. Schvaneveldt, R. W. & Gomez, R. L. Attention and probabilistic sequence learning. *Psychol. Res.* **61**, 175–190 (1998).

288. Vandenberghe, M., Schmidt, N., Fery, P. & Cleeremans, A. Can amnesic patients learn without awareness? New evidence comparing deterministic and probabilistic sequence learning. *Neuropsychologia* **44**, 1629–1641 (2006).

289. Skewes, J. C. & Gebauer, L. Brief Report: Suboptimal auditory localization in Autism spectrum disorder: support for the Bayesian account of sensory symptoms. *J. Autism Dev. Disord.* **46**, 2539–2547 (2016).

290. Marshall, L. *et al.* Pharmacological fingerprints of contextual uncertainty. *PLoS Biol.* **14**, (2016).

291. Hardwick, R. M., Rottschy, C., Miall, R. C. & Eickhoff, S. B. A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage* **67**, 283–297 (2013).

292. Belmonte, M. K. *et al.* Autism and abnormal development of brain connectivity. *J. Neurosci.* **24**, 9228–9231 (2004).

293. Just, M. A., Cherkassky, V. L., Keller, T. A. & Minshew, N. J. Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain* **127**, 1811–1821 (2004).

294. Maximo, J. O., Cadena, E. J. & Kana, R. K. The implications of brain connectivity in the neuropsychology of autism. *Neuropsychol. Rev.* **24**, 16–31 (2014).

295. Courchesne, E. & Pierce, K. Why the frontal cortex in autism might be talking only to itself: Local over-connectivity but long-distance disconnection. *Curr. Opin. Neurobiol.* **15**, 225–230 (2005).

296. O'Reilly, C., Lewis, J. D. & Elsabbagh, M. Is functional brain connectivity atypical in autism?

A systematic review of EEG and MEG studies. *PLoS One* **12**, 1–28 (2017).

297. Cook, J. L., Barbalat, G. & Blakemore, S. J. Top-down modulation of the perception of other people in schizophrenia and autism. *Front. Hum. Neurosci.* **6**, 1–10 (2012).

298. Frith, C. What do imaging studies tell us about the neural basis of autism? *Novartis Found. Symp.* **251**, 149–176 (2003).

299. Beker, S., Foxe, J. J. & Molholm, S. Ripe for solution: Delayed development of multisensory processing in autism and its remediation. *Neurosci. Biobehav. Rev.* **84**, 182–192 (2018).

300. Feldman, J. I. *et al.* Audiovisual multisensory integration in individuals with autism spectrum disorder: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **95**, 220–234 (2018).

301. Chita-Tegmark, M. Attention allocation in ASD: A review and meta-analysis of eye-Tracking studies. *Rev. J. Autism Dev. Disord.* **3**, 209–223 (2016).

302. Stavropoulos, K. K. M. & Carver, L. J. Oscillatory rhythm of reward: Anticipation and processing of rewards in children with and without autism. *Mol. Autism* **9**, 1–15 (2018).

303. Schultz, R. T. Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *Int. J. Dev. Neurosci.* **23**, 125–141 (2005).

304. Dawson, G., Webb, S. J. & McPartland, J. Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Dev. Neuropsychol.* **27**, 403–424 (2005).

305. Bird, G., Press, C. & Richardson, D. C. The role of alexithymia in reduced eye-fixation in autism spectrum conditions. *J. Autism Dev. Disord.* **41**, 1556–1564 (2011).

306. Chita-Tegmark, M. Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Res. Dev. Disabil.* **48**, 79–93 (2016).

307. Dawson, G., Bernier, R. & Ring, R. H. Social attention: A possible early indicator of efficacy in autism clinical trials. *J. Neurodev. Disord.* **4**, 1–12 (2012).

308. Reisinger, D. L. *et al.* Atypical Social Attention and Emotional Face Processing in Autism Spectrum Disorder: Insights From Face Scanning and Pupillometry. *Front. Integr. Neurosci.* **13**, 1–14 (2020).

309. Cook, J. L., Den Ouden, H., Heyes, C. M. & Cools, R. The social dominance paradox. *Curr. Biol.* **24**, 2812–2816 (2014).

310. Worthy, D. A., Hawthorne, M. J. & Otto, A. R. Heterogeneity of strategy use in the Iowa gambling task: A comparison of win-stay/lose-shift and reinforcement learning models. *Psychon. Bull. Rev.* **20**, 364–371 (2013).

311. Crawley, D. *et al.* Modeling cognitive flexibility in autism spectrum disorder and typical development reveals comparable developmental shifts in learning mechanisms. *PsyArXiv* 10.31234/osf.io/h7jcm (2019).

312. Rushworth, M. F. S., Mars, R. B. & Sallet, J. Are there specialized circuits for social cognition and are they unique to humans? *Curr. Opin. Neurobiol.* **23**, 436–442 (2013).

313. Heyes, C. M. Blackboxing: Social learning strategies and cultural evolution. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, (2016).

314. Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T. & Hutchison, K. E. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16311–16316 (2007).

315. Doll, B. B., Bath, K. G., Daw, N. D. & Frank, M. J. Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J. Neurosci.* **36**, 1211–1222 (2016).

316. Cooper, J. C., Dunne, S., Furey, T. & O'Doherty, J. P. Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. *J. Cogn. Neurosci.* **24**, 106–118 (2012).

317. Lohrenz, T., Bhatt, M., Apple, N. & Montague, P. R. Keeping up with the Joneses: Interpersonal Prediction Errors and the Correlation of Behavior in a Tandem Sequential Choice Task. *PLoS Comput. Biol.* **9**, e1003275 (2013).

318. Barnett, J. H., Xu, K., Heron, J., Goldman, D. & Jones, P. B. Cognitive effects of genetic variation in monoamine neurotransmitter systems: A population-based study of COMT, MAOA, and 5HTTLPR. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **156**, 158–167 (2011).

319. Wilson, W. J. The Rescorla-Wagner Model , Simplified. *Unpublished* 1–4 (2012).

320. Mill, J., Asherson, P., Browes, C., Souza, U. D. ' & Craig, I. UTR VNTR: Evidence From Brain and Lymphocytes Using Quantitative RT-PCR. *Am. J. Med. Genet. (Neuropsychiatric Genet.* **114**, 975–979 (2002).

321. Spencer, T. J. *et al.* Functional genomics of attention-deficit/hyperactivity disorder (adhd) risk alleles on dopamine transporter binding in ADHD and healthy control subjects. *Biol. Psychiatry* **74**, 84–89 (2013).

322. Willeit, M. & Praschak-Rieder, N. Imaging the effects of genetic polymorphisms on radioligand binding in the living human brain: A review on genetic neuroreceptor imaging of monoaminergic systems in psychiatry. *Neuroimage* **53**, 878–892 (2010).

323. Matsumoto, M. *et al.* Catechol O-methyltransferase mRNA expression in human and rat brain: Evidence for a role in cortical neuronal function. *Neuroscience* **116**, 127–137 (2003).

324. Sesack, S. R., Hawrylak, V. A., Matus, C., Guido, M. A. & Levey, A. I. Dopamine axon varicosities in the prelimbic division of the rat prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter. *J. Neurosci.* **18**, 2697–2708 (1998).

325. Slifstein, M. *et al.* COMT genotype predicts cortical-limbic D1 receptor availability measured with [11C]NNC112 and PET. *Mol. Psychiatry* **13**, 821–827 (2008).

326. Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S. C. & Heekeren, H. R. Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17951–17956 (2009).

327. Greenberg, B. D. *et al.* Genetic variation in the serotonin transporter promoter region affects serotonin uptake in human blood platelets. *Am. J. Med. Genet. - Neuropsychiatr. Genet.* **88**, 83–87 (1999).

328. Heils, A., Mößner, R. & Lesch, K. P. The human serotonin transporter gene polymorphism-basic research and clinical implications. *J. Neural Transm.* **104**, 1005–1014 (1997).

329. Lesch, K. P. *et al.* Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science (80-.).* **274**, 1527–1531 (1996).

330. Canli, T. & Lesch, K. P. Long story short: The serotonin transporter in emotion regulation and social cognition. *Nat. Neurosci.* **10**, 1103–1109 (2007).

331. Rendell, L. *et al.* Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends Cogn. Sci.* **15**, 68–76 (2011).

332. Daw, N. D., Kakade, S. & Dayan, P. Opponent interactions between serotonin and dopamine. *Neural Networks* **15**, 603–616 (2002).

333. Cohen, J. Y. & Grossman, C. D. *Dorsal raphe serotonergic neurons regulate behavior on multiple timescales*. *Handbook of Behavioral Neuroscience* vol. 31 (Elsevier, 2020).

334. Doya, K., Miyazaki, K. W. & Miyazaki, K. Serotonergic modulation of cognitive computations. *Curr. Opin. Behav. Sci.* **38**, 116–123 (2021).

335. Clarke, H. F., Walker, S. C., Dalley, J. W., Robbins, T. W. & Roberts, A. C. Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cereb. Cortex* **17**, 18–27 (2007).

336. Boulougouris, V. & Robbins, T. W. Enhancement of spatial reversal learning by 5-HT2C receptor antagonism is neuroanatomically specific. *J. Neurosci.* **30**, 930–938 (2010).

337. Grossman, C. D., Bari, B. A. & Cohen, J. Y. Serotonin neurons modulate learning rate through uncertainty. *Curr. Biol.* **32**, 586-599.e7 (2022).

338. De Deurwaerdère, P. & Di Giovanni, G. Serotonergic modulation of the activity of mesencephalic dopaminergic systems: Therapeutic implications. *Prog. Neurobiol.* **151**, 175–236 (2017).

339. Hanna, G. L. *et al.* Serotonin transporter and seasonal variation in blood serotonin in families with obsessive-compulsive disorder. *Neuropsychopharmacology* **18**, 102–111 (1998).

340. Garbarino, V. R., Gilman, T. L., Daws, L. C. & Gould, G. G. Extreme enhancement or depletion of serotonin transporter function and serotonin availability in autism spectrum disorder. *Pharmacol. Res.* **140**, 85–99 (2019).

341. Anderson, G. M. *et al.* Serotonin transporter promoter variants in autism: Functional effects and relationship to platelet hyperserotonemia. *Mol. Psychiatry* **7**, 831–836 (2002).

342. Tunbridge, E. M. *et al.* Which Dopamine Polymorphisms Are Functional? Systematic Review and Meta-analysis of COMT, DAT, DBH, DDC, DRD1–5, MAOA, MAOB, TH, VMAT1, and VMAT2. *Biol. Psychiatry* **86**, 608–620 (2019).

343. Faraone, S. V., Spencer, T. J., Madras, B. K., Zhang-James, Y. & Biederman, J. Functional effects of dopamine transporter gene genotypes on in vivo dopamine transporter functioning: A meta-analysis. *Mol. Psychiatry* **19**, 880–889 (2014).

344. Van Dyck, C. H. *et al.* Increased dopamine transporter availability associated with the 9-repeat allele of the SLC6A3 gene. *J. Nucl. Med.* **46**, 745–751 (2005).

345. Heinz, A. *et al.* Genotype influences in vivo dopamine transporter availability in human striatum. *Neuropsychopharmacology* **22**, 133–139 (2000).

346. Campbell-Meiklejohn, D. K., Simonsen, A., Frith, C. D. & Daw, N. D. Independent neural computation of value from other people's confidence. *J. Neurosci.* **37**, 673–684 (2017).

347. Biele, G., Rieskamp, J. & Gonzalez, R. Computational models for the combination of advice and individual learning. *Cogn. Sci.* **33**, 206–242 (2009).

348. Shanks, D. R. Forward and backward blocking in human contingency judgement. *Q. J. Exp. Psychol. Sect. B* **37**, 1–21 (1985).

349. Kehoe, E. J., Schreurs, B. G. & Graham, P. Temporal primacy overrides prior training in serial compund conditioning of the rabbit's nictitating membrane response. *Anim. Learn. Behav.* **15**, 455–464 (1987).

350. Mackintosh, N. J. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* **82**, 276–298 (1975).

351. Pearce, J. M. & Hall, G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).

352. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 1–11 (2001).

353. Anderson, B. A. A value-driven mechanism of attentional selection. *J. Vis.* **13**, 1–16 (2013).

354. Hickey, C., Chelazzi, L. & Theeuwes, J. Reward changes salience in human vision via the anterior cingulate. *J. Neurosci.* **30**, 11096–11103 (2010).

355. Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward. *Brain Res. Rev.* **28**, 309–369 (1998).

356. Pessoa, L. & Engelmann, J. B. Embedding reward signals into perception and cognition. *Front. Neurosci.* **4**, (2010).

357. Doll, B. B., Jacobs, W. J., Sanfey, A. G. & Frank, M. J. Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Res.* **1299**, 74–94 (2009).

358. Atlas, L. Y., Doll, B. B., Li, J., Daw, N. D. & Phelps, E. A. Instructed knowledge shapes feedback- driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *Elife* **5**, 1–26 (2016).

359. Li, J., Delgado, M. R. & Phelps, E. A. How instructed knowledge modulates the neural systems of reward learning. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 55–60 (2011).

360. Koban, L., Jepma, M., Geuter, S. & Wager, T. D. What's in a word? How instructions, suggestions, and social information change pain and emotion. *Neurosci. Biobehav. Rev.* **81**, 29–42 (2017).

361. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).

362. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).

363. Chelazzi, L., Perlato, A., Santandrea, E. & Della Libera, C. Rewards teach visual selective attention. *Vision Res.* **85**, 58–72 (2013).

364. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).

365. Beck, D. M. & Kastner, S. Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Res.* **49**, 1154–1165 (2009).

366. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Psychol.* **68**, 47–72 (2017).

367. Garvert, M. M., Moutoussis, M., Kurth-Nelson, Z., Behrens, T. E. J. & Dolan, R. J. Learning-induced plasticity in medial prefrontal cortex predicts preference malleability. *Neuron* **85**, 418–428 (2015).

368. Ereira, S. *et al.* Social training reconfigures prediction errors to shape Self-Other boundaries. *Nat. Commun.* **11**, 1–14 (2020).

369. Behrens, T. E. J., Hunt, L. T. & Rushworth, M. F. S. The computation of social behavior. *Science (80-.).* **324**, 1160–1164 (2009).

370. Langdon, A. J., Sharpe, M. J., Schoenbaum, G. & Niv, Y. Model-based predictions for dopamine. *Curr. Opin. Neurobiol.* **49**, 1–7 (2018).

371. Bromberg-Martin, E. S., Matsumoto, M., Hong, S. & Hikosaka, O. A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J. Neurophysiol.* **104**, 1068–1076 (2010).

372. Sharpe, M. J. & Schoenbaum, G. Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. *Neurobiol. Learn. Mem.* **153**, 131–136 (2018).

373. Sadacca, B. F., Jones, J. L. & Schoenbaum, G. Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *Elife* **5**, 1–13 (2016).

374. Jones, J. L. *et al.* Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science (80-.).* **338**, 953–956 (2012).

375. Moran, R., Dayan, P. & Dolan, R. J. Human subjects exploit a cognitive map for credit assignment. *Proc. Natl. Acad. Sci. U. S. A.* **118**, 1–12 (2021).

376. Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–325 (2013).

377. Brazil, I. A. *et al.* Psychopathy-related traits and the use of reward and social information: A computational approach. *Front. Psychol.* **4**, 1–11 (2013).

378. Tarantola, T., Kumaran, D., Dayan, P. & De Martino, B. Prior preferences beneficially influence social and non-social learning. *Nat. Commun.* **8**, 817 (2017).

379. Schultz, W. Behavioral dopamine signals. *Trends Neurosci.* **30**, 203–210 (2007).

380. Gläscher, J., Daw, N., Dayan, P. & Doherty, J. P. O. States versus Rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**, 585–595 (2011).

381. Camps, M., Cortés, R., Gueye, B., Probst, A. & Palacios, J. M. Dopamine receptors in human brain: Autoradiographic distribution of D2 sites. *Neuroscience* **28**, 275–290 (1989).

382. Lidow, M. S., Goldman-Rakic, P. S., Gallager, D. W. & Rakic, P. Distribution of dopaminergic receptors in the primate cerebral cortex: Quantitative autoradiographic analysis using [3H]raclopride, [3H]spiperone and [3H]SCH23390. *Neuroscience* **40**, 657–671 (1991).

383. Grace, A. A. Dopamine. in *Neuropsychopharmacology: The Fifth Generation of Progress* 120–132 (2002).

384. Dienes, Z. Using Bayes to get the most out of non-significant results. *Front. Psychol.* **5**, 1–17 (2014).

385. Lee, M. D. & Wagenmakers, E. J. *Bayesian cognitive modeling: A practical course*. (Cambridge University Press, 2013). doi:10.1017/CBO9781139087759.

386. Lindström, B., Haaker, J. & Olsson, A. A common neural network differentially mediates

direct and social fear learning. *Neuroimage* **167**, 121–129 (2018).

387. Sulzer, D., Cragg, S. J. & Rice, M. E. Striatal dopamine neurotransmission: Regulation of release and uptake. *Basal Ganglia* **6**, 123–148 (2016).

388. Korn, C. *et al.* Distinct roles for dopamine clearance mechanisms in regulating behavioral flexibility. *Mol. Psychiatry* (2021) doi:10.1038/s41380-021-01194-y.

389. Yael, D. *et al.* Haloperidol-induced changes in neuronal activity in the striatum of the freely moving rat. *Front. Syst. Neurosci.* **7**, 1–11 (2013).

390. Wächtler, C. O. *et al.* Region-specific effects of acute haloperidol in the human midbrain, striatum and cortex. *Eur. Neuropsychopharmacol.* **35**, 126–135 (2020).

391. Schmitz, Y., Benoit-Marand, M., Gonon, F. & Sulzer, D. Presynaptic regulation of dopaminergic neurotransmission. *J. Neurochem.* **87**, 273–289 (2003).

392. Benoit-Marand, M., Borrelli, E. & Gonon, F. Inhibition of dopamine release via presynaptic D2 receptors: Time course and functional characteristics in vivo. *J. Neurosci.* **21**, 9134–9141 (2001).

393. Ford, C. P. The role of D2-autoreceptors in regulating dopamine neuron activity and transmission. *Neuroscience* **282**, 13–22 (2014).

394. Grace, A. A. Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nat. Rev. Neurosci.* **17**, 524–532 (2016).

395. Belujon, P. & Grace, A. A. Regulation of dopamine system responsivity and its adaptive and pathological response to stress. *Proc. R. Soc. B Biol. Sci.* **282**, (2015).

396. Diederen, K. M. J. *et al.* Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *J. Neurosci.* **37**, 1708–1720 (2017).

397. Dugast, C., Brun, P., Sotty, F., Renaud, B. & Suaud-Chagny, M. F. On the involvement of a tonic dopamine D2-autoinhibition in the regulation of pulse-to-pulse-evoked dopamine release in the rat striatum in vivo. *Naunyn. Schmiedebergs. Arch. Pharmacol.* **355**, 716–719 (1997).

398. Garris, P. A. *et al.* A role for presynaptic mechanisms in the actions of nomifensine and haloperidol. *Neuroscience* **118**, 819–829 (2003).

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

400. Sternberg, S. Memory-scanning: mental processes revealed by reaction-time experiments. *Am. Sci.* **57**, 421–457 (1969).

401. Rouder, J. N., Morey, R. D., Speckman, P. L. & Province, J. M. Default Bayes factors for ANOVA designs. *J. Math. Psychol.* **56**, 356–374 (2012).

402. Van Den Bergh, D. *et al.* A tutorial on conducting and interpreting a bayesian ANOVA in JASP. *Annee Psychol.* **120**, 73–96 (2020).

403. Rybicki, A. J. *et al.* Intact predictive motor sequence learning in autism spectrum disorder. *Sci. Rep.* **11**, 1–11 (2021).

404. Ego, C. *et al.* Behavioral characterization of prediction and internal models in adolescents with autistic spectrum disorders. *Neuropsychologia* **91**, 335–345 (2016).

405. Karaminis, T. *et al.* Atypicalities in perceptual adaptation in autism do not extend to perceptual causality. *PLoS One* **10**, 1–16 (2015).

406. Manning, C., Neil, L., Karaminis, T. & Pellicano, E. The effects of grouping on speed discrimination thresholds in adults, typically developing children, and children with autism. *J. Vis.* **15**, (2015).

407. Spanò, G., Peterson, M. A., Nadel, L., Rhoads, C. & Edgin, J. O. Seeing can be remembering: Interactions between memory and perception in typical and atypical development. *Clin. Psychol. Sci.* **4**, 254–271 (2016).

408. Gidley Larson, J. C., Bastian, A. J., Donchin, O., Shadmehr, R. & Mostofsky, S. H. Acquisition of internal models of motor tasks in children with autism. *Brain* **131**, 2894–2903 (2008).

409. Mathys, C. *et al.* Uncertainty in perception and the Hierarchical Gaussian filter. *Front. Hum. Neurosci.* **8**, 825 (2014).

410. Cannon, J., O'Brien, A. M., Bungert, L. & Sinha, P. Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. *Autism Res.* **14**, 604–630 (2021).

411. Deschrijver, E., Bardi, L., Wiersema, J. R. & Brass, M. Behavioral measures of implicit theory of mind in adults with high functioning autism. *Cogn. Neurosci.* **7**, 192–202 (2016).

412. Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G. & Jones, W. Two-year-olds with autism orient to nonsocial contingencies rather than biological motion. **459**, 257–261 (2009).

413. Blake, R. *et al.* Visual Recognition of Biological Motion Is Impaired in Children with Autism Published by : Sage Publications , Inc . on behalf of the Association for Psychological Science Stable URL : http://www.jstor.org/stable/40063786 Accessed : 08-04-2016 13 : 48 UTC. **14**, 151–157 (2016).

414. Wagenmakers, E. J. *et al.* Bayesian inference for psychology. Part II: Example applications with JASP. *Psychon. Bull. Rev.* **25**, 58–76 (2018).

415. Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group studies - Revisited. *Neuroimage* **84**, 971–985 (2014).

416. Loomes, R., Hull, L. & Mandy, W. P. L. What Is the Male-to-Female Ratio in Autism Spectrum Disorder? *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 466–474 (2017).

417. Gonzalez-Gadea, M. L. *et al.* Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *J. Neurophysiol.* **114**, 2625–2636 (2015).

418. Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 1–23 (2010).

419. Rybicki, A. J., Sowden, S. L., Schuster, B. . & Cook, J. L. Dopaminergic challenge dissociates learning from primary versus secondary sources of information. *Elife* **11:e74893**, (2022).

420. Rogers, R. D. The roles of dopamine and serotonin in decision making: Evidence from pharmacological experiments in humans. *Neuropsychopharmacology* **36**, 114–132 (2011).

421. Schweighofer, N. *et al.* Low-serotonin levels increase delayed reward discounting in humans. *J. Neurosci.* **28**, 4528–4532 (2008).

422. Alcalá-López, D. *et al.* Computing the social brain connectome across systems and states. *Cereb. Cortex* **28**, 2207–2232 (2018).

423. Piva, M. *et al.* The dorsomedial prefrontal cortex computes task-invariant relative subjective value for self and other. *Elife* **8**, 1–29 (2019).

424. Strange, B. A., Duggins, A., Penny, W., Dolan, R. J. & Friston, K. J. Information theory, novelty and hippocampal responses: Unpredicted or unpredictable? *Neural Networks* **18**, 225–230 (2005).

425. Harrison, L. M., Duggins, A. & Friston, K. J. Encoding uncertainty in the hippocampus. *Neural Networks* **19**, 535–546 (2006).

426. Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D. & Iverson, G. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* **16**, 225–237 (2009).

427. van Doorn, J. *et al.* The JASP Guidelines for Conducting and Reporting a Bayesian Analysis. *Psychon. Bull. Rev.* (2020) doi:10.3758/s13423-020-01798-5.

428. Wagenmakers, E. J. A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* **14**, 779–804 (2007).

429. Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.* **6**, 65–70 (1979).

430. Abdi, H. Holm's Sequential Bonferroni Procedure. in *Encyclopedia of Research Design* (ed. Salkind, N.) (Sage, 2010).

431. Daunizeau, J., Adam, V. & Rigoux, L. VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. *PLoS Comput. Biol.* **10**, (2014).

432. Hu, X. Z. *et al.* Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *Am. J. Hum. Genet.* **78**, 815–826 (2006).

433. Wendland, J. R., Martin, B. J., Kruse, M. R., Lesch, K. P. & Murphy, D. L. Simultaneous genotyping of four functional loci of human SLC6A4, with a reappraisal of 5-HTTLPR and rs25531. *Mol. Psychiatry* **11**, 224–226 (2006).

434. Carver, C. S. & White, T. L. Behavioral Inhibition, Behavioral Activation, and Affective Responses to Impending Reward and Punishment: The BIS/BAS Scales. *J. Pers. Soc. Psychol.* **67**, 319–333 (1994).

435. Lovibond, P. F. & Lovibond, S. H. *Manual for the Depression Anxiety Stress Scales.* (1995).

436. Davis, M. H. A Mulitdimensional Approach to Individual Differences in Empathy. *J. Pers. Soc. Psychol.* **44**, 113–126 (1983).

437. Beck, A. T., Steer, R. A. & Brown, G. Beck Depression Inventory-II. in (APA PsycTests, 1996).

438. Porges, S. W. Body Perception Questionnaire (BPQ) Manual. *Stress Int. J. Biol. Stress* 1–7 (1993).

439. Watson, D., Clark, L. & Tellegen, A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* **54**, 1063–1070 (1988).

440. Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J. & Friston, K. J. Bayesian model selection for group studies. *Neuroimage* **46**, 1004–1017 (2009).

441. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology (Berl).* **191**, 507–520 (2007).

442. Cools, R., Gibbs, S. E., Miyakawa, A., Jagust, W. & D'Esposito, M. Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J. Neurosci.* **28**, 1208–1212 (2008).

Appendix 1

Supplementary material for Chapter 2

Intact predictive motor sequence learning in autism spectrum disorder

Rybicki, A. J., Galea, J. M., Schuster, B. A., Hiles, C., Fabian, C., & Cook, J. L.

Scientific Reports **11**:20693 (2021). DOI: 10.1038/s41598-021-00173-1

Methods

Participant inclusion criteria

All participants were fluent English speakers, with no history of neurological disease, muscular dystrophy or cerebral palsy, and had an intelligence quotient (IQ) score, measured using the two-subscale measure of the Wechsler Abbreviated Scale of Intelligence- 2 (WASI-2; Wechsler, D., 2011), greater than 75.

Autism spectrum group clinical information

Sixteen (57%) of the participants reported direct family history of developmental, neurological or severe psychiatric disorders. Six subjects had an additional diagnosis of a psychiatric disorder: one with bipolar disorder, one with epilepsy, two with attention deficit disorder (ADHD), three with generalized anxiety disorder and two with major depression. Sixteen (57%) of the sample were on prescribed psychiatric medication, with antidepressants the most commonly prescribed medication (twelve subjects), followed by ADHD medication (three subjects) and epilepsy (one subject). Six subjects (21%) had a learning disorder and two (7%) reported developmental delay. 2-subscale IQ scores ranged from 78-142 (mean (SD) = 108.68 (16.31) All analyses were repeated with the exclusion of sub-groups based on clinical information (see Supplementary results).

Serial reaction time behavioural task

Participants completed a computerised motor sequence learning task run on MATLAB 2018b (MathWorks, Inc.; http://www.mathworks.co.uk/) with Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent_2000) for which they sat in front of a computer screen (approx. 30 cm distance from screen), with the keyboard placed in front of their dominant hand. They were instructed to place each of their fingers separately on the keyboard letters V, B, N and M and to maintain this position throughout the task. A Razer DeathStalker keyboard, with ultra-polling at 1000Hz enabled, was used for precise measurement of RT.

For each individual trial, the following events were presented to the participant: A warning cue was displayed for 250 milliseconds (ms), followed by a fixation cross (1000 ms). After this, one of four imperative stimuli (IS) were shown in the centre of the screen for 250 ms (Fig. 1a). During this time, the participant was required to respond to the displayed stimuli. Stimulus presentation was followed by the fixation cross (2500 ms). Each IS was associated with a specific keyboard key, which corresponded to a specific finger press action (V, B, N, or M). Participants were required to learn the association between the IS and action (e.g., middle finger press) and execute the associated action when presented with an IS. Participants were instructed to respond as fast as possible, without sacrificing accuracy. Participants had to complete at least two training blocks of 60 trials in which they scored over 90% accuracy, before they could progress to the main experiment. If performance on the second training block was less than 90% correct, another block was added, this procedure was repeated until performance exceeded 90% correct. During the training blocks, participants received textual on-screen feedback as to whether their response was correct ("Correct!!!") or incorrect ("Wrong"). During training, stimuli were presented in an unpredictable order, with equal probabilities (25%) of each stimulus appearing on each trial. After the training round, they were asked to respond as quickly and accurately as possible to the presented symbols. They were shown the following instructions on the computer screen before starting the task:

"The experiment will start soon. Please remember to stay focused all the time! Please get ready when you see the warning cue, this will allow you to respond fast and accurately"

In the main experiment, participants completed seven blocks of 100 trials with self-paced rest intervals between the blocks. Feedback was not given in the main experiment. In the main experiment, IS order followed different sequences depending on condition, with predictable and unpredictable sequences presented in different blocks (Fig. 1b). For blocks one, four and seven (Fig. 1c) the stimulus sequence was unpredictable, with an equal probability (0.25) of each stimulus appearing on each trial. For blocks two, three, five and six, stimulus presentation followed a predictable pattern (see Fig. 1d for the easy predictable probabilities). Each stimulus was drawn from a predictable first-order Markov sequence, where the current stimulus t was dependent on the stimulus presented at the previous trial, t-1. Therefore,

predictable sequences were generated by sampling from a distribution specified in a transition matrix which quantified the dependence between stimuli. For the first predictable condition (Fig. 1d), the sequence followed an easy pattern in which IS order 1-2-3-4 (corresponding to the keyboard presses V-B-N-M) occurred with high probability, requiring the participant to respond with the natural order of the fingers, i.e., index, middle, ring and little finger. For the difficult-predictable condition stimuli followed a less natural predictable pattern whereby the stimuli order 1-4-2-3 (V-M-B-N) occurred with high probability.

Trial-by-trial surprise

Surprise was quantified on a trial-by-trial basis in a stimulus-specific manner, with the surprise (S) of observing a stimulus *i* on trial *t* after experiencing stimulus type *j* on trial *t*-1 calculated using the assumption that subjects behaved as "ideal" observers, beginning each block with the prior expectation of all stimulus pairings being equally probable, and updating the conditional probability of each pairing using a Bayesian update scheme [424],[425]. On each trial (*t*), subjects were presented with one of four IS, with the conditional probability of an IS on a given trial estimated from the previous occurrences of IS on the preceding trials. Specifically, the conditional probability (*E*) of an IS at trial *t*, $p(E_t)$, was estimated from the number of occurrences of IS *i* up to trial *t* (n_i^t, where *i* indexes the IS type and *t* the trial number) (Equation 1). Thus, the estimate of *i* at a given trial *t* was defined by:

$$p_t(E_t = i) = \frac{n_i^t + 1}{\sum_i (n_i^t + 1)}, (p_0 E_0 = i) = \frac{1}{4} \qquad \text{(Equation 1)}$$

Due to the probabilistic structure of the first-order Markov sequence, the IS occurring on the previous trial (*t*-1) could be used to form the prediction for the IS on trial *t*, allowing an approximation of the joint probability distribution for each IS pair to be estimated from the count of previous occurrences of the IS pair up to trial *t* ($n_{ij,}^t$ *i* = current IS type and *j* = previous IS type, Equation 2).

$$p_t(E_t = i, E_{t-1} = j) = \frac{n_{ij}^t + 1}{\sum_{i,j}\left(n_{ij}^t + 1\right)} \quad \text{(Equation 2)}$$

The surprise (*S*) of observing IS type *i* on trial *t* after experiencing IS type *j* on trial *t*-1 was therefore calculated as the negative log of the IS pair's predicted joint probability (Equation 3).

$$S(E_t = i, E_{t-1} = j) = -\log 2\left(p(E_t = i, E_{t-1} = j)\right) \quad \text{(Equation 3)}$$

Surprise was therefore stimulus-specific, representing the unexpectedness of the current IS, given the IS at trial *t*-1 and was high when an IS pairing was infrequent and low when the paring was frequent or occurred at a high probability (Fig. S1). Surprise was low overall during the predictable sequences, with occasional violations when unlikely surprising IS pairs occurred. In the unpredictable blocks, all events were equally as surprising as stimulus-pair probabilities were all fixed at 0.25.

Inverse Efficiency scores

Due to the extensive practice period, error rate was low across all conditions. Further to this, reaction time and accuracy were not independent in this task, as participants were required to respond within a narrow time frame. As this could lead to a speed-accuracy trade-off, whereby variable measures could lead to contradictory conclusions about the effect of group, we included a measure which combined speed (inverse RT) and accuracy. The inverse efficiency score (IES) [283] was utilized. This measure divided RT by 1 minus the proportion of errors (PE) or the proportion of correct responses (Equation 4). IES scores were calculated for each condition and for surprising and unsurprising trials separately.

$$IES = \frac{RT}{(1 - PE)}$$ (Equation 4)

Bayesian statistical testing

Bayesian statistical testing was implemented as a supplement to null hypothesis significance tests, to compare the likelihood of the data under the null and alternative hypothesis and provide an estimate for the amount of evidence represented in the data, in order to investigate if null results represent a true lack of a difference between the groups [384]. Bayesian statistical testing was implemented in JASP, based on the R package "BayesFactor" [401].

A Bayesian t-test framework [426] was used to determine if there was a difference in surprise-related slowing between the groups, for both the easy and difficult conditions. A Bayes factor, comparing the fit of data under the null hypothesis and the alternative hypothesis was estimated [427],[428] for each condition, whereby the null hypothesis (H_0) postulates that there are no differences between groups for surprise-related slowing scores. A two-sided alternative hypothesis was used, allowing the effect size (δ) to take both positive and negative values, with a default Cauchy prior distribution for a two-sample t-test, specifically, a zero-centred Cauchy distribution with a scale of 0.707 [427].

Bayesian repeated measures ANOVAs were utilized to investigate the effects of surprise, condition and group on RT and IES scores. The JASP framework for repeated measures ANOVA was used [402], whereby candidate models M and their condition-effect parameters β were compared, resulting in Bayes factors for each candidate model, quantifying the relative predictive performance of different models, as well as inclusion Bayes factors for predictors of interest. The inclusion Bayes factor (BF_{incl}) for a given predictor quantifies the change in odds from the prior probability that the predictor is included in the model to the probability of inclusion in the model after seeing the data, and were computed by comparing all models with a predictor against all models without that predictor, i.e., comparing models that contain the effect of interest to equivalent models stripped of the effect. For example, an inclusion Bayes factor for an effect of 3 for a given predictor *i* can be interpreted as stating that models which include the predictor *i* are 3 times more likely to describe the observed data than models without the predictor. In short, the inclusion Bayes factor is interpreted as the evidence given the observed data for including a certain predictor in the model. The inverse

of this, the Bayes exclusion factor (BF_{excl}) therefore represents the evidence for excluding a certain predictor. For example, a BF_{excl} value of 3 for a given predictor *j* means that models that *exclude* the predictor *j* are 3 times more likely given the data than models which include it. For all Bayesian analyses, the Bayes factor quantifies the relative evidence for one theory or model over another. We followed the classification scheme used in JASP [385] to classify the strength of evidence given by the Bayes factors, with BF_{01} between one and three considered as weak evidence, between three and ten as moderate evidence and greater than ten as strong evidence for the null hypothesis respectively. An annotated .jasp file containing all analysis is available at https://osf.io/cax4g/.

Results

AQ and TAS scores as a predictor of surprise-related slowing

As AQ and TAS scores differed significantly between the groups, we conducted separate linear regression analyses to investigate the potential effect of AQ and TAS score severity on surprise-related slowing. Results demonstrated that neither AQ ($\beta = 0.031$, $R^2 = 0.001$, $t(62) = 0.241$, $p = 0.810$) nor TAS scores ($\beta = 0.019$, $R^2 = 0.000$, $t(62) = 0.147$, $p = 0.884$) predicted the extent of surprise-related slowing. Repeated the above analyses with sequence learning as the dependent variable yielded similar results: neither AQ ($\beta = 0.139$, $R^2 = 0.019$, $t(62) = 1.096$, $p = 0.277$) nor TAS scores ($\beta = 0.115$, $R^2 = 0.115$, $t(62) = 0.900$, $p = 0.371$) predicted the extent of sequence learning.

Comorbidity and Medication

We re-ran all analyses with the exclusion of participants with current psychiatric medication use (n = 16). There were no significant differences between for any of the main measures of interest between the full sample and the sample excluding medication use, including sequence learning (all main/interaction effect(s) of group: all p values > 0.05, all $\eta^2 < 0.01$) and surprise-related slowing (all main/interaction effect(s) of group: all p values > 0.05, all $\eta^2 < 0.01$). Repeating the analyses with the exclusion of participants with comorbid psychiatric conditions (n = 7) yielded similar results (all main/interaction effect(s) of group: all *p* values > 0.05, all $\eta^2 < 0.01$).

Motor execution

Performance in the unpredictable blocks enable us to obtain a measure of participants' motor ability, as, since there are no sequences to learn in the unpredictable blocks, any differences in speed/error are likely to reflect motor execution ability. No significant differences were found between CTRL and ASD in RT (unpaired t-test: $t(61) = 0.951$, $p = 0.345$, $d = 0.241$, $BF_{01} = 2.642$) or IES (unpaired t-test: $t(61) = 0.513$, $p = 0.610$, $d = 0.130$, $BF_{01} = 3.462$) when we examined unpredictable trials (averaged across 3 blocks). Bayesian independent samples t-tests supported these results, with $BF_{01} = 2.642$ providing anecdotal evidence for no difference in RT between groups, and $BF_{01} = 3.462$ providing moderate evidence for no difference in IES scores between groups. We then investigated whether motor execution (RT/IES in unpredictable blocks) was correlated with other measures. No correlation was found between RT in the unpredictable blocks and sequence-learning in either the easy ($r = 0.211$, $p = 0.098$) or difficult ($r = -0.012$, $p = 0.924$) predictable blocks. In addition, there was no correlation between RT in the unpredictable blocks and the extent of surprise-related slowing in either the easy ($r = 0.164$, $p = 0.200$) or difficult ($r = -0.032$, $p = 0.805$) predictable blocks, suggesting that our measures (surprise-related slowing and sequence learning) are independent of motor execution.

Fig. S1

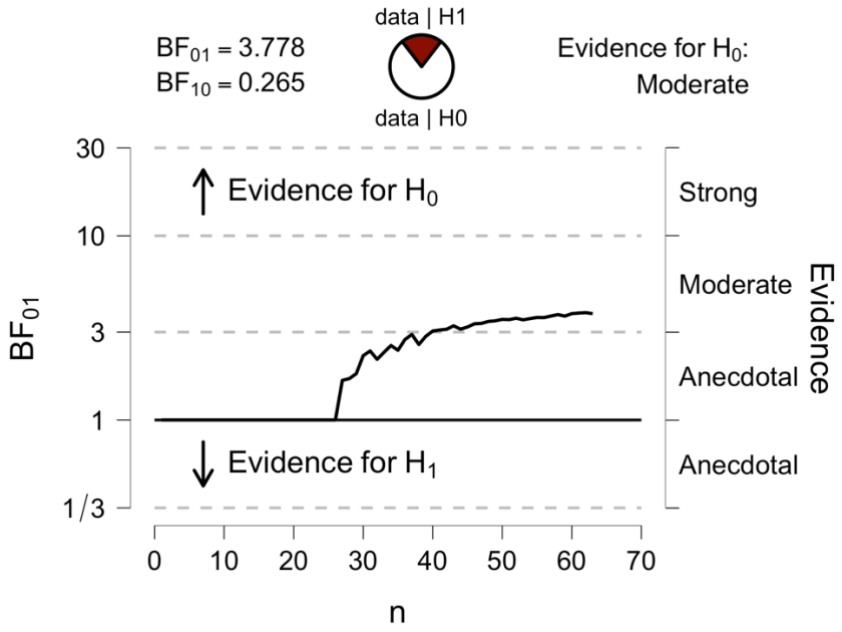


Fig. S1. Sequential analysis. Sequential analysis represents a visualisation of evidence as data are collected. The black line represents accumulation of evidence for the null hypothesis (BF_{01}) that groups do not differ in the extent of surprise-related slowing during the easy-predictable condition.

Fig. S2

(a)

(b)

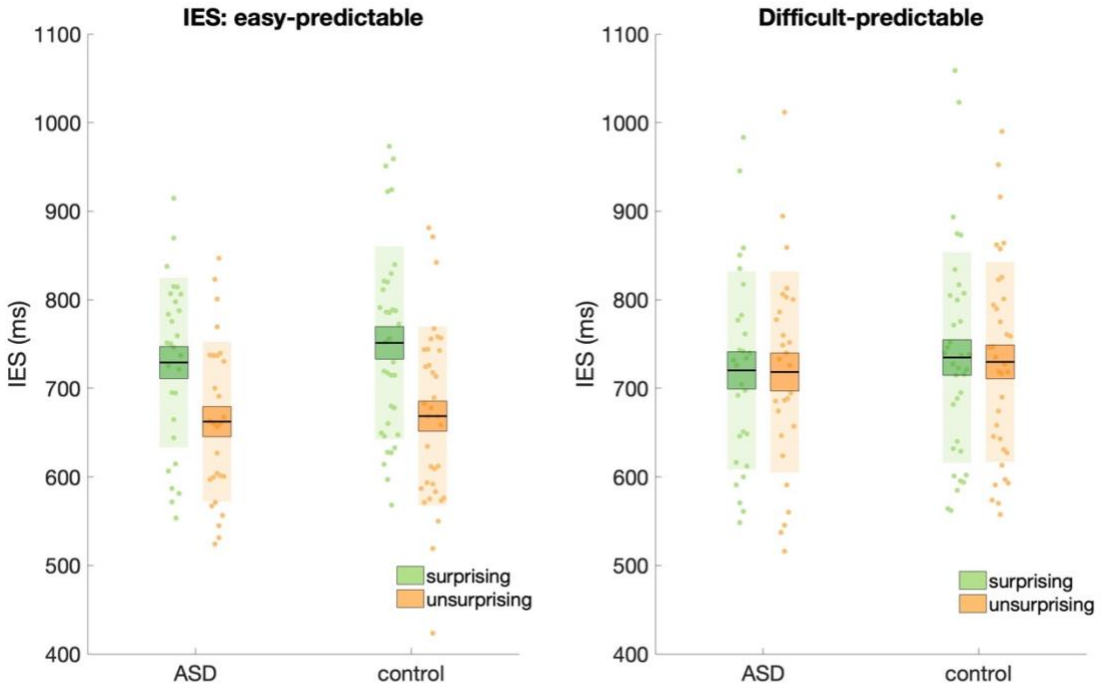


Fig. S2. IES scores. IES scores were lower for unsurprising trials for the easy but not the difficult predictable conditions. No differences were observed between groups. Data points indicate individual participants. The mean is the thick black horizontal line, and 1 standard error of the mean (SE) is represented by the shaded box around the mean. Standard deviation (SD) is the shaded region.

Table S1. Clinical information for the autistic individuals

	Total sample (%) **N=28**
Major depression	2 (7%)
Generalized anxiety disorder	3 (11%)
Bipolar disorder	1 (4%)
ADHD	3 (11%)
Epilepsy	1 (3%)
No current disorder*	22 (78%)

Appendix 2

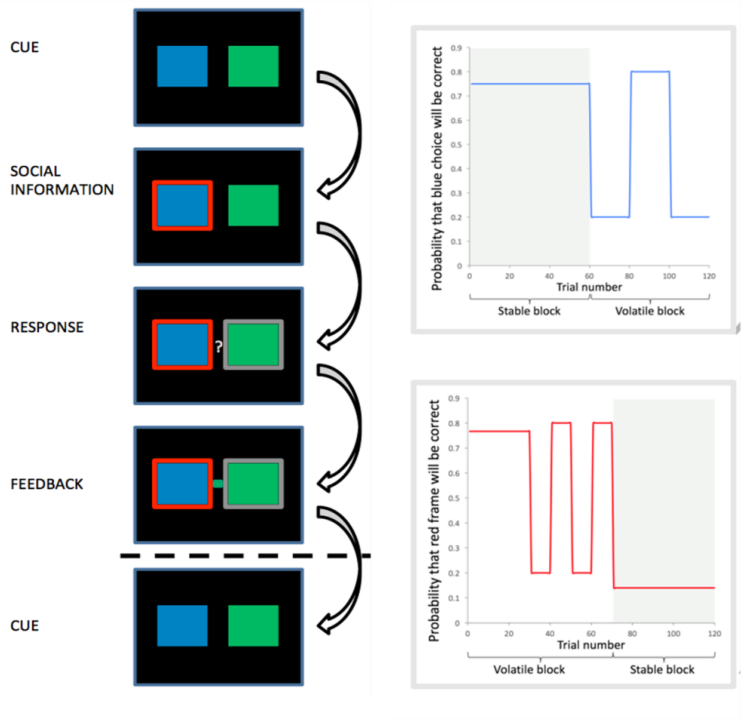
Supplementary methods and supplementary results for Chapter 3

2.1 Social Learning Task

The behavioural task (social learning task (SLT)) lasted approximately 35 minutes. Participants were seated approximately 30cm from a computer screen. Stimuli were displayed using PsychToolBox and the task was programmed using MATLAB R2017b (The MathWorks, Natick, MA). Before the main task, participants completed a step-by-step on-screen practice task (10 trials) in which they learnt to choose between the two options to obtain a reward and learned that the "advice" represented by the frame(s) could help in making the correct choice in some phases. To ensure that participants were making a conceptual distinction between the social and individual learning sources, participants were required to complete a short pre-task quiz (Appendix 2.1.1), testing their knowledge, after the practice task. Participants were required to repeat the practice round until they achieved 100% correct score in the quiz, meaning that all participants understood the structure of the task and that the red shape represented *social* information. Participants were informed as to whether they had earned a £5 bonus after the session. However, due to ethical considerations, all participants received the bonus.

On each trial participants were required to make a choice between a blue and green box in order to win points. Participants could also use an additional, secondary, source of information - a red frame surrounding either the blue or green box – to help make their decision. Participants were informed that the frame represented the most popular choice made by a group of participants who had previously completed the task. They were also informed that the task followed 'phases' wherein sometimes the blue, but at other times the green choice, was more likely to result in reward and sometimes the social information predominantly indicated the correct box, but at other times it predominantly surrounded the incorrect box (Suppl. Fig. 2.1A). After making their choice participants received outcome information in the form of a blue or green indicator. The indicator primarily informed participants about whether the blue or green box had been rewarded on the current trial. Whether the social information surrounded the correct or incorrect box could, secondarily, be inferred from the indicator. For example, if the red frame indicated that the social group had

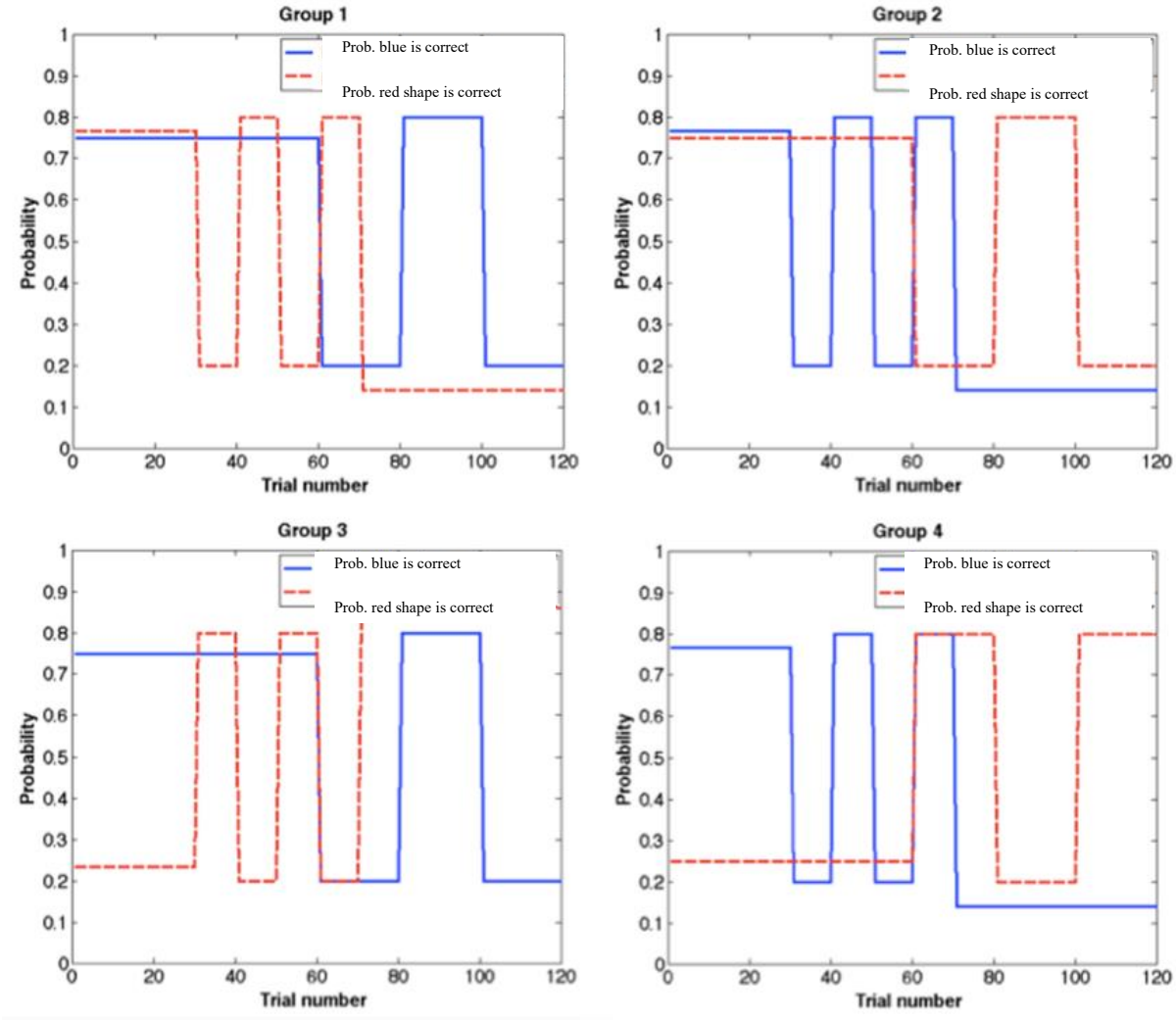
chosen the blue shape, and the blue shape was shown to be correct, participants could infer that the social information had therefore been correct on that trial. Both the probability of reward associated with the blue/green stimuli and the utility of the social information, varied according to separate probabilistic schedules, with participants randomly assigned to one of four groups (Appendix 2.1.1). All schedules included stable phases, where the probability of reward was constant for 50-60 trials, and volatile phases, wherein probabilities reversed every 10–30 trials, with outcomes determined according to these schedules, for both the probability of blue being correct and the probability of the red shape indicating the correct answer. After making their selection, participants saw a small blue or green box which informed them whether a blue or green choice had been rewarded on the current trial. From this information the participant inferred whether the social information (red frame) was correct or incorrect.



Supplementary Figure 2.1. Social learning task. **A**. Participants selected between a blue and a green box to gain points. On each trial, the blue and green boxes were presented first. After 1-4 seconds (s), one of the boxes was highlighted with a red frame, representing the social information. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a green or blue box in the middle of the screen. **B.** Example of pseudo-randomised probabilistic schedule. The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of the blue box/frame being correct (top) and the probability of the red (social) box/frame being correct (bottom).

2.1.1 Randomisation groups

Reward outcomes (blue/green correct) and the veracity of social information (correct/incorrect), were governed by four different pseudo-randomisation schedules (adapted from Behrens et al [95], with participants randomly assigned to one of the four groups. Both individual reward information and social information varied separately between stable phases, where the probability of reward was constant, and volatile phases, in which the probability switched every 10-20 trials. Participants were informed that correct choices would be rewarded, and thus to aim to accumulate points to obtain a reward. For example, the randomisation schedule for group 1 was the same as that employed by Behrens et al [95]. During the first 60 trials, the individual reward history was stable, with a 75% probability of blue being correct. During the next 60 trials, the reward history was volatile, switching between 80% green correct and 80% blue correct every 20 trials. Meanwhile, during the first 30 trials, social information was stable, with 75% of choices being correct. During the next 40 trials, the social information was volatile, switching between 80% incorrect and 80% correct every 10 trials. During the final 50 trials, social information was once again stable, with 85% of choices being incorrect. Randomisation schedules for groups 2, 3, and 4 were inverted and counterbalanced versions of schedule 1 (Suppl. Fig. 2.2).



Supplementary Figure 2.2. Randomisation schedules. The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of blue being correct and the probability of the social information indicating the correct answer. Probability schedules were counterbalanced between participants. Solid blue lines show the probability of blue being the correct choice, dashed red lines show the probability of the social information being correct. Schedules 1-4 are displayed here.

2.1.2 Pre-task quiz

(Correct answers are indicated with a star)

Q1) What does it mean if the blue square is on the left side of the screen?

A) The blue box is the correct answer

B) The blue box is more likely to be correct

C) Nothing, the blue and green boxes change at random *

D) The game is entering a new phase

Q2) What does the red frame signify?

A) The most likely answer

B) The output from a (possibly biased) roulette wheel

C) Nothing, the information is deliberately misleading

D) What previous players believed the correct answer was at some point in the game *

Q3) A small square appears between the two larger squares when an answer is given. What does it represent?

A) The type of phase the game is in

B) The correct answer *

C) The answer you gave

D) Your previous response that you gave

2.2 Data analysis

All statistical analyses were conducted using MATLAB R2017b (The MathWorks, Natick, MA) and JASP (JASP Team (2020). JASP (Version 0.14) [Computer software]). We used the standard $p < .05$ criteria for determining if significant effects were observed, with a Holm correction [429] applied for multiple comparisons, to control for type I family-wise errors. Holm correction was used as it progressively adapts the threshold values for significance, resulting in an increase in power, relative to other correction methods [430]. Where appropriate, data were transformed to meet assumptions of normality for parametric testing.

2.2.1 Bayesian statistical testing

For all analyses, Bayesian statistical testing was implemented as a supplement to null hypothesis significance testing, allowing for comparison of the likelihood of observed data under the null and alternative hypothesis. All Bayesian tests were carried out using JASP software (JASP Team (2020)).

Bayesian t-tests [426] were used to determine if null results represented a true lack of a difference between genotype groups. A Bayes factor, comparing the fit of data under the null and the alternative hypothesis was estimated for each comparison [427],[428], whereby the null hypothesis (H_0) postulates that there are no differences between groups. A two-sided alternative hypothesis was used, allowing the effect size (δ) to take both positive and negative values. We used a default Cauchy prior distribution for a two-sample t-test, specifically, a zero-centred Cauchy distribution with a scale of 0.707 [427]. In addition, Bayes inclusion factors (BF_{incl}) were included when conducting repeated measures ANOVAs (RM-ANOVA). The JASP framework was used [402], whereby candidate models M and their related condition-effect parameters β were compared, resulting in Bayes factors for each candidate model, quantifying the relative predictive performance of different models, as well as inclusion Bayes factors for predictors of interest.

The inclusion Bayes factor (BF_{incl}) for a given predictor represents the evidence, given the observed data, for including a certain predictor in the model. BF_{incl} quantifies the change in odds from the prior probability that the predictor is included in the model to the posterior probability of inclusion after seeing the data and is computed by comparing all models with a

given predictor against all models without that predictor, i.e., comparing models that contain the effect of interest to equivalent models stripped of the effect. For example, an inclusion Bayes factor for an effect of 3 for a given predictor *i* or interaction can be interpreted as stating that models which include the predictor *i* are 3 times more likely to describe the observed data than models without the predictor or interaction. Similarly, the inverse, a Bayes exclusion factor (BF_{excl}) for a given predictor or interaction, represents the relative likelihood that a model *without* the predictor *i* can best explain the observed data [384]. For all Bayesian analyses, we followed a widely used classification scheme [385], in which BF_{10}/BF_{incl} values between one and three are considered weak evidence, between three and ten as moderate evidence and greater than ten as strong evidence for the alternative hypothesis H1. In addition, BF_{01} values between 1 and 0.33 are considered weak evidence, between 0.33 and 0.1 as moderate evidence, and smaller than 0.1 as strong evidence for the null hypothesis respectively.

2.2.2 Optimal learner model

The influence of each information source (primary and secondary) on choices was quantified by regressing two "optimal learners" against subjects' choices. The first comprised an optimal "individual learner model", which was generated by using a Bayesian learner algorithm [88] to simulate an optimal learner who learns solely from individual information (the blue and green stimuli). The second comprised a "social learner model" which simulated an optimal learner who learns solely from the social information (red stimuli). The Bayesian learner algorithm [88] describes an optimal approach to tracking reward probabilities in a changing environment. It assumes an underlying probability of an outcome being correct and tracks this probability across time, as well as maintaining an estimate of the rate of change of probabilities, i.e., volatility. All probabilities are updated in a Markovian fashion, meaning there is no requirement to store the full history of decision outcomes or statistics of the environment [88]. Thus, on each trial, the individual learner model represented the reward probability associated with a blue choice, derived through learning, in an optimal fashion, exclusively from information about reward outcomes and ignoring the social information. The social learner model represented the probability, based on the (reward-weighted) social information, that the social information was correct. From the social learner model, on each trial, the reward probability of a blue choice was calculated, that would have been derived if a

participant had been learning optimally, exclusively from the social information (i.e., ignoring individual reward outcomes). Subsequently both models were regressed separately against each individual participant's choice data using binomial logistic regression, with model predictions from the primary and secondary models as continuous predictor variables and participant response as the dependent variable (0/1). For each participant, this produced two parameter estimates, or standardised beta weights, each representing the degree to which individual experience and social information explained choices. For example, a participant whose choices were more strongly influenced by the social information than the individual information, would have a high social $\beta_{optimal}$ value, and a low individual $\beta_{optimal}$ value.

2.3 Computational modelling framework

Participant response was modelled using a Rescorla-Wagner (RW) learning model [24], consisting of a perceptual model and an action selector. The model relies on the assumption that updates to choice behaviour are based on prediction errors, i.e., the difference between an expected and the actual outcome. Participants were assumed to update their beliefs about reward outcomes, or the state of the environment based on sensory feedback in the form of reward, and to use this feedback to make decisions about the next action (response model).

2.3.1 Perceptual model

The RW predictors used in our learning models consisted of a modified version of a simple learning model, with a free parameter, learning rate (α), which varied between 0 and 1.

$$V_{(i+1)} = V_i + \alpha(r_i - V_i)$$

The predicted value on each trial (V_i) is updated on each trial based on the prediction error (PE), or the difference between the expected and the actual reward ($r_i - V_i$), weighted by the learning rate α. α thus captures the extent to which the PE updates the estimated value on the next trial. In line with previous work [93],[309], we used an extended version of this learning model, with separate α values for volatile and stable environmental phases. In a stable environment, learning rate will optimally be low, and reward outcomes over many trials will be taken into account. In a volatile environment, however, an increased learning rate is

optimal, as more recent trials are used to update choice behaviour [88]. Rescorla-Wagner predictors were simultaneously estimated for parameters relating to learning from both individual and social reward information. Consequently, the model generated the predicted value of going with the individual source (going with the blue frame; $V\_individual(i+1)$) and the predicted value of the social information (going with the group recommendation; $V\_social(i+1)$) and provided four α estimates: $\alpha_{individual\_stable}$, $\alpha_{individual\_volatile}$, α_{social_stable}, $\alpha_{social\_volatile}$.

2.3.2 Response model

An action selector was utilised whereby information from learning from both information sources (individual and social reward) was integrated. The action selector predicts the probability that the blue choice will be rewarded on a given trial, and was based on the softmax function (TAPAS toolbox - available at http://www.translationalneuromodeling.org/tapas) and adapted by Diaconescu and colleagues [222],[230]. This response model is identical to that used by Cook and colleagues [93] and reproduced here with permission. The value of individual and social information was combined using the following:

$$vB_{(i+1)} = \zeta\left(V_{social_{advice_{weighted(i+1)}}}\right) + (1-\zeta)\left(V_{individual(i+1)}\right)$$

whereby the belief that blue would be rewarded (vB) comprises the integrated value of individual information ($V_{individual}$) and social information. $V_{social_{advice_{weighted}}}$ comprises the "social" advice weighted by the probability of social advice accuracy. ζ is a free parameter which describes the weighting of social relative to individual information in belief integration.

The probability that this belief would determine participant choice was described by a unit square sigmoid function, describing how learned belief values are translated into choices.

$$P(y_{(i_1)} = 1 \,||vB_{(i+1)}) = \frac{vB_{(i+1)}{}^{\beta}}{vB_{(i+1)}{}^{\beta} + (1-vB_{(i+1)})^{\beta}}$$

The participant-specific free parameter β, the inverse of the decision temperature, describes the extent to which estimated value of choices determines actual participant choice: as β decreases, decision noise increases and decisions become more stochastic; as β increases, decisions become more deterministic towards the higher value option.

2.3.3 Model fitting

Optimisation of free parameter values was performed as per Cook and colleagues [93], using a quasi-Newton optimisation algorithm specified in the TAPAS toolbox (quasinewton_optim_config.m). The function maximised the log-joint posterior density over all parameters given the data and the generative model. α values were estimated in logit space (see tapas_logit.m), i.e., a logistic sigmoid transformation of native space (tapas_logit(x) = ln(x/(1-x)); x = 1/(1+exp(-tapas_logit(x)))). An uninformative prior, allowing for individual differences in learning rate was used for α: tapas_logit (0.2, 1), with a variance of 1. Initial values were fixed at logit (0.5, 1). The prior for β was set to log (48), with a variance of 1, and the prior for ζ was set at 0 with a variance of 10^2 (logit space), i.e., an equal weighting for information derived from individual-value and social-value learning (0.5). All prior choices were based on previous work. Maximum-a-posteriori (MAP) estimates for all model parameters were calculated using the HGF toolbox version 3 (OSF link). All code used is adapted from the open-source software package TAPAS (available at http://www.translationalneuromodeling.org/tapas).

2.3.4 Optimal model parameters

For all analyses, estimated α values were compared to optimal α estimates. An optimal learner model, with the same architecture and priors as the model employed in the current task, was fit to 100 synthetic datasets, resulting in average optimal learning rates: $\alpha_{\text{optimal\_primary\_stable}} = 0.16$, $\alpha_{\text{optimal\_primary\_volatile}} = 0.21$, $\alpha_{\text{optimal\_secondary\_stable}} = 0.17$, $\alpha_{\text{optimal\_secondary\_volatile}} = 0.19$. Scores representing the difference between (untransformed) α estimates and optimal α scores were calculated ($\alpha_{diff} = \alpha - \alpha_{\text{optimal}}$).

2.4 Computational modelling validation

2.4.1 Model comparison

Although there was *a priori* evidence to model choice behaviour with the chosen model, which featured four α estimates: $\alpha_{\text{individual\_volatile}}$, $\alpha_{\text{individual\_stable}}$, $\alpha_{\text{social\_volatile}}$, $\alpha_{\text{social\_stable}}$, we explored whether participants in ASD and CTRL groups might be solving the task in structurally different ways. To this end, we adapted our learning model, resulting in eight possible computational models representing different strategies that participants might use to solve the task (note that we do not imply that participants are explicitly aware of their strategy). All models were variations of the classic Rescorla-Wagner model. Based on WSLS results, we included a family of models which included separate learning rates from learning from wins and losses (Models 5-8).

A formal model comparison was carried out using Group level Bayesian model selection (BMS), to evaluate which model provided the (relative) best fit to the observed data. The VBA toolbox , specifically random-effects BMS (using the VBA_groupBMC_btwConds.m function), was utilised [431]. Random effects group BMS computes an approximation of the model evidence relative to the other models, i.e., the probability of the data y given a model m, $p(y|m)$, with log model evidence here approximated with Akaike Information Criterion (AIC) values. The posterior probability that a model has generated the observed data, relative to other models is estimated, as well as the exceedance probability, representing the likelihood that a given model is more likely than other included models in the set. Analysis across both groups allows us to test the hypothesis that the same model produced observed data for the ASD and CTRL groups. Specifically, we investigated whether the same model could explain data from both groups, or whether there was sufficient evidence to conclude that the data from the ASD and CTRL groups are best fit by differing models. For example, it could be that a model that includes the factor volatility provides the best fit to the data produced by the CTRL group, whereas for the ASD group the best fit is provided by a model that does not include volatility as a factor.

Model 1 was a classic Rescorla-Wagner model:

$$V_{(i+1)} = V_i + \alpha\varepsilon_i$$

with $\varepsilon_i = r_i - V_i$, the difference between the actual and the expected reward or prediction error (PE).

Model 2 was an extension of Model 1, with separate learning rates (α) for learning from primary value and secondary value learning sources:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary}\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)} + \alpha_{secondary}\varepsilon_i$$

Model 3 had a single learning rate α for primary/secondary learning, but separate learning rates for volatile and stable blocks:

$$V_{(i+1)} = V_i + \alpha_{volatile}\varepsilon_i + \alpha_{stable}\varepsilon_i$$

Model 4 had four separate learning rates α for volatile and stable and primary and secondary learning:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary_{volatile}}\varepsilon_i + \alpha_{primary_{stable}}\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)} + \alpha_{secondary_{volatile}}\varepsilon_i + \alpha_{secondary_{stable}}\varepsilon_i$$

As an exploratory measure, we further extended Models 1-4 to include separate learning rates corresponding to learning from rewarded trials and unrewarded trials separately, i.e., learning from wins and losses.

Model 5:

$$V_{(i+1)} = V_i + \alpha_{reward}\,\varepsilon_i + \alpha_{unreward}\,\varepsilon_i$$

Model 6:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary_{reward}}\,\varepsilon_i + \alpha_{primary_{unreward}}\,\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)} + \alpha_{secondary_{reward}}\,\varepsilon_i + \alpha_{secondary_{unreward}}\,\varepsilon_i$$

Model 7:

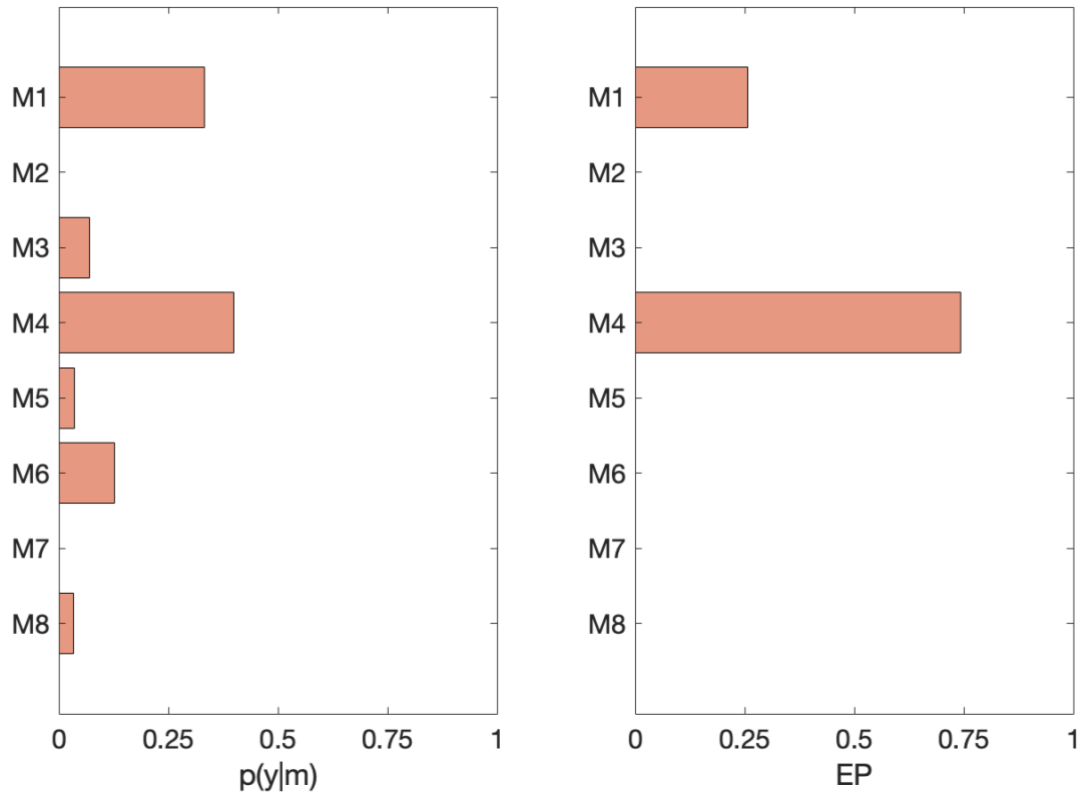
$$V_{(i+1)} = V_i + \alpha_{volatile_{reward}}\varepsilon_i + \alpha_{stable_{reward}}\varepsilon_i + + \alpha_{volatile_{unreward}}\varepsilon_i + \alpha_{stable_{unreward}}\varepsilon_i$$

Model 8:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary_{volatile_{reward}}}\varepsilon_i + \alpha_{primary_{stable_{reward}}}\varepsilon_i + + \alpha_{primary_{volatile_{unreward}}}\varepsilon_i + \alpha_{primary_{stable_{unreward}}}\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)}V_{(i+1)} + \alpha_{secondary_{volatile_{reward}}}\varepsilon_i + \alpha_{secondary_{stable_{reward}}}\varepsilon_i + \alpha_{secondary_{volatile_{unreward}}}\varepsilon_i + \alpha_{secondary_{stable_{unreward}}}\varepsilon_i$$

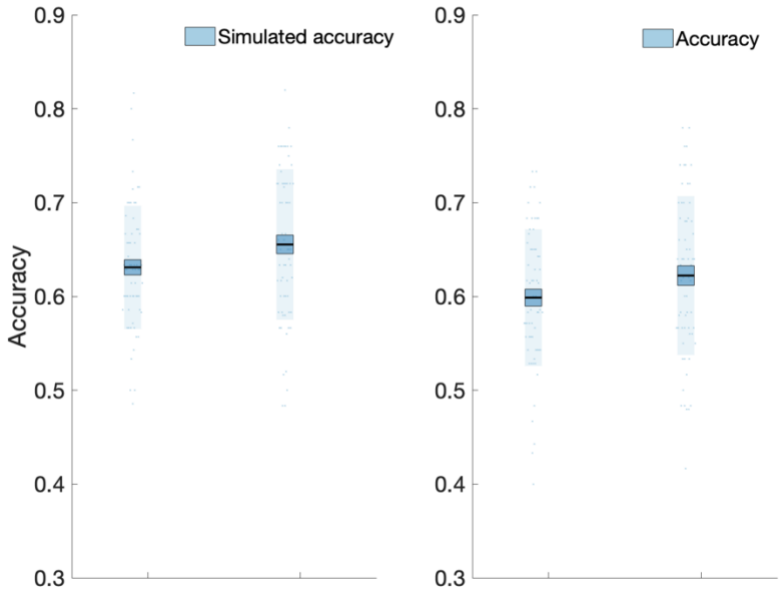
BMS, using AIC scores, revealed a high probability that the data from the ASD and CTRL groups were fit by the same model with the posterior probability that the two groups had the same model frequencies = 0.854 (Suppl. Fig. 2.3). Thus, we failed to reject the null hypothesis of no difference between groups. The previously chosen model (Model 4) was the winning model; an adapted Rescorla-Wagner model with separate α values for stable and volatile environments and separate α values for individual and social learning.



Supplementary Figure 2.3. Between-groups model selection (BMS) A. Estimated posterior model frequencies (p($y|m$)). B. Exceedance probabilities for each model (EP). Models (1-8) were compared based on AIC scores for each participant (n = 64) across both ASD and CTRL groups.

2.4.2 Model Validation

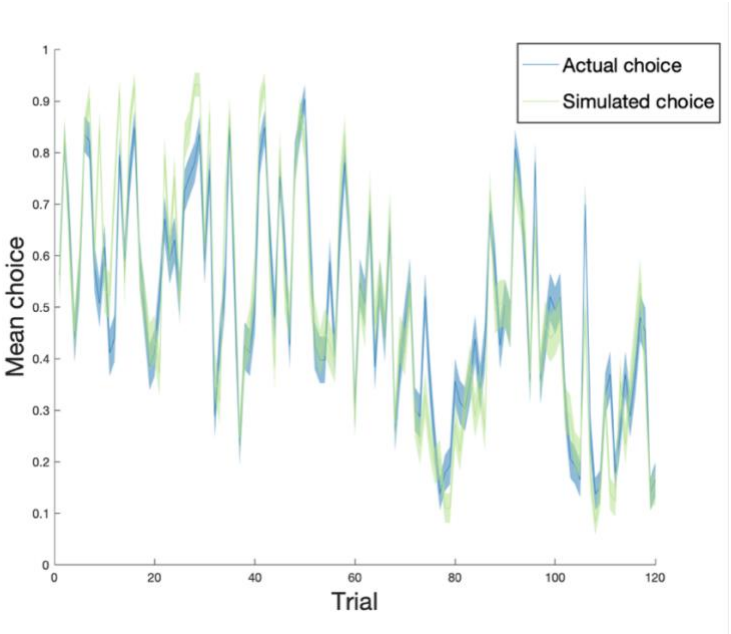
To demonstrate that the chosen model (model 4) accurately described participant behaviour, we subsequently simulated response data for each participant, using estimated model parameter values (tapas_simModel.m). Simulated accuracy and calculated accuracy were significantly correlated for each participant (r = 0.623, p <.001) (Suppl. Fig. 2.4).



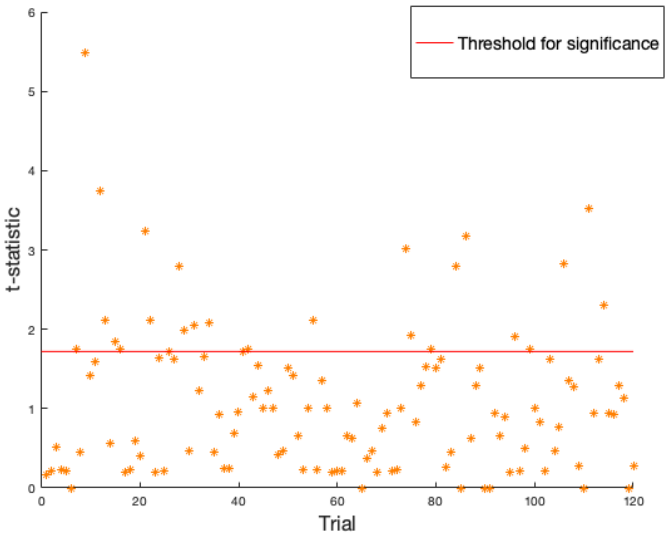
Supplementary Figure 2.4. Model simulations (left) and participant response data (right). Mean accuracy is displayed separately for volatile and stable environmental phases. Boxes = standard error of the mean, shaded region = standard deviation, individual datapoints (n = 65) are displayed.

In additon, actual and simulated choice were compared, to ensure that the chosen model was indeed capturing participants' choice behaviour. On each trial (1-120), we compared mean and simulated choice (averaged across subjects), using a paired t-test (Suppl. Fig. 2.5A). We then employed bootstrapped paired t-tests (n = 100) with resampled distributions, to isolate significant differences. We defined significant differences as those occurred in less than 5% of t-tests (Suppl. Fig. 2.5B).

A



B



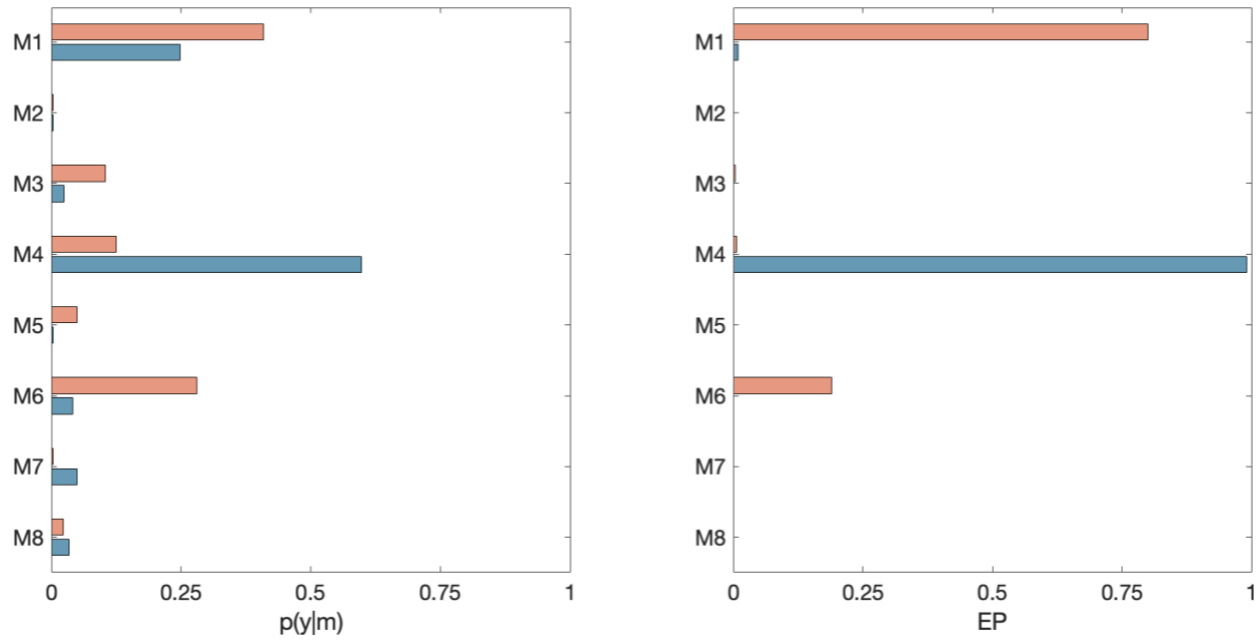
Supplementary Figure 2.5. A. Participant choice data and model simulations. Line indicates mean choice across all participants. Shaded region = standard error of the mean. **B.** T-statistics for comparison between actual and simulated choice. Choices significantly differed on 26 trials.

Parameter recovery

To ensure that parameter estimates could be recovered, model parameters were estimated from simulated data for each participant. All recovered parameters correlated significantly with estimated parameters (all $p < 0.001$).

2.4.3 Estimated parameters from other models

Interestingly, although BMS provided evidence for model 4 overall, and there was *a priori* evidence to model choice behaviour with this model, which featured four α estimates: $\alpha_{\text{individual\_volatile}}$, $\alpha_{\text{individual\_stable}}$, $\alpha_{\text{social\_volatile}}$, $\alpha_{\text{social\_stable}}$, more heterogeneity was observed within the ASD group. Within this group, the posterior probability of the same model explaining the data was equal to 0.607, compared to 1.000 for the CTRL group (Suppl. Fig. 2.6). Within the ASD group, model 1, comprising of only 2 free parameters, α and β, was the winning model, followed by model 6, with separate α values for individual/social and separate α values for wins/losses, as well as β and ζ. Thus, suggesting that the previously described model (Model 4) does not explain all observed variance within this group and that participants in the ASD and CTRL groups might be solving the task in structurally different ways. We further explored our data by considering parameter estimates from alternative computational models. Specifically, we investigated whether estimated parameters from models 1 and 6 differed between the ASD and CTRL groups.



Supplementary Figure 2.6. A. Estimated posterior model frequencies (p(y|m)). **B.** Exceedance probabilities for each model (EP). Models (1-8) were compared based on AIC scores for each participant separately for ASD (n = 29) and CTRL groups (n = 35). Red bars indicate ASD, and blue bars indicate CTRL groups respectively.

Model 1

Independent t-tests showed that neither α (t(63) = 1.125, p = 0.265, d = 0.281) nor β values (t(63) = 0.659, p = 0.512, d = 0.164) differed between ASD and CTRL groups.

Model 6

A RM ANOVA was carried out on α values, with information source (social, individual) and reward (wins, losses) as WS factors and group (ASD, CTRL) as BS factor. A main effect of information and an information by reward interaction effect were observed. However, no main/interaction effect(s) involving group were observed. Independent t-tests between groups showed that neither ζ (t(63) = 0.874, p = 0.385, d = 0.218) nor β values (t(63) = 0.603, p = 0.549, d = 0.150) differed between ASD and CTRL groups.

2.5 Extended statistical analysis

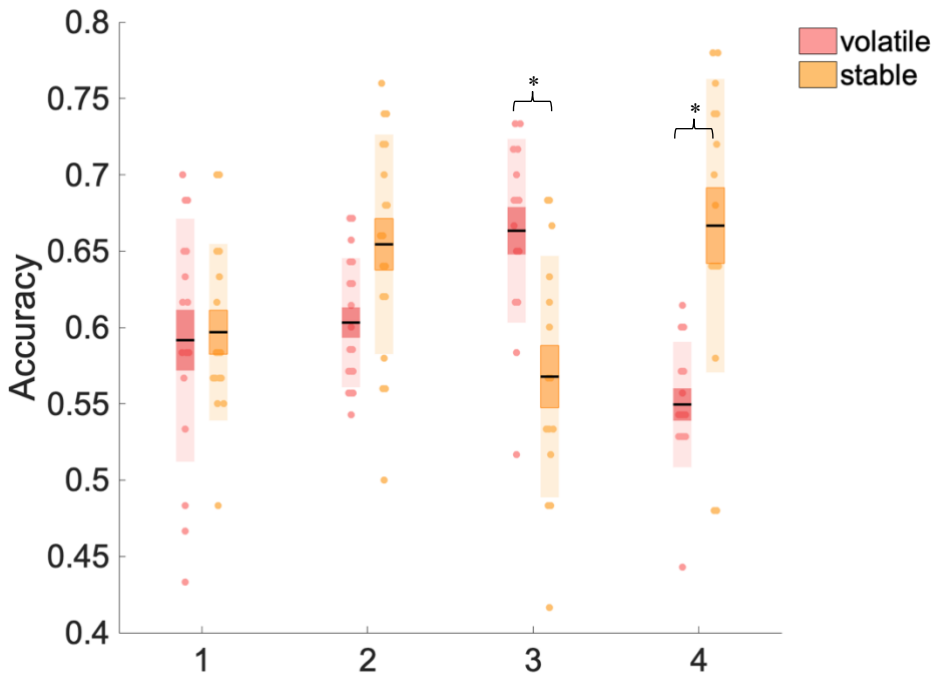
Supplementary Table S2.1. Untransformed learning rates

	$\alpha_{individual\_volatile}$		$\alpha_{individual\_stable}$		$\alpha_{social\_volatile}$		α_{social_stable}	
	Ctrl	**ASD**	**Ctrl**	**ASD**	**Ctrl**	**ASD**	**Ctrl**	**ASD**
n	35	29	35	29	35	29	35	29
Mean	0.334	0.387	0.321	0.411	0.147	0.134	0.139	0.161
Std. Error of Mean	0.033	0.038	0.036	0.039	0.025	0.018	0.018	0.021
Minimum	0.075	0.035	0.029	0.067	0.009	0.020	0.010	0.004
Maximum	0.728	0.727	0.739	0.724	0.528	0.409	0.588	0.444

Note: Ctrl refers to non-autistic group, ASD refers to autistic group.

2.5.1. Accuracy and reaction time

Accuracy was compared between groups, randomisation schedules, and volatile and stable phases. A repeated measures analysis of variance (RM-ANOVA), with within-subjects factor volatility (stable, volatile), and between-subjects factors group (ASD, CTRL) and randomisation schedule (1-4) demonstrated no difference in accuracy between ASD (mean (standard error) $\bar{x}$ ($\sigma_{\bar{x}}$) accuracy = 0.607 (0.009)), and CTRL groups ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.617 (0.009); F (1,56) = 0.676 p = 0.414, η_p^2 = 0.012). Furthermore, no difference in accuracy was observed between volatile ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.602 (0.009)) and stable phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.623 (0.011); F (1,56) = 1.835, p = 0.181, η_p^2 = 0.032). However, a schedule by volatility interaction was observed (F (3,56) = 11.297, p < 0.001, η_p^2 = 0.377) (Suppl. Fig. 2.7). For participants in schedules 1 and 2, no difference in accuracy was observed between stable and volatile phases (p_{holm} > 0.05). However, for schedules 3 and 4, accuracy significantly differed as a function of volatility; for schedule 3, higher accuracy was observed in volatile ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.663 (0.016) compared to stable phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.568 (0.020), t(64) = 3.455, p_{holm} = 0.024). For schedule 4, however, higher accuracy was observed in stable ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.667 (0.025) compared to volatile phases ($\bar{x}$ ($\sigma_{\bar{x}}$) = 0.550 (0.011), t(64) = 4.553 , p_{holm} = 0.001). However, as participants were counterbalanced to different randomisation schedules, with the proportion of participants assigned to each schedule not differing between groups (X^2 (1, N = 64) = 1.658, p = 0.646), this pattern of results does not affect between-group comparisons. Finally, mean reaction time (RT) did not significantly vary between ASD ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.455 (0.130) and CTRL groups ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.240 (0.130); F (1,62) = 1.605, p = 0.210, η_p^2 = 0.025).

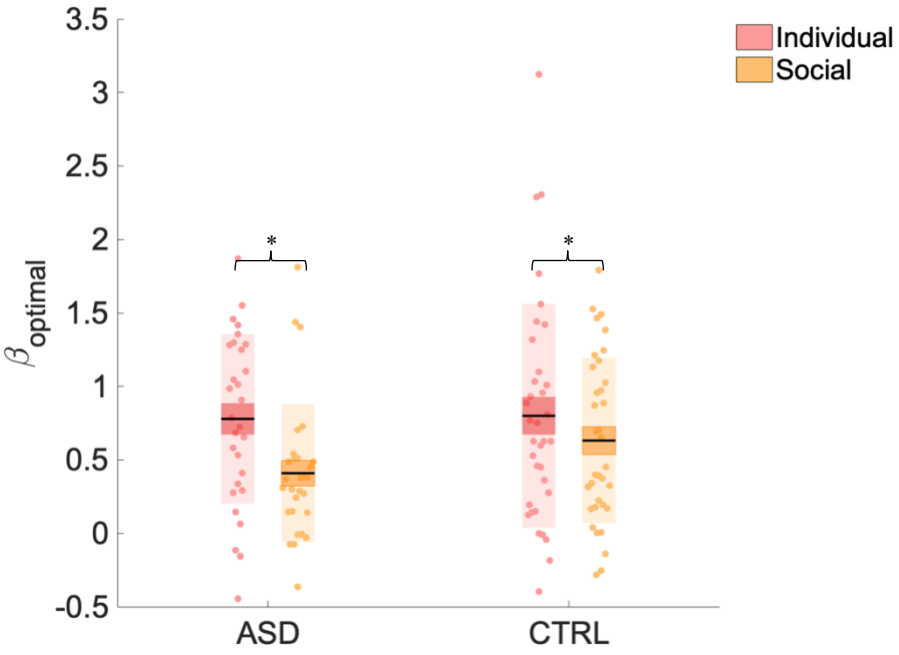


Supplementary Figure 2.7. Mean accuracy across randomisation groups. A significant randomisation by schedule interaction was observed. Data points indicate mean accuracy for individual participants (n = 64), bold point indicates the mean, bold line indicates standard error of the mean (1 SEM), * indicates statistical significance ($p_{holm} < 0.05$).

2.5.2 Optimal learner model

We tested whether participants in both ASD and CTRL groups learned in a more optimal fashion from the individual versus social source of information. A Bayesian learner model was used to create two optimal models (1) an optimal individual learner, and (2) an optimal social learner (Appendix 2.2). Subsequently we regressed both models against participants' choice data, resulting in two $\beta_{optimal}$ values capturing the extent to which a participant made choices according to the optimal individual, and optimal social learner models respectively. $\beta_{optimal}$ values were submitted to a RM-ANOVA with factors information source (individual, social) and group (ASD, CTRL). A main effect of information source was observed ($F(1,62) = 4.429$, $p = 0.039$, $\eta_p^2 = 0.067$), with $\beta_{optimal}$ values significantly higher for individual ($\bar{x}(\sigma_{\bar{x}}) = 0.795\ (0.076)$), compared with social information ($\bar{x}(\sigma_{\bar{x}}) = 0.526\ (0.076)$) (Suppl. Fig. 2.8). However, no main effect of group ($F(1,62) = 2.211$, $p = 0.151$, $\eta_p^2 = 0.033$, $BF_{excl} = 3.053$) or, crucially, no information by group interaction effect ($F(1,62) = 0.615$, $p = 0.436$,

$\eta_p^2 = 0.010$, $BF_{excl} = 2.563$) was observed. Results show that the extent to which a participant made choices based on individual and social information did not differ between the groups.



Supplementary Figure 2.8. Beta weights ($\beta_{optimal}$) for individual and social information across ASD and CTRL groups. $\beta_{optimal}$ values were significantly higher for individual, versus social information. No differences were observed between ASD and CTRL groups. Data points indicate $\beta_{optimal}$ for individual participants (n = 64), bold point indicates the mean, bold line indicates standard error of the mean (1 SEM), * indicates statistical significance ($p < 0.05$). $\beta_{optimal}$ for individual information was significantly lower than $\beta_{optimal}$ for social information. $\beta_{optimal}$ did not differ between groups.

Appendix 3

Supplementary material for Chapter 4

3.1. Genotyping Analysis

Genetic analyses were carried out in at the department of Human Genetics of the Radboud University Nijmegen Medical Centre. Participants were asked to donate saliva samples. High molecular weight DNA was isolated from saliva using Oragene kits according to the manufacturer protocol (DNA Genotek Inc., Kanata, Ontario, Canada).

In total, 803 participants were genotyped for variants in the dopamine transporter gene (SLC6A3/ DAT1), the serotonin transporter gene (SLC6A4/SERT/5HTT) and the Catechol-o-methyl-transferase gene (COMT). 5% duplicate and blank samples were included as quality controls to control for random genotyping error. The genotyping assay was validated before use.

Genotyping of DAT1/ SLC6A3

Genotyping of the 40 base pair variable number of tandem repeats (VNTR) polymorphism in the 3' untranslated region (UTR) of the SLC6A3 encoding the DAT1 transporter was carried out below as previously described by den Ouden et al [236]. After DNA extraction, amplification of genomic DNA was carried out with 0.33 µM of a forward primer (NED-5'-TGTGGTGTAGGGACGGCCTGAGAG-3') and reverse primer (5'-CTTCCTGGAGGTCACGGCTCAAGG-3') in 1x AmpliTaq Gold® 360 Mastermix (Applied Biosytems, Nieuwerkerk a/d Ijssel, The Netherlands). The cycling conditions for the polymerase chain reaction (PCR) started with 10 min at 92°C, followed by 35 cycles of 30 sec at 95°C, 30 sec at the optimized annealing temperature of 64°C, and 1 min 72°C, then followed by an extra 7 min 72°C. The amplifications were performed in a PTC200 Multicycler (MJ- Research via Biozym, Landgraaf, The Netherlands). Subsequent to this, determination of allele length was performed by direct analysis on an automated capillary sequencer (ABI3730, Applied Biosystems) using standard conditions.

The 10R and 9R alleles were analysed, with 330 subjects homozygous for the 10R allele, 32 homozygous for the 9R allele, and 208 subjects heterozygous (9/10). Participants with different allele variants (such as 10/11, 9/11) were excluded. Testing for Hardy–Weinberg equilibrium did not show deviations from the expected distribution ($\chi^2(3) = 5.284$, $p = 0.152$).

Genotyping of SERT/5-HTT/rs25531

Genotyping of the 5HTTLPR polymorphism was performed by simple sequence length analysis. PCR was performed on 50 ng of genomic DNA using 0.5 µM fluorescently labelled forward primer (FAM-5'-GGCGTTGCCGCTCTGAATGC-3') and reverse primer (5'-GAGGGACTGAGCTGGACAACCAC-3'), 0.25 mM dNTPs, 1x PCR optimization buffer A (30 mM Tris-HCl pH 8.5, 7.5mM (NH4)2SO4, 0.75 mM MgCl2), 10% DMSO and 0.4 U AmpliTaq Gold® DNA Polymerase (Applied Biosystems). The cycling conditions for the PCR started with 12 min at 95°C, followed by 35 cycles of 1 min at 94°C, 1 min at the optimized annealing temperature (57.5°C), and 2 min. 72°C, then followed by an extra 10 min 72°C. Subsequent determination of the length of the alleles was performed by direct analysis on an automated capillary sequencer (ABI3730, Applied Biosystems) using standard conditions. The SNP present in the 5HTTLPR (rs25531) was genotyped using Taqman analysis (assay ID: Taqman assay: C_25746809_50; Applied Biosystems). Genotyping was carried out in a volume of 10 µl containing 10 ng of genomic DNA, 5 µl of ABgene Mastermix (2x; ABgene Ltd., Hamburg, Germany), 0.125 µl of the Taqman assay and 3.875 µl of H2O. Amplification was performed on a 7500 Fast Real-Time PCR System starting with 15 minutes at 95ºC, followed by 50 cycles of 15 seconds at 95ºC, 1 minute at 60ºC. Genotypes were scored using the algorithm and software supplied by the manufacturer (Applied Biosystems). The assay has been validated by digesting the 5HTT PCR product with MspI (New England Biolabs, Ipswich, USA) and separating the restriction fragments on a 2% agarose gel. This resulted in restriction fragments of 340 bp,130 bp and 60 bp for the LA allele, fragments of 175 bp, 165 bp, 130 bp and 60 bp for the LG allele, and fragments of 300 bp, 130 bp and 60 bp for the S allele. As LG is comparable to the S-allele with regard to gene expression, we grouped the S and LG alleles together for the behavioural analysis, resulting in L'L', S'L', and S'S' genotypes [432],[433]. 185 subjects were homozygous for the s allele, 123 homozygous for the l allele, and 262 subjects were heterozygous (S/L). These

genotyping results did not show deviations from Hardy–Weinberg equilibrium ($\chi^2(3) = 0.185$, $p = 0.667$).

Genotyping of COMT

Genotyping of the COMT Val158Met polymorphism was carried out using Taqman analysis. Amplification was carried out using PCR. 166 subjects were homozygous for the val allele, 134 homozygous for the met allele, and 270 subjects were heterozygous (val/met). Testing for Hardy–Weinberg equilibrium did not show deviations from the expected distribution ($\chi^2(3) = 0.215$, $p = 0.643$).

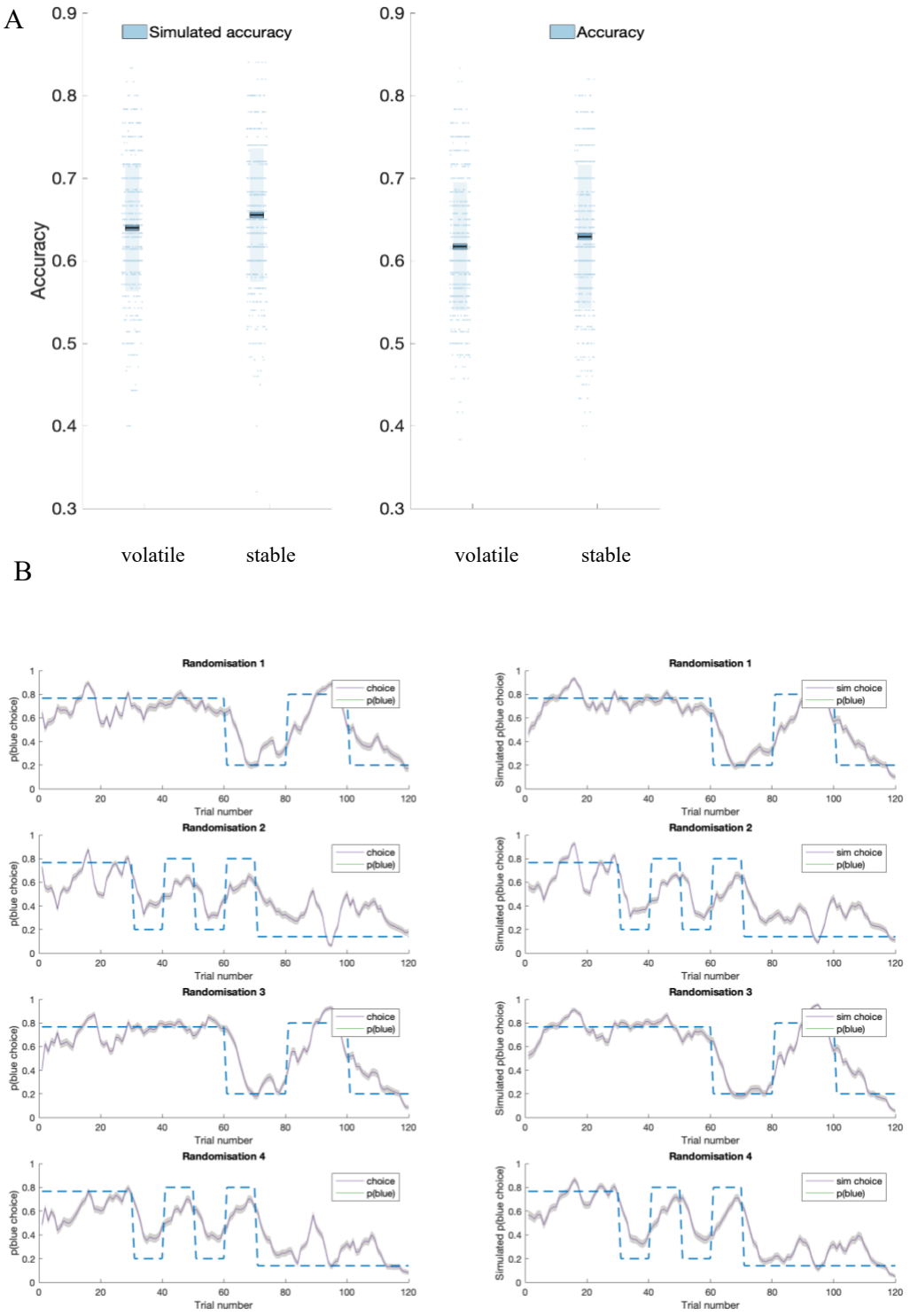
3.2 Computational modelling

3.2.1 Model comparison

RFX-BMS model comparison, with eight models was conducted. The winning model was a model with a single α value, based on F values (LME), with model 4 the next best performing model. Based on AIC values, however, model 4 outperformed model 1, with both an exceedance probability and posterior model probability of 1. As we had a strong a priori hypothesis concerning different learning rates for social and individual information, model 4 was chosen as the winning model.

3.2.2 Model Simulation

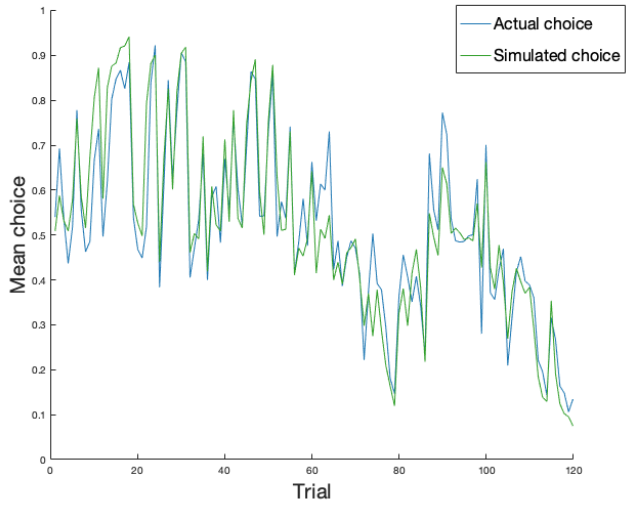
We simulated response data for each participant, using estimated model parameter values (tapas_simModel.m). Simulated and calculated accuracy were significantly correlated for stable ($r = 0.665$, $p < 0.001$) and volatile phases ($r = 0.627$, $p < 0.001$) (Suppl. Fig. 3.1).



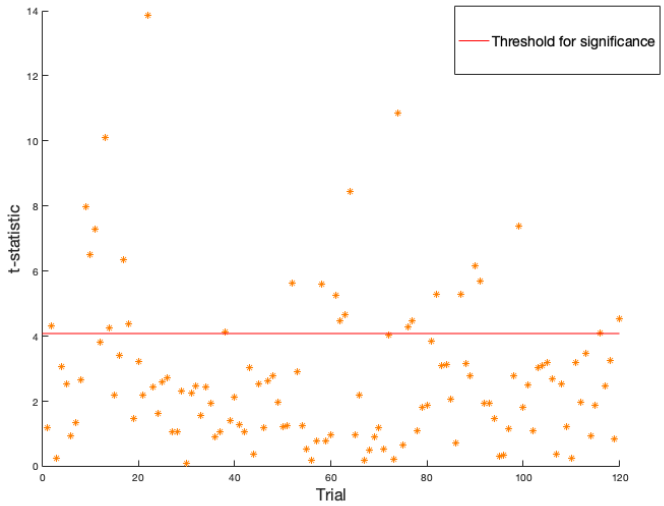
Supplementary Figure S3.1. A. Model simulations (left) and participant response data (right). Mean accuracy is displayed separately for volatile and stable environmental phases. Boxes = standard error of the mean, shaded region = standard deviation, individual datapoints are displayed. **B.** Participant data (left) juxtaposed against model simulations (right) Running average, across 5 trials of blue choices for probabilistic randomisation schedules 1 to 4. Shaded region = standard error of the mean.

We next compared actual and simulated choice, to ensure that the computational model was indeed capturing participants' choice behaviour. On each trial (1-120), we compared mean and simulated choice (averaged across subjects), using a paired t-test (Suppl. Fig. 3.2A). We then employed bootstrapped paired t-tests (n = 100) with resampled distributions, to isolate significant differences. Analysis revealed 26 trials where significant differences were found. We defined significant differences as those occurred in less than 5% of t-tests (Suppl. Fig. 3.2B).

A



B



Supplementary Figure S3.2.A. Participant choice data and model simulations. Line indicates mean choice across all participants. Shaded region = standard error of the mean. B. T-statistics for comparison between actual and simulated choice.

Parameter recovery

Finally, and to ensure that parameter estimates could be recovered, model parameters were estimated from simulated data for each participant, using the same model. All recovered parameters correlated significantly with estimated parameters (all $p < 0.001$).

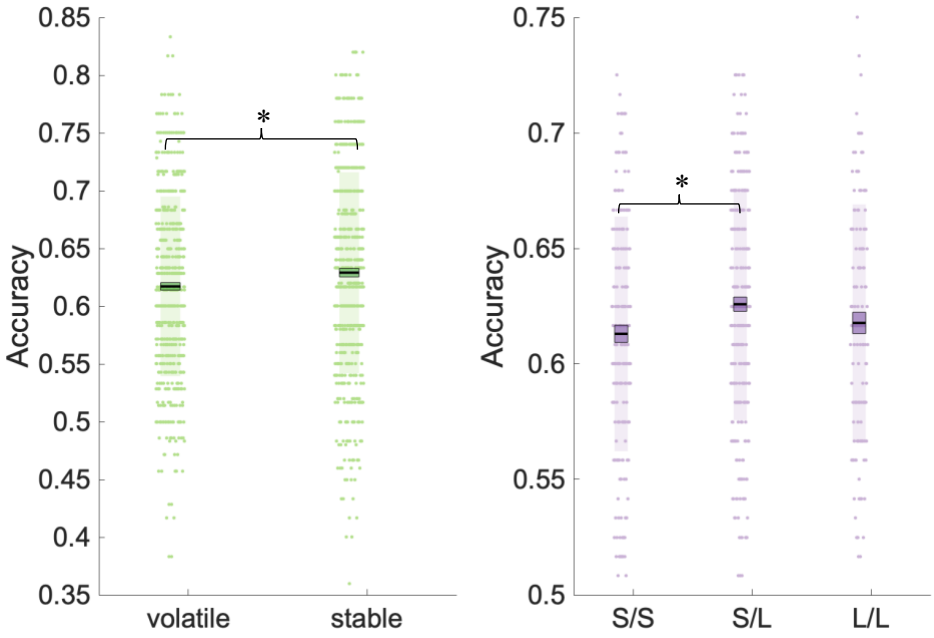
3.3 Extended statistical analysis

3.3.1. Accuracy

Mean accuracy scores were submitted to a RM-ANOVA, with volatility and schedule as predictor variables. A significant main effect of volatility was observed ($F(1,566) = 8.063$, $p = 0.005$, $\eta^2p = 0.006$), with higher accuracy in stable (mean (standard error) $\bar{x}(\sigma_{\bar{x}}) = 0.629$ (0.003)) as compared to volatile phases ($\bar{x}(\sigma_{\bar{x}}) = 0.617$ (0.003) (Suppl. Fig. 3.3A). A main effect of schedule was observed ($F(3,566) = 3.024$, $p = 0.029$, $\eta^2p = 0.016$) with mean accuracy for schedule 2 ($\bar{x}(\sigma_{\bar{x}}) = 0.632$ (0.004)) significantly higher than schedule 1 ($\bar{x}(\sigma_{\bar{x}}) = 0.614$ (0.004); $t(566) = 2.975$, $p_{holm} = 0.018$), and non-significantly higher than for schedules 3 ($\bar{x}(\sigma_{\bar{x}}) = 0.622$ (0.004); $t(566) = 1.722$, $p_{holm} = 0.428$) and 4 ($\bar{x}(\sigma_{\bar{x}}) = 0.625$ (0.004); $t(566) = 1.242$, $p_{holm} = 0.588$. Finally, a significant schedule by volatility interaction was observed ($F(3,566) = 100.323$, $p < 0.001$, $\eta^2p = 0.347$). Accuracy did not vary significantly across schedule in volatile phases ($p_{holm} > 0.05$). However, in stable phases, accuracy differed a function of schedule. Mean accuracy for schedule 1 ($\bar{x}(\sigma_{\bar{x}}) = 0.575$ (0.004)) was numerically lower than for schedule 2 ($\bar{x}(\sigma_{\bar{x}}) = 0.622$ (0.004); $t(566) = -2.826$, $p_{holm} = 0.090$ and significantly lower than for schedule 4 ($\bar{x}(\sigma_{\bar{x}}) = 0.662$ (0.004); $t(566) = -4.218$, $p_{holm} < 0.001$). Similarly, mean accuracy for schedule 3 ($\bar{x}(\sigma_{\bar{x}}) = 0.530$ (0.004) was significantly lower than for both schedule 2 ($t(566) = -5.702$, $p_{holm} < 0.001$) and schedule 4 ($t(566) = -6.512$, $p_{holm} < 0.001$). Schedules 1 and 3 did not differ significantly ($t(566) = 2.197$, $p_{holm} = 0.372$), nor did schedules 2 and 4 ($t(566) = -2.471$, $p_{holm} = 0.202$),

Separate RM-ANOVAs to assess the effect of different genotypes on accuracy were carried out for DAT, COMT and SERT, with no main/interaction effect(s) involving DAT or COMT genotype on accuracy (all $p > 0.05$). However, SERT genotype had a significant effect on accuracy ($F(2,558) = 3.398$, $p = 0.034$, $\eta^2p = 0.012$), with higher mean accuracy for

individuals with the S/L allele ($\bar{x}(\sigma_{\bar{x}})$ = 0.626 (0.003)) compared to the S/S carriers ($\bar{x}(\sigma_{\bar{x}})$ = 0.613 (0.004), t(567) = 2.643, p_{holm} = 0.025) and numerically higher when compared with L/L carriers ($\bar{x}(\sigma_{\bar{x}})$ = 0.618 (0.005)), t(567) = 1.467, p_{holm} = 0.286) (Suppl. Fig. 3.3B). The above analysis was repeated with the inclusion of age as a covariate predictor variable. Age was a significant negative predictor of accuracy (F (1,557) = 13.106, p < 0.001, η^2p = 0.023). Interestingly, with the inclusion of age, volatility was no longer a significant predictor of accuracy (F (1,566) = 0.272, p = 0.602, η^2p < 0.001), driven by a significant negative relationship between age and accuracy in volatile phases (r = -0.155, p < 0.001). However, schedule, volatility by schedule and SERT genotype remained significant predictors.

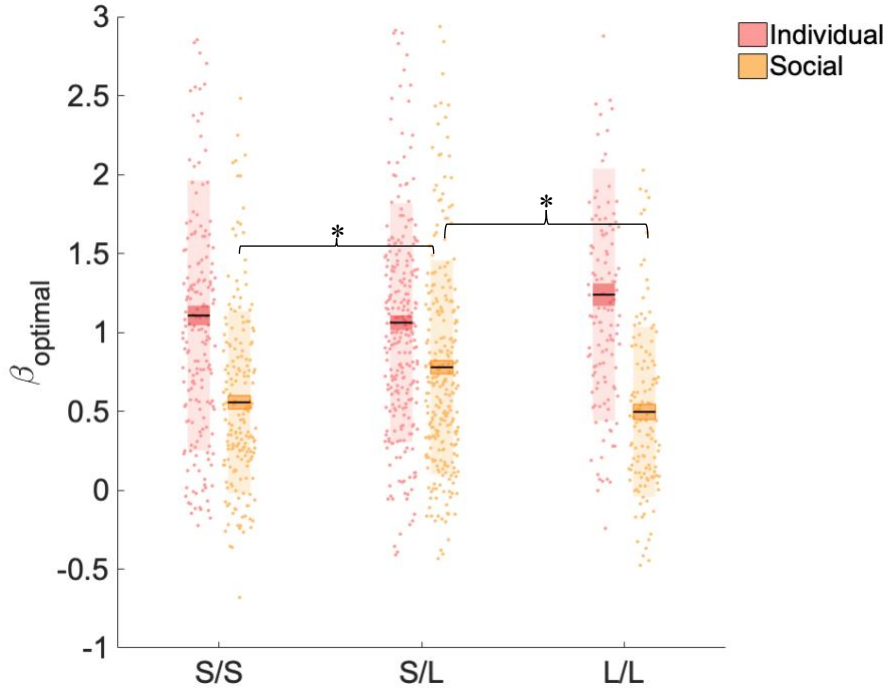


Supplementary Figure S3.3. A. Accuracy for stable and volatile phases. Accuracy was higher overall during the stable relative to the volatile environmental phases. **B. Accuracy as a function of SERT genotype**. Accuracy was higher for the S/L genotype: significantly compared to the S/S group and numerically compared to the L/L group. Data points indicate α estimates for individual participants (n = 570), boxes = standard error of the mean, shaded region = standard deviation, * indicates statistical significance (p < 0.05).

3.2.2. Optimal learner beta weights varied as function of SERT genotype

A RM-ANOVA with information type as the IV and ideal learner beta weights as the DV revealed significantly higher beta weights for individual ($\beta_{optimal\ \_individual}$ $\bar{x}(\sigma_{\bar{x}})$ = 1.144

(0.031)) compared with social information ($\beta_{optimal\_social}$: $\bar{x}(\sigma_{\bar{x}})$ = 0.618 (0.031), F(1, 567) = 107.624, p <0.001, η^2p = 0.160). Thus, the majority of participants were giving greater weight to individual information while making decisions. Separate RM-ANOVAs, with within-subject factor of information type (individual, social), and between-subject factor of genotype were carried out for DAT, COMT and SERT. No main or interaction effects were observed as a function of DAT or COMT genotypes, although there was a significant main effect of information type in both analyses. For analysis involving SERT, however, a significant SERT by information type interaction was observed (F (2, 567) = 7.338, p < 0.001, η^2p = 0.025) (Suppl. Fig. 3.4). $\beta_{optimal\_social}$ scores were significantly higher for individuals carrying the S/L genotype ($\bar{x}(\sigma_{\bar{x}})$ = 0.780 (0.042)), compared with both the S/S ($\bar{x}(\sigma_{\bar{x}})$ = 0.557 (0.043); t (567) = 3.241, p_{holm} = 0.006) and the L/L genotypes ($\bar{x}(\sigma_{\bar{x}})$ = 0.497 (0.048)); t(567) = 3.622, p_{holm} = 0.002). For $\beta_{optimal\_individual}$, however, beta weights did not significantly differ across genotypes (p_{holm} > 0.05). We thus investigated whether an increased weighting of social information was a driver of increased accuracy. A backwards regression model was deployed, with beta weights as predictor variables and accuracy as the dependent variable. Accuracy was significantly predicted by the full model (R = 0.731, F (2, 567) = 324.831, p < 0.001), with both $\beta_{optimal\_social}$ (t (567) = 20.200, p < 0.001) and $\beta_{optimal\_individual}$ (t (567) = 21.151, p < 0.001) significant positive predictors of accuracy.



Supplementary Figure S3.4. Optimal learner model. $\beta_{optimal\_social}$ varied as a function of SERT genotype. Data points indicate $\boldsymbol{\alpha}$ estimates for individual participants (n = 570), boxes = standard error of the mean, shaded region = standard deviation, * indicates statistical significance ($p < 0.05$).

3.2.3. SERT genotype: s - and l-carriers

Previous work has grouped together individuals with either one or two copies of the *S* (short) variant separately to those who are homozygous for the L (long) variant. As an exploratory measure, we re-ran all analyses with SERT genotype as a factor in two groups: S carriers (S/S and S/L) and L homozygotes (L/L). No effects of SERT genotype (S or L) were observed on accuracy (all $p > 0.05$). For optimal learner values, however, a RM-ANOVA with optimal beta weights as the DV and information and SERT genotype as the IVs revealed a significant SERT by information interaction ($F(1,568) = 8.765$, $p = 0.003$, $\eta^2p = 0.009$). Post hoc t-tests showed that $\beta_{optimal\_individual}$ varied significantly as a function of genotype, with higher $\beta_{optimal\_individual}$ for l-carriers ($\bar{x}(\sigma_{\bar{x}}) = 1.239\ (0.072)$) compared with s-carriers ($\bar{x}(\sigma_{\bar{x}}) = 1.081\ (0.038)$), $t(568) = 2.164$, $p_{holm} = 0.031$). Similarly, $\beta_{optimal\_social}$ varied as a function of genotype, with higher $\beta_{optimal\_social}$ for s-carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.688\ (0.031)$) compared with l-carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.497\ (0.048)$); $t(568) = 2.997$, $p_{holm} = 0.003$).

Finally, all model parameters were re-analysed. A RM-ANOVA with learning rate α as the DV and information, volatility and SERT genotype as the IVs revealed a significant main effect of information type ($F(1,568) = 180.896$, $p < 0.001$, $\eta^2p = 0.242$). No other main or interaction effects were observed (all $p > 0.05$, all $\eta^2p < 0.001$). There was, however, a significant effect of SERT on ζ values ($F(1,568) = 12.145$, $p < 0.001$, $\eta^2p = 0.021$), with higher values for s-allele carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.469\ (0.012)$) compared with l-allele carriers ($\bar{x}(\sigma_{\bar{x}}) = 0.376\ (0.024)$).

3.3. Instruction script - online SLT

In this game you can win points by making choices between two shapes. You will also see choices made by other players which you can use to help you with your choice. The entire game should take approximately 15 to 20 minutes. You should finish it in one session, and you should work on your own without the help of others.

Welcome. You have a choice: either choose the blue shape or the green shape. One shape is correct - guessing which one it is will give you points. Try picking a shape by clicking on one. [Participant responds]

Feedback: After you make a choice, the correct shape appears in the middle. You didn't really have a lot of information though. After each round you need to move your mouse into the dashed area to continue. Start the next round by moving your mouse into the area, then pick a shape. Maybe the same shape will be right again? [Participant responds]

Blue was right again! Things happen in phases in this game. Right now, it looks like you are

in a phase where blue is most likely to be correct. TIP: Here's a little tip - the shapes can be used by people who find it difficult to see the difference between blue and green. If you don't have a problem with seeing these colours you can ignore the shapes and just focus on the colours. Move your mouse into the area to start the next round and try picking blue again. [Participant responds]

And blue again! It certainly looks as though you are in a blue phase but make sure you pay attention to what the right answers are because the phase that you are in can change at any time.

TIP: Here's a tip - ignore which side of the screen the shapes are on - it's the colour that is important! Try again. Perhaps the other shape is right this time? [Participant responds]

Green! This time the green shape was right! The chance of each shape being right or wrong will change as you play, so pay attention! [Participant responds]

A Little Help. To help you decide - one of the shapes will be highlighted. This is the most popular choice selected by a group of 4 people who previously played this task. Here they think the green shape will be correct. Try picking a shape now. [Participant responds]

A caution. You see? This time the others got it right! Be careful though because we have mixed up the order of the other people's trials so that their choices will also follow phases. It looks like, right now, you could be in a phase where the group's information is useful - perhaps these are trials from the end of their experiment where they had developed a good idea of what was going on. [Participant responds]

Move your mouse into the dashed area and then try picking a shape. Do you think that the group's information will be useful? [Participant responds]

Yep, it certainly looks like the group information is useful right now but be careful, this could change! Sometimes you will see less useful information - for example from the beginning of their experiment where they didn't have a very good idea of what was going on.
Getting it right, gives you points. Get enough points and you could earn a silver or even a gold prize. [Participant responds]

Ready? No more practice. Now we'll start with the real game! Good luck.

Supplementary Table S3.1. Genotype frequencies (number and percentage)

SERT	N
L/L	123 (21.58 %)
S/L	262 (45.96 %)
S/S	185 (32.46 %)
DAT	
9/9	32 (5.61%)
9/10	208 (36.49 %)
10/10	330 (57.89 %)
COMT	
val/val	166 (29.12 %)
val/met	270 (47.37 %)
met/met	134 (23.51 %)

Note: The final sample which included complete genotype data for all three polymorphisms consisted of 693 subjects. A subset of participants was then excluded based on performances in the behavioural task, resulting in a final sample of 570.

Appendix 4

Supplementary material for Chapter 5

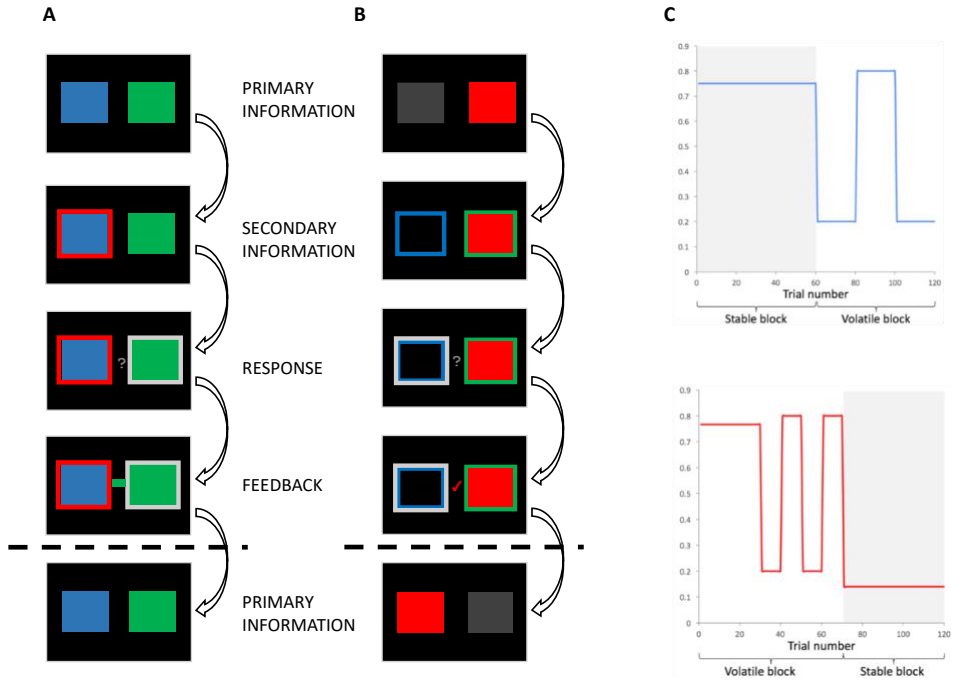
4.1 Social Learning Task

The behavioural task (social learning task (SLT)) lasted approximately 35 minutes. Participants were seated approximately 30cm from a computer screen. Stimuli were displayed using PsychToolBox and the task was programmed using MATLAB R2017b (The MathWorks, Natick, MA). Before the main task, participants completed a step-by-step on-screen practice task (10 trials) in which they learnt to choose between the two options to obtain a reward and learned that the "advice" represented by the frames could help in making the correct choice in some phases. To ensure that participants were making a conceptual distinction between the social and individual learning sources, participants were required to complete a short pre-task quiz (Appendix 1.1.3), testing their knowledge, after the practice task. Participants were required to repeat the practice round until they achieved 100% correct score in the quiz, meaning that all participants understood the structure of the task and that the red box represented *social* information. Furthermore, after the experiment, participants completed a feedback questionnaire (Appendix 1.1.4). Participants were informed as to whether they had earned a £5 bonus after the session. However, due to ethical considerations, all participants received the bonus.

Social-primary group

For the social-primary group the social information source was the primary source of learning. On each trial participants were presented with two grey placeholders. One placeholder was filled with a red box, indicating the group's choice. Blue/green frames then appeared around the placeholders. As in the individual-primary group, participants were informed that the task followed 'phases' wherein sometimes going with, but at other times going against, the group's choice was more likely to result in reward and sometimes the blue frame predominantly indicated the correct box, whereas at other times the green frame predominantly indicated the correct box. After making their choice participants received outcome information in the form of a tick/cross indicator (Suppl. Fig.4.1B). The indicator

primarily informed participants about whether the social group had been rewarded (and thus going with them would have resulted in points scoring but going against them would not) on the current trial. Whether the blue/green frame surrounded the correct or incorrect option could, secondarily, be inferred from the indicator. As in the individual-primary task, both the probability of reward associated with the blue/green stimuli and the utility of the social information varied according to probabilistic schedules (Appendix 4.1C). All other aspects of the task structure were the same as previously described in the individual-primary task group (displayed for reference – Suppl. Fig.4.1A).



Supplementary Figure S4.1. Social learning task. **A.** Individual-primary group. Participants selected between a blue and a green box to gain points. On each trial, the blue and green boxes were presented first. After 1-4 seconds (s), one of the boxes was highlighted with a red frame, representing the social information. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a green or blue box in the middle of the screen. **B.** Social-primary group. Participants selected between going with, or against a red box, which represented the social information. On each trial, the red box was displayed. After 1-4s, blue and green frames appeared. After 0.5–2s, a question mark appeared, indicating that participants were able to make their response. Response was indicated by a silver frame surrounding their choice. After a 1-3s interval, participants received feedback in the form of a tick or a cross. This feedback informed participants if going with the group was correct or incorrect, from this feedback participants could infer whether the blue or green frame was correct. **C.** Example of pseudo-randomised probabilistic schedule. The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of the blue box/frame being correct (top) and the probability of the red (social) box/frame being correct (bottom).

4.2 Pre-task quiz

(Correct answers are indicated with a star)

1. What does it mean if the green frame is around the red box?

A) The green box is the correct answer

B) Previous players believed the correct answer was green *

C) The left box is more likely to be correct

D) The answer you gave is correct

2. What does the red tick signify?

A) The most likely answer is red

B) The red box is incorrect

C) Nothing, the information is deliberately misleading

D) The group of previous players chose the correct answer *

3. What does it mean if the red box is on the left?

A) Nothing, the colour of the boxes changes at random.

B) The correct answer is on the left

C) The answer you gave is on the left

D) Previous players believed that the correct answer was the frame on the left. *

4.3 Post-task feedback quiz

(Likert scale 1-5)

1. Did you have a strategy? If so, what was it?

2. Did you use the group's suggestions (red box) to help you to make your decision?

3. Did you pay attention to which colour (green/blue) was more likely to be correct?

4. How hard did you find the task?

5. How clear were the task instructions? Were they easy to understand?

4.4 Instruction scripts

Social-primary group

Instruction script - version 1

Welcome. You have a choice between two shapes. One of the shapes is filled with red. This indicates the most popular choice selected by a group of 4 people who previously played this task. One shape is correct – this means that the group were either correct or incorrect. When the question mark appears, try picking a shape by pressing the left or right keyboard buttons. *[Participant responds]*

Feedback: After you make a choice, a tick or cross will appear in the middle. This tells you if the group of previous players were correct or incorrect. Here they think the blue shape (filled with red) will be correct. Try picking a shape now. *[Participant responds]*

The group were correct! This means that this time the others got it right and picked the correct colour.

Things happen in phases in this game. The game could be in a phase where the group are more likely to be correct. Have another go. *[Participant responds]*

The group were correct again! The blue shape was right again. It certainly looks as though you are in a phase where the blue phase but make sure you pay attention to what the right answers are because the phase that you are in can change at any time. Here's a tip - ignore which side of the screen the shapes are on - it's the colour that is important! *[Participant responds]*

The others got it right again. It looks like, right now, you could be in a phase where the group's information is useful. Perhaps these are trials from the end of their experiment, when they had developed a pretty good idea of what was going on. Be careful though because we have mixed up the order of the other people's trials so that their choices will also follow phases. Try again. Perhaps the other shape is right this time? *[Participant responds]*

Green! This time the green shape was right! The chance of each shape being right or wrong will change as you play, so pay attention! The group were incorrect this time. Remember that sometimes you will see less useful information from the group - for example from the beginning of their experiment where they didn't have a very good idea of what was going on. Have another go... *[Participant responds]*

This time the green shape was right! The chance of each shape being right or wrong will change as you play, so pay attention. The group were correct too. It looks like, right now, you could be in a phase where the group's information is useful. Try to be as accurate as possible. Getting it right, gives you points. Get enough points and you could earn a silver or even a gold prize! Have another go... *[Participant responds]*

Things happen in phases in this game. Remember, the tick or cross in the middle tells you if the group were correct or incorrect. That means that the shape with the red box was the correct choice. Have another go... *[Participant responds]*

The group were correct this time. The tick in the middle tells you that they picked the correct choice. There will now be a short quiz. Pick one more shape and then we'll head to the real game! *[Participant responds]*

Instruction script - version 2

Welcome. You have a choice between going with, or against advice from a group. Below you can see a blue and green frame, one frame is filled with a red box: this indicates the most popular choice selected by a group of 4 people who previously played this task. One frame is correct. You can pick the same frame as the group have picked or choose to go against the group's advice. When the question mark appears, make your selection by pressing the left or right keyboard buttons. *[Participant responds]*

Feedback: After you make a choice, a tick or cross will appear in the middle. This tells you if the group of previous players were correct or incorrect. This time they were correct! This means that the frame filled with the red square was the correct frame. Here they think the blue frame (filled with red) will be correct. Try picking a frame now. *[Participant responds]*

The group were correct! This means that this time the others got it right and picked the correct colour. Things happen in phases in this game. The game could be in a phase where the group are more likely to be correct. Have another go. *[Participant responds]*

The group were correct again! The blue frame was right again. It certainly looks as though you are in a phase where the group are correct but make sure you pay attention to the feedback because the phase that you are in can change at any time. Blue and green can also go through phases: it looks like you might be in a phase where the blue frame is more likely to be correct. Try again. *[Participant responds]*

The others got it right again. It looks like, right now, you could be in a phase where the group's information is pretty useful. Perhaps these are trials from the end of their experiment, when they had developed a pretty good idea of what was going on. Be careful though because we have mixed up the order of the other people's trials so that their choices will follow phases. Try again. *[Participant responds]*

The group were incorrect this time. This time the green frame was correct. The chance of each frame being right or wrong will change as you play, so pay attention! Remember that sometimes you will see less useful information from the group - for example from the beginning of their experiment where they didn't have a very good idea of what was going on. Have another go... *[Participant responds]*

The group were correct this time. The chance of each frame being right or wrong will change as you play, so pay attention. Try to be as accurate as possible.

Getting it right, gives you points. Get enough points and you could earn a silver or even a gold prize! Have another go... *[Participant responds]*

Things happen in phases in this game. Remember, the tick or cross in the middle tells you if the group were correct or incorrect. That means that the frame filled with the red was the correct choice. Have another go... *[Participant responds]*

The group were correct this time. The tick in the middle tells you that they picked the correct choice. There will now be a short quiz. Pick one more time and then we'll head to the real game! [*Participant responds]*

Appendix 5

Supplementary material for Chapter 6

Dopaminergic challenge dissociates learning from primary versus secondary sources of information

Rybicki, A. J., Sowden, S. L., Schuster, B. A., & Cook, J. L.

eLife 2022;11:e74893 DOI: 10.7554/eLife.74893

Appendix 5.1

Inclusion criteria

Participant is willing and able to give informed consent for participation in the study.

Aged between 18 and 45.

BMI in the range of 18.5 – 29.5
Resting blood pressure in the range of 90/60 (low) to 140/90 (high)

Electrocardiogram QT (heart rate corrected) interval < .42

Exclusion criteria

Participation in another drug study in the 3 weeks previous.

Personal or first-degree family history of cardiovascular disease, specifically hypotension, arrhythmias or valvular disease, stroke

Neurological abnormalities or traumas, kidney disease or liver disease

Inherited blood conditions

Psychiatric or psychological conditions (including depression and anxiety disorders)

Known learning disability

Anybody found to have an elongated Q-T interval following single lead ECG examination

Low heart rate

Low or high blood pressure

Any regular medication - excluding the oral contraceptive pill

Recent recreational drugs use or alcohol and drug dependency

Known allergy to any medication

Current pregnancy or breastfeeding

Previous participant in a drug study

Lack of sleep in last 24 hours.

Lack of food or drink in last 12 hours

Primary sensory impairment (e.g., uncorrected visual or hearing impairment)

Lactose intolerant

Insufficient English to be able to consent to take part in the study

Baseline cognitive measures and mood ratings

Approximately one week prior to drug/placebo administration, participants completed a battery of self-report questionnaire measures: Autism Spectrum quotient (AQ)[124], Toronto Alexithymia Scale (TAS 20) [280], Behavioural Inhibition/Activation Scale (BIS-BAS) [434], the Depression Anxiety and Stress Scale (DASS 21) [435], Interpersonal Reactivity Index (IRI) [436], Beck's Depression Inventory (BDI) [437] and Body Perception Questionnaire (BPQ) [438]. Self-report questionnaire scores are summarised in Table 1. The individual-primary group did not differ significantly on any measure from the social-primary group. The group that received haloperidol (HAL) on day 1 did not differ significantly on any of the baseline measures from the group that received placebo (PLA) on day 1 ($p < 0.05$). Mood and fatigue were monitored three times per day during each test day, i) before capsule intake, ii) two hours post-capsule intake upon start task battery, and iii) upon completion of the task battery. The mood ratings consisted of the Positive and Negative Affect Scale (PANAS) [439]. A self-report scale was used to monitor fatigue. 24% of participants reported that they did not know on which day they had taken an active drug. Out of the remaining participants, 84% of participants correctly reported that they thought they had received an active drug. No adverse side effects were reported. Blood pressure, heart rate and blood oxygenation levels were monitored five times over the course of the testing days; before drug/placebo administration, and then at one, two and three and a half hour intervals thereafter. Measures were taken for a final time immediately before the end of the testing day.

Appendix 5.1 - Table 1. Self-report questionnaire scores (n = 31)

Self-report questionnaires	Individual-primary group	Social-primary group	t (29)	p-value
AQ	9.412 (4.556)	6.500 (4.179)	1.910	0.065
TAS-20	39.529 (6.947)	40.313 (7.981)	-0.301	0.765
BIS-BAS	50.647 (6.855)	51.125 (5.536)	-0.219	0.828
DASS-Stress	3.176 (4.231)	3.875 (2.306)	-0.583	0.723
DASS-Anxiety	1.353 (2.178)	1.938 (2.516)	-0.715	0.564
DASS-Depression	1.706 (1.863)	2.313 (3.005)	-0.702	0.480
IRI	66.235(15.114)	66.375(10.645)	-0.031	0.976
BDI	3.176 (3.746)	3.438 (2.732)	-0.227	0.822
BPQ	52.176(29.473)	46.688(18.650)	0.635	0.221

Note: Mean (standard deviation) scores are reported. Significance level for the between-group differences are reported. Autism Spectrum quotient (AQ) [124], Toronto Alexithymia Scale (TAS 20) [280], Behavioural Inhibition/Activation Scale (BIS-BAS)[434], the Depression Anxiety and Stress Scale (DASS 21) [435], Interpersonal Reactivity Index (IRI) [436], Beck's Depression Inventory (BDI) [437] and Body Perception Questionnaire (BPQ) [438].

Drug effects on mood and tiredness

Positive and negative affect (PANAS) scores were submitted to separate RM-ANOVAs, with within-subjects (WS) factors time (baseline/start testing/end testing) and drug (HAL/PLA). For both positive and negative scores, a main effect of time was observed. Both positive ($F(2,62) = 8.286$, $p < 0.001$, $\eta_p^2 = 0.211$), and negative scores decreased over time ($F(2,62) = 6.020$, $p = 0.004$, $\eta_p^2 = 0.163$). A drug by time interaction was observed for positive scores ($F(2,62) = 7.353$, $p = 0.001$, $\eta_p^2 = 0.192$), with simple effects analysis demonstrating that positive scores decreased over time under haloperidol ($p < 0.001$), but not placebo ($p = 0.994$). A main effect of drug was observed on negative scores ($F(1,31) = 4.749$, $p = 0.037$, $\eta_p^2 = 0.133$), with higher negative affect scores under haloperidol ($\bar{x}\ (\sigma_{\bar{x}}) = 10.771\ (0.557)$) compared with placebo ($\bar{x}\ (\sigma_{\bar{x}}) = 9.491(0.557)$).

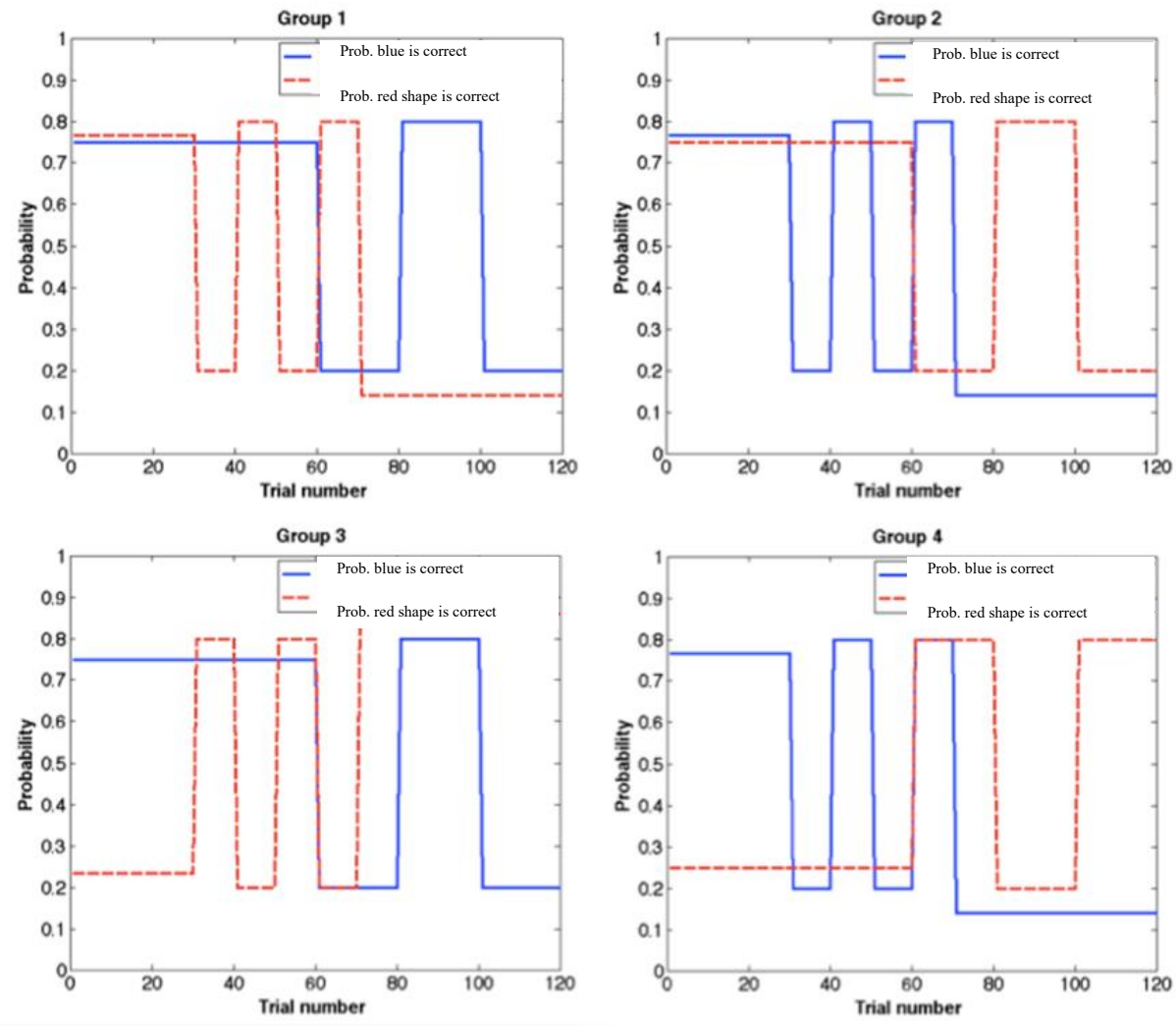
Self-reported fatigue ratings (Likert scale: 1-10, with higher scores referring to higher levels of fatigue) were submitted to a RM-ANOVA, with WS factors time (T1-T5) and drug (HAL/PLA). A main effect of time was observed, with fatigue rising across time ($F(4,88) = 6.652$, $p < 0.001$, $\eta_p^2 = 0.232$). No main or interaction effect(s) involving drug were observed.

Appendix 5.2

Randomisation groups

For both the social-primary and individual-primary group, the probability of reward associated with the blue/green stimuli (individual information) and the red stimuli (social information) were governed by different pseudo-randomisation schedules, adapted from Behrens et al [95]. Schedules were counterbalanced between participants to ensure that learning could not be explained in terms of differences in learning between schedules with increased/decreased, or early/late occurring, volatility. The individual-primary group (schedules 1,3) were sub-divided into two groups, such that half started with predominantly correct social information, and half with predominantly incorrect social information, with the same true for the social-primary group (schedules 2,4). The primary information source was always less volatile overall compared to the secondary information source, irrespective of whether it was social or individual. To give an example, the randomisation schedule for group 1 was the same as that employed by Behrens et al [95]. During the first 60 trials, the individual reward history was stable, with a 75% probability of blue being correct. During the next 60 trials, the reward history was volatile, switching between 80% green correct and 80% blue correct every 20 trials. Meanwhile, during the first 30 trials, social information was stable, with 75% of choices being correct. During the next 40 trials, the social information was volatile, switching between 80% incorrect and 80% correct every 10 trials. During the final 50 trials, social information was once again stable, with 85% of choices being incorrect. Randomisation schedules for groups 2, 3, and 4 were inverted and counterbalanced versions of schedule 1 (Appendix 5.2 - Fig. 1).

Appendix 5.2 - Figure 1.



Appendix 5.2- Figure 1. Randomisation schedules. The probability of reward varied according to probabilistic schedules, including stable and volatile blocks for both the probability of blue being correct and the probability of the social information indicating the correct answer. Probability schedules were counterbalanced between participants. Solid blue lines show the probability of blue being the correct choice, dashed red lines show the probability of the social information being correct. Schedules 1-4 are displayed here.

Supplementary Methods

i. Visual working memory task

Participants completed a visual working memory (VWM) task, adapted from the Sternberg VWM Task [400], and programmed using MATLAB R2017b. Participants were first presented with instructions followed by practice trials. Upon completion of the practice trials, participants completed 60 experimental trials across 5 blocks. On each trial, a fixation cross was displayed in the centre of screen (fixation duration varied randomly between 500-1000 ms). Then participants were presented with a list of letters, (varying between 5 – 9 consonants in length, with letters randomly selected from the alphabet on each trial) for 1000 ms, followed by a blue fixation cross for 3000 ms. Following this, a single test letter was displayed (for a maximum of 4000 ms), requiring participants to determine whether the letter was taken from the previously displayed list. For 50% of trials, the letter had been present on the previous list and on 50% of trials, it had not. Participants responded by pressing 1-3 on the keyboard (1 – Yes, 2 - No, 3 – Unsure). The total task duration was approximately 10 minutes. Responses (accuracy) and response time (time from test letter displayed until participant response) were recorded for each trial.

ii. Go-NoGo learning

An adapted version of a probabilistic Go/No-Go Task [215] was employed, presented using MATLAB R2017b. In this task, a 'Go' response measures sensitivity to reward, whereas a 'No-Go' response measures sensitivity to punishment. Participants were presented with 4 different stimuli, each with a probabilistic value of reward (80%, 60%, 40%, 20%) and instructed to accumulate as many points as possible and to avoid losing points, achieved by selecting or withholding a response to the given stimuli. For example, if selected, stimuli A would result in gaining a point on 80% of trials and losing a point on 20% of trials. Participants were informed that points would be rewarded with monetary compensation; however, due to ethical considerations, all participants were awarded £5 at the end, regardless of task performance. Participants first completed 4 blocks of a practice stage, where single stimuli were presented (40 trials/block, with each stimulus presented 10 times per block). Reward feedback was provided, allowing learning of the probabilistic value of each stimulus. This was followed by 6 testing blocks (40 trials/block) displaying either single stimuli

(training stimuli) or novel pairs of stimuli on each trial, whereby participants were required to respond based on the *combined* probabilistic value of the pairs. Testing blocks contained positive pairs with a high associated probabilistic reward value, equal pairs (equally probable reward value), and negative pairs, with a high probabilistic value for punishment. Participants could respond via a 'Go' (space bar press) or 'No-Go' (withhold response) response. Feedback was not provided during testing blocks. In all trials, a fixation cross was presented for 250-750ms, followed by stimuli presentation for 1000ms and a response period for 250ms. Task performance was calculated as the difference in 'Go' response for stimuli (novel pairs and single stimuli) with a high probability of reward under HAL and PLA conditions, for each participant separately.

Appendix 5.3

Model fitting

Optimisation of free parameter values was performed as per Cook and colleagues [93], using a quasi-Newton optimisation algorithm specified in TAPAS toolbox - quasinewton_optim_config.m. The function maximised the log-joint posterior density over all parameters given the data and the generative model. **α** values were estimated in logit space (see tapas_logit.m), i.e., a logistic sigmoid transformation of native space (tapas_logit(x) = ln(x/(1-x)); x = 1/(1+exp(-tapas_logit(x)))). An uninformative prior, allowing for individual differences in learning rate was used for **α:** tapas_logit (0.2, 1), with a variance of 1. Initial values were set at logit (0.5, 1), with a variance of 1. Initial values were allowed to vary, to allow for inter-individual differences in prior preferences for the extent to which individual would conform to the group choice. The prior for **β** was set to log (48), with a variance of 1, and the prior for ζ was set at 0 with a variance of 10^2 (logit space), i.e., an equal weighting for information derived from primary and secondary learning (0.5). Prior choices were based on previous work [93]. Maximum-a-posteriori (MAP) estimates for all model parameters were calculated using the HGF toolbox version 3 (https://osf.io/398w4/files/). All code used is adapted from the open-source software package TAPAS (available at http://www.translationalneuromodeling.org/tapas).

Model comparison

We based our choice of perceptual model on previous work by Cook and others [93], wherein a systematic comparison of three alternative models was conducted, to determine which best explained observed choice behaviour. Here we repeated Cook et al.'s model comparison and added four further extensions of the classic model, thus we compared eight alternative models in total. A formal model comparison was carried out using Bayesian model selection using the VBA toolbox [440].

Data were initially analysed with eight models. All models were variations of the classic Rescorla-Wagner model. Group level Bayesian model selection (BMS) was used to evaluate which model provided the (relative) best fit to the observed data. The VBA toolbox [431], specifically random-effects BMS (using the VBA_groupBMC_btwConds.m function), was utilised. Random effects group BMS computes an approximation of the model evidence relative to the other models, i.e., the probability of the data y given a model m, p($y|m$), with log model evidence here approximated with F values. The posterior probability that a model has generated the observed data, relative to other models is estimated, and the exceedance probability, or the likelihood that a given model is more likely than other included models in the set, is estimated. Analysis across both conditions allows us to test the hypothesis that the same model produced observed data under both haloperidol and placebo conditions.

Model 1 was a classic Rescorla-Wagner model:

$$V_{(i+1)} = V_i + \alpha\varepsilon_i$$

with $\varepsilon_i = r_i - V_i$, the difference between the actual and the expected reward or prediction error (PE).

Model 2 was an extension of Model 1, with separate learning rates (α) for learning from primary value and secondary value learning sources:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary}\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)} + \alpha_{secondary}\varepsilon_i$$

Model 3 had a single learning rate α for primary/secondary learning, but separate learning rates for volatile and stable blocks:

$$V_{(i+1)} = V_i + \alpha_{volatile}\varepsilon_i + \alpha_{stable}\varepsilon_i$$

Model 4 had four separate learning rates α for volatile and stable and primary and secondary learning:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary_{volatile}}\varepsilon_i + \alpha_{primary_{stable}}\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)} + \alpha_{secondary_{volatile}}\varepsilon_i + \alpha_{secondary_{stable}}\varepsilon_i$$

As an exploratory measure, we further extended Models 1-4 to include separate learning rates corresponding to learning from rewarded trials and unrewarded trials separately, i.e., learning from wins and losses.

Model 5:

$$V_{(i+1)} = V_i + \alpha_{reward}\,\varepsilon_i + \alpha_{unreward}\,\varepsilon_i$$

Model 6:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary_{reward}}\,\varepsilon_i + \alpha_{primary_{unreward}}\,\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)} + \alpha_{secondary_{reward}}\,\varepsilon_i + \alpha_{secondary_{unreward}}\,\varepsilon_i$$

Model 7:

$$V_{(i+1)} = V_i + \alpha_{volatile_{reward}}\varepsilon_i + \alpha_{stable_{reward}}\varepsilon_i + \alpha_{volatile_{unreward}}\varepsilon_i + \alpha_{stable_{unreward}}\varepsilon_i$$

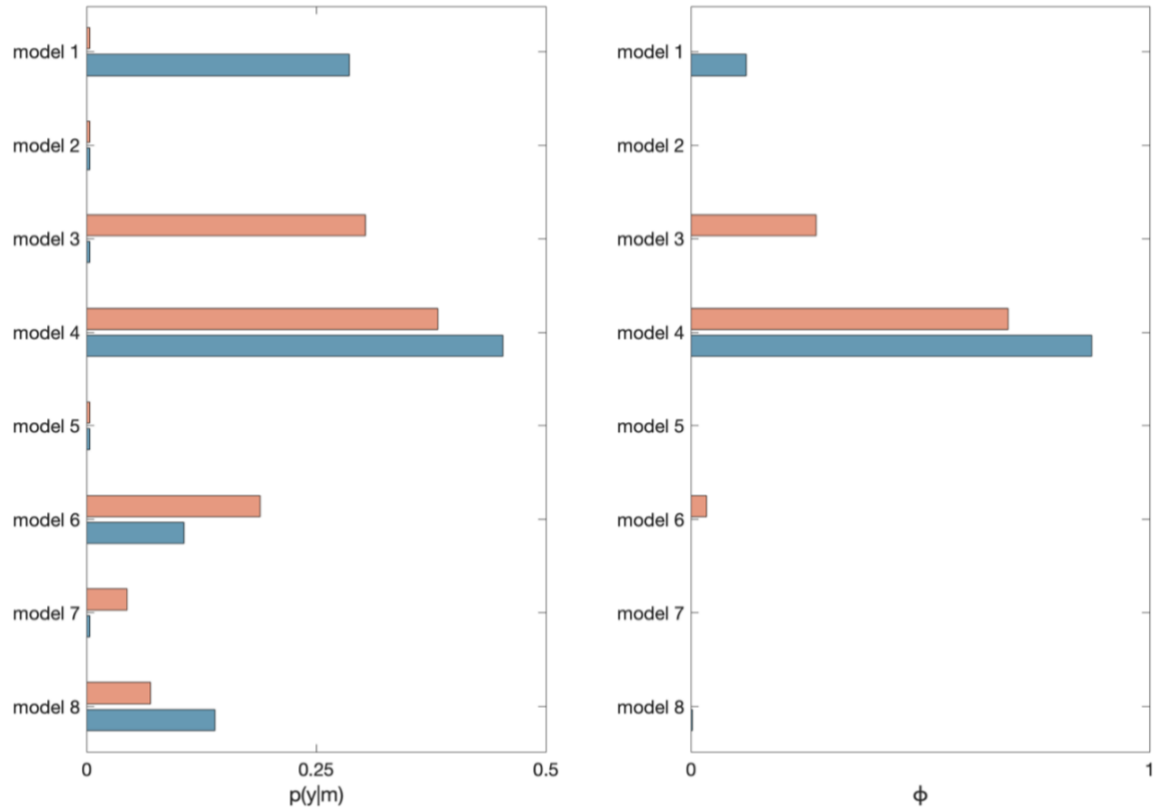
Model 8:

$$V_{primary(i+1)} = V_{primary(i)} + \alpha_{primary_{volatile_{reward}}}\varepsilon_i + \alpha_{primary_{stable_{reward}}}\varepsilon_i + + \alpha_{primary_{volatile_{unreward}}}\varepsilon_i + \alpha_{primary_{stable_{unreward}}}\varepsilon_i$$

$$V_{secondary(i+1)} = V_{secondary(i)}V_{(i+1)} + \alpha_{secondary_{volatile_{reward}}}\varepsilon_i + \alpha_{secondary_{stable_{reward}}}\varepsilon_i + \alpha_{secondary_{volatile_{unreward}}}\varepsilon_i + \alpha_{secondary_{stable_{unreward}}}\varepsilon_i$$

We ran a between-groups model comparison, to ensure that the same model could explain the observed data under both placebo and haloperidol. When comparing all models, Model 4 performed best, with an exceedance probability approaching 1. The exceedance probability that the same model (Model 4) had produced data under both conditions was equal to 1. For condition 1 (placebo), the posterior probabilities that the observed data had produced the model was equal to 10.329 for Model 3 and 12.998 for Model 4, with the probability that the data was produced by the winning model p(H1|y) = 0.762. For group 2 (haloperidol), Model 4 had a posterior probability of 15.417 (p(H1|y) = 0.998). For the between-groups assessment, the posterior probability p(H1|y) = 0.999 and the protected exceedance probability (ϕ) was equal to 0.999.

Appendix 5.3- Figure 1

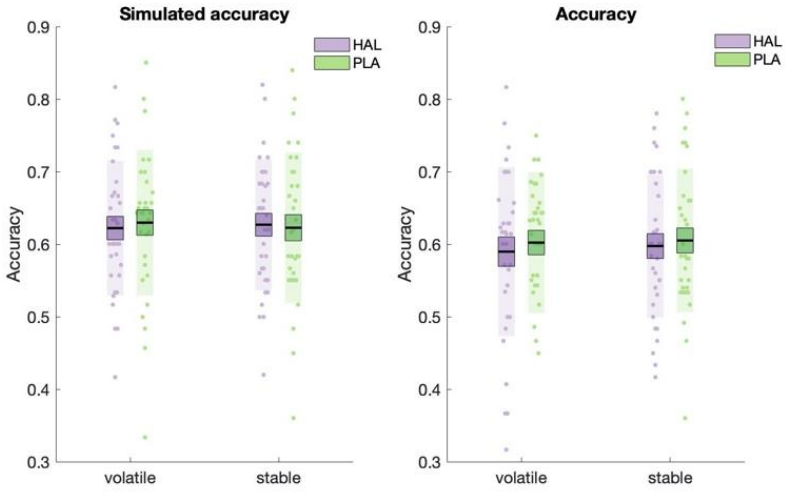


Appendix 5.3 - Figure 1. Model comparison. Results from random-effects Bayesian model selection. Exceedance Probability and posterior model probability for models 1-8. p($y|m$) = posterior model probability, ϕ = exceedance probability, HAL = blue, PLA = red.

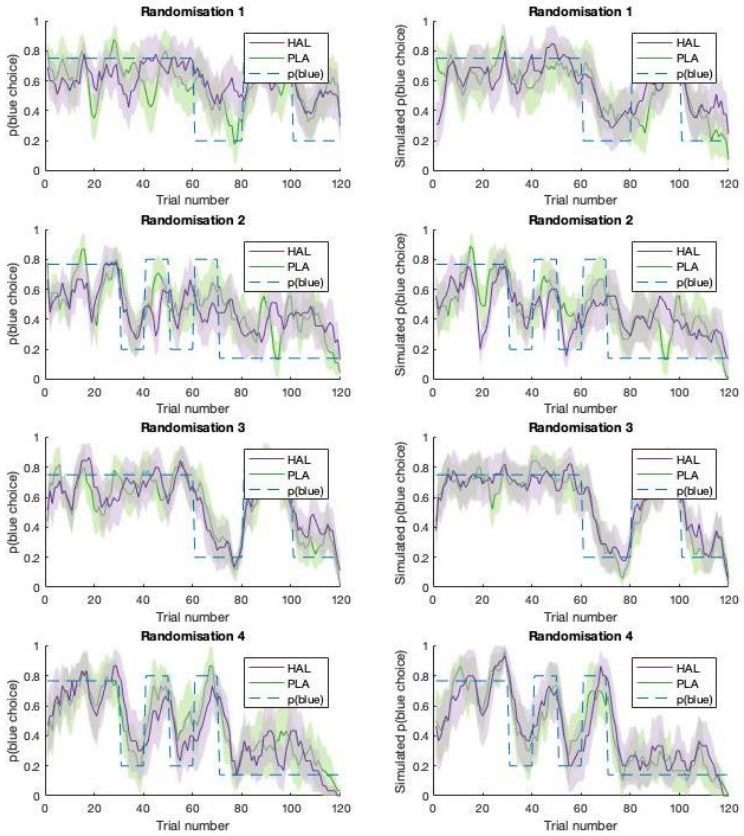
Model Validation

To demonstrate that the chosen model (model 4) accurately described participant behaviour, we simulated response data for each participant, using estimated model parameter values (tapas_simModel.m). Accuracy did not significantly differ between actual and simulated accuracy for PLA (t = -0.866, p = 0.394) or HAL conditions (t = -0.280, p = 0.781) (Appendix 5.3 -Fig. 2A). Simulated and calculated accuracy were significantly correlated for each participant, under both placebo (r = 0.487, p = 0.005) and haloperidol conditions (r = 0.712, p <.001) (Appendix 5.3 -Fig. 2B).

A


Simulated accuracy
Accuracy
HAL
PLA
Accuracy
0.9
0.8
0.7
0.6
0.5
0.4
0.3
volatile
stable

B


Randomisation 1
Randomisation 2
Randomisation 3
Randomisation 4
HAL
PLA
p(blue)
p(blue choice)
Simulated p(blue choice)
Trial number

Appendix 5.3 - Figure 2. A. Model simulations (left) and participant response data (right). Mean accuracy is displayed separately for volatile and stable environmental phases, under HAL (purple) and PLA (green). Boxes = standard error of the mean, shaded region = standard deviation, individual datapoints are displayed. HAL = haloperidol, PLA = placebo. **B.** Participant data (left) juxtaposed against model simulations (right) Running average, across 5 trials of blue choices for probabilistic randomisation schedules 1 to 4. Shaded region = standard error of the mean.

In addition, to formally test model predictions of choice behaviour, for each participant we calculated the average value that the model estimated for the options chosen by the participant (collapsed across HAL and PLA conditions), and the average value that the model estimated for the options that were not chosen by the participant. If the chosen model was accurately describing participants' choice behaviour, then average estimated values for chosen options should be significantly higher than for the unchosen options. Indeed, a paired samples t-test illustrated that, model-derived values for chosen options ($\bar{x}(\sigma_{\bar{x}})$ = 0.607 (0.008)) were significantly greater than values for unchosen options ($\bar{x}(\sigma_{\bar{x}})$ = 0.393 (0.008); $t(30) = 12.558$, $p < 0.001$).

To ensure that parameter estimates could be recovered, we simulated response data for each participant, based on estimated model parameters, using the function tapas_simModel.m from the TAPAS toolbox. Model parameters were subsequently estimated from simulated data and averaged over 100 iterations for each participant, separately for HAL and PLA conditions. All recovered parameters correlated significantly with estimated parameters under both HAL ($\alpha_{primary:}$ $r = 0.991$, $p < 0.001$, $\alpha_{secondary:}$ $r = 0.961$, $p < 0.001$) and PLA ($\alpha_{primary:}$ $r = 0.975$, $p < 0.001$, $\alpha_{secondary:}$ $r = 0.984$, $p < 0.001$) treatment conditions. A RM ANOVA on recovered parameters showed the same pattern of results as with estimated parameters, including a significant interaction effect for our main interaction of interest (drug by information source: ($F(1,29) = 4.027$, $p = 0.054$, $\eta_p^2 = 0.122$)).

Appendix 5.4

Extended statistical analyses

i. Learning rate analysis (n = 41)

A RM-ANOVA, with (square-root transformed) learning rate (α) as the DV and predictors information source, volatility, drug, and group was carried out on estimates from the mixed model analysis which included all participants who completed at least one study day (N = 41). A significant main effect of information was observed (F (1,234) = 3.944, p = 0.048, beta estimate ($\sigma_{\bar{x}}$) = 0.019 (0.010); t = 1.986, CI [0 - 0.04]), with higher mean values for $\alpha_{primary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.429 (0.018)) compared with $\alpha_{secondary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.391 (0.018)).

A significant volatility by information interaction (F (1, 234) = 4.676, p = 0.032, beta estimate ($\sigma_{\bar{x}}$) = 0.021 (0.010), t = -2.162, CI [0 - 0.04]) was observed. Post hoc comparisons revealed that, under stable phases, $\alpha_{primary}$ values ($\bar{x}(\sigma_{\bar{x}})$ = 0.461 (0.023)) were significantly greater than $\alpha_{secondary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.381 (0.023); z = 2.933, p_{holm} = 0.007), with no difference between α in volatile phases (z = -0.125, p_{holm} = 0.901). No main effect of group was observed, however, there was a significant information by group interaction (F (1, 234) = 32.471, p < 0.001, beta estimate ($\sigma_{\bar{x}}$) = 0.05 (0.010); t = 5.700, CI [0.04-0.07]). Post hoc comparisons revealed that, for the individual-primary group, $\alpha_{primary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.455 (0.026)) was significantly greater than $\alpha_{secondary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.307 (0.026); z = 5.351, p_{holm} < 0.001). For the social-primary group, however, $\alpha_{secondary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.475 (0.025)) was significantly greater than $\alpha_{primary}$ ($\bar{x}(\sigma_{\bar{x}})$ = 0.404 (0.025); z = 2.667, p_{holm} = 0.015).

A significant volatility by group interaction was observed (F (1,234) = 4.168, p = 0.042, beta estimate ($\sigma_{\bar{x}}$) = 0.020 (0.010); t = 2.042, CI [0 - 0.04]). For the individual-primary group, $\alpha_{volatile}$ (estimate ($\sigma_{\bar{x}}$) = 0.351 (0.026)) was (marginally) significantly lower than α_{stable} (estimate ($\sigma_{\bar{x}}$) = 0.41 (0.026); z = -2.192, p_{holm} < 0.057). For the social-primary group, however, $\alpha_{volatile}$ (estimate ($\sigma_{\bar{x}}$) = 0.449 (0.025)) and α_{stable} (estimate ($\sigma_{\bar{x}}$) = 0.431 (0.025)) did not significantly differ (z = 0.672, p_{holm} = 0.502). Most importantly, as with the analysis reported in the main text, a significant drug by information interaction was observed (F (1,234) = 3.727, p = 0.054, beta estimate ($\sigma_{\bar{x}}$) = 0.020 (0.010); t = 1.93, CI [0.00 – 0.04]. Post hoc comparisons demonstrated that, under PLA there was a significant difference between $\alpha_{primary}$ (beta estimate ($\sigma_{\bar{x}}$) = 0.451 (0.023)) and $\alpha_{secondary}$ (estimate ($\sigma_{\bar{x}}$) = 0.375 (0.023); z = 2.727, p_{holm} = 0.026, uncorrected p = 0.006). This difference was nullified under HAL

($\alpha_{primary}$ estimate ($\sigma_{\bar{x}}$) = 0.408 (0.023) and $\alpha_{secondary}$ (estimate ($\sigma_{\bar{x}}$) = 0.407 (0.023); z = 0.040, p_{holm} = 0.968, uncorrected p = 0.968). There was no significant group x information source x drug interaction (F (1,234) = 0.029, p = 0.866, beta estimate ($\sigma_{\bar{x}}$) = -0.002 (0.010); t = -0.169, CI [-0.02 - 0.02]).

ii. Accuracy

An analysis of accuracy was conducted in participants who had completed both study days (n=31), to explore whether there was any systematic variation as a function of randomization schedule, and across drug and placebo conditions and volatile and stable phases. A RM-ANOVA, with within-subjects factors drug (HAL, PLA) and volatility (stable, volatile), and between-subjects factor group (social-primary, individual-primary) and randomisation schedule (1-4), demonstrated no difference in accuracy between haloperidol ($\bar{x}(\sigma_{\bar{x}})$ = 0.601(0.011)), and placebo ($\bar{x}(\sigma_{\bar{x}})$ = 0.614 (0.011); F (1,27) = 1.161, p = 0.291, η_p^2 = 0.041). However, a significant main effect of schedule was observed (F (3,27) = 3.004, p = 0.048, η_p^2 = 0.250), with the lowest accuracy observed for schedule 1 ($\bar{x}(\sigma_{\bar{x}})$ = 0.558 (0.019)). Although accuracy for schedule 1 was lower than for schedule 2 ($\bar{x}(\sigma_{\bar{x}})$ = 0.619 (0.018); t (27) = -2.358, p_{holm} = 0.129), schedule 3 ($\bar{x}(\sigma_{\bar{x}})$ = 0.614 (0.018); t(27) = -2.162, p_{holm} = 0.159) and schedule 4 ($\bar{x}(\sigma_{\bar{x}})$ = 0.637 (0.020); t(27) = -2.748, p_{holm} = 0.063); these differences were no longer significant after correction for multiple comparisons. Mean accuracy for schedules 2-4 did not significantly differ from each other (all p-values = 1.000). In addition, there was a significant interaction effect between schedule and volatility (F (3,27) = 7.527, p < 0.001, ηp^2 =0.455). For all schedules except for schedule 3, there was no significant difference in accuracy between volatile and stable phases (all p>0.05). However, for schedule 3, accuracy was significantly higher for volatile ($\bar{x}(\sigma_{\bar{x}})$ = 0.675 (0.022)) over stable phases ($\bar{x}(\sigma_{\bar{x}})$ = 0.533 (0.022); t(27) = (3.656), p_{holm} = 0.027). Accuracy was significantly higher for the social-primary group ($\bar{x}(\sigma_{\bar{x}})$ = 0.629 (0.013)), compared with the individual-primary group ($\bar{x}(\sigma_{\bar{x}})$ = 0.586 (0.013); F (1,29) = 5.196, p = 0.030, η_p^2 = 0.152) and no other main effects or interactions were observed (all p>0.05).

iii. Relationship between accuracy scores and parameters from model-based analyses

A backwards regression with PLA accuracy as the dependent variable, and $\alpha_{primary}$ and $\alpha_{secondary}$ (collapsed across volatile and stable phases), initial values $V_{primary(i)}$ and $V_{secondary(i)}$, β and ζ as predictors, was carried out. PLA accuracy was marginally significantly predicted by a model with $\alpha_{secondary}$ as a single predictor ($R = 0.347$, $F(1,29) = 3.981$, $p = 0.055$). Under haloperidol, a backward regression with HAL accuracy as the dependent variable, and $\alpha_{primary}$, $\alpha_{secondary}$, $V_{primary(i)}$, $V_{secondary(i)}$, β and ζ as predictors, revealed that HAL accuracy was significantly predicted by the full model. Within the model, $\alpha_{primary}$ was the only significant predictor (Table 2). Removing predictors did not significantly improve the fit of the model (R^2change < 0.001, F change (1,25) = -0.064, p = 1.000).

Appendix 5.4 - Table 1

Coefficients from regression model with HAL accuracy as the dependent variable.

	β	β (SEM)	standardised β	t	p
constant	0.431	0.089		4.840	<0.001
$\alpha_{primary}$	0.195	0.077	0.431	2.532	0.018*
$\alpha_{secondary}$	0.076	0.119	0.127	0.642	0.527
$V_{primary(i)}$	0.121	0.090	0.230	1.342	0.192
$V_{secondary(i)}$	0.033	0.131	0.050	0.249	0.806
β	0.002	0.001	0.329	1.698	0.102
ζ	0.045	0.043	0.189	1.066	0.297

Note: * indicates statistical significance

iv. Go, No-go control task

To further investigate the neurochemical mechanisms underlying the observed decrease in α_{primary} under haloperidol, we measured performance on a probabilistic Go, No-go control task, adapted from Frank and colleagues[215]. Previous research (using a similar low, acute dose of haloperidol) resulted in enhancement of learning from positive reinforcement, indexed by an increase in learning from positive feedback [215], suggested to be mediated via pre-synaptic antagonistic effects on phasic dopamine (DA) signalling. As an exploratory measure, participants were stratified into two subgroups based on performance during this task; those with a higher change in 'Go' performance for high reward trials under haloperidol, and those with a lower change in 'Go' performance under haloperidol, relative to placebo. For the participants who demonstrated increased 'Go' performance under haloperidol (n = 12), a significant drug by information effect was observed on the main behavioural task ($F(1,10) = 4.773$, $p = 0.054$, $\eta_p^2 = 0.323$). However, this effect was not observed in participants with reduced 'Go' performance under haloperidol (n = 19; $F(1,17) = 2.001$, $p = 0.175$, $\eta_p^2 = 0.105$). Thus, suggesting that the observed effect of haloperidol on learning rate for primary information was driven by a subgroup of participants who exhibited increased 'Go' performance under haloperidol (relative to placebo). Given that such effects on Go performance have been linked to pre-synaptic antagonistic effects on phasic DA signalling [215] these results suggest that the effects we observed on α_{primary} are likely mediated by effects of haloperidol on phasic DA signalling.

While an increase in Go performance suggests pre-synaptic effects of haloperidol on phasic dopamine release, the effects of haloperidol are also mediated via antagonism of heteroreceptors on non-dopaminergic neurons[215], resulting in a reduction in tonic dopamine signalling. These tonic effects are commonly indexed by a slowing of response [383],[441]. Indeed, haloperidol had a significant effect on (log) reaction time (RT), with higher reaction times observed under haloperidol ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.580 (0.147) seconds(s)) when compared with placebo ($\bar{x}$ ($\sigma_{\bar{x}}$) = 1.242 (0.150), $p = 0.002$, $\eta^2 = 0.292$). We therefore investigated whether there was a relationship between ΔRT and $\Delta\alpha$ under haloperidol. A median split (ΔRT) resulted in two subgroups of participants. Separate RM-ANOVAs, with (square root) learning rate estimates (α) as the dependent variable, and information, volatility and task group as the predictor variables were carried out for each subgroup. For the subgroup of participants who

showed the greatest increase in RT (slowing of response) under haloperidol (n=15), the drug by information interaction no longer reached significance (F (1,13) = 0.106, p = 0.750, η_p^2 = 0.008). The opposite pattern of results was observed for the subgroup of participants (n =16) with a ΔRT below the median change (a reduced slowing of response under haloperidol): here a significant drug by information interaction effect was observed (F (1,14) = 10.846, p = 0.005, η_p^2 = 0.437). Results show that, for the subgroup of participants who showed the greatest slowing of response (ΔRT), haloperidol did not significantly affect learning rates. Given that response slowing has been linked to tonic dopamine this pattern of results further reinforces the idea that our observed effects on α_{primary} are likely mediated by effects of haloperidol on phasic, not tonic, DA.

v. Effect of randomisation schedule and drug day on model parameters

Randomisation schedule (1-4) and drug day (i.e., haloperidol administered on testing day 1 or 2) were included as predictor variables in all analyses (with both n = 31 and n = 41 samples), with no main/interaction effect(s) observed (all F< 1, all p > 0.05). Additionally, testing session was used to check for the presence of practice effects. Testing session (session 1 or 2) was included as a predictor variable in all analysis, with no main/interaction effect(s) observed (all F< 1, all p > 0.05).

vi. Effects of baseline verbal working memory (VWM) on model parameters

As there is evidence to suggest that effects of dopamine manipulation are dependent on baseline DA synthesis, with working memory capacity shown to predict dopamine synthesis in healthy adults [442], we stratified participants into high and low verbal working memory (VWM) groups, based on mean baseline (under placebo) accuracy scores on a verbal working memory task [400]. VWM (high/low) was included as a predictor in a mixed model analysis (n = 31). A Type III RM-ANOVA conducted on model estimates revealed a significant interaction between VWM and information type (F(1,189) = 5.932, p = 0.016, beta estimate (SE) = 0.026 (0.010), t = 2.436, CI [0.00 – 0.05]) with planned contrasts revealing that, for low VWM participants, $\alpha_{secondary}$ values ($\bar{x}(\sigma_{\bar{x}})$ = 0.364 (0.031) were significantly lower than $\alpha_{primary}$ values ($\bar{x}(\sigma_{\bar{x}})$ = 0.447 (0.031); z(30) = 2.820, p_{holm} = 0.010). There was no

significant difference between $\alpha_{primary}$ and $\alpha_{secondary}$ for high VWM participants (z(30) = -0.641, p_{holm} = 0.522). No other main or interaction effects of VWM on α values were observed (all F < 0.01, all p > 0.05). Additionally, the pattern of results was unchanged from the previous analysis excluding VWM, with the drug by information interaction effect remaining significant (F (1,189) = 3.967, p = 0.048, beta estimate (SE) = 0.021 (0.010), t = 1.992, CI [0.00 – 0.04]). Finally, while including baseline VWM as continuous predictor variable in a RM-ANOVA, no main or interaction effect(s) of VWM on α values were observed. Additionally, neither gender, age nor BMI interacted with any outcome variables (all F < 0.01, all p > 0.05). Results suggest that the observed decrease in $\alpha_{primary}$ under haloperidol is not related to variation in working memory capacity.

Appendix 5.5

Instruction scripts

Individual-primary group

Welcome. You have a choice: either choose the blue shape or the green shape. One shape is correct – guessing which one it is will give you points. To help you to choose, one of the shapes is filled with red. This indicates the most popular choice selected by a group of 4 people who previously played this task. When the question mark appears, try picking a shape by pressing the left or right keyboard buttons. [Participant responds]

Feedback: After you make a choice, a tick or cross will appear in the middle. This tells you if the group of previous players were correct or incorrect.

Here they think the blue shape (filled with red) will be correct. Try picking a shape now. [Participant responds]

Blue is correct! This means that this time the others got it right.

Things happen in phases in this game. The game could be in a phase where the blue shape is more likely to be correct. Have another go. [Participant responds]

And blue again! It certainly looks as though you are in a blue phase but make sure you pay attention to what the right answers are because the phase that you are in can change at any time. Here's a tip – ignore which side of the screen the shapes are on – it's the colour that is important! [Participant responds]

The others got it right again. It looks like, right now, you could be in a phase where the group's information is useful. Perhaps these are trials from the end of their experiment, when they had developed a pretty good idea of what was going on. Be careful though because we have mixed up the order of the other people's trials so that their choices will also follow phases. Try again. Perhaps the other shape is right this time? [Participant responds]

Green! This time the green shape was right! The chance of each shape being right or wrong will change as you play, so pay attention! The group were incorrect this time. Remember that sometimes you will see less useful information from the group – for example from the beginning of their experiment where they didn't have a very good idea of what was going on. Have another go... [Participant responds]

This time the green shape was right! The chance of each shape being right or wrong will change as you play, so pay attention. The group were correct too. It looks like, right now, you could be in a phase where the group's information is useful. Try to be as accurate as possible. Getting it right, gives you points. Get enough points and you could earn a silver or even a gold prize! Have another go... [Participant responds]

Things happen in phases in this game. Remember, the tick or cross in the middle tells you if the group were correct or incorrect. That means that the shape with the red box was the correct choice. Have another go... [Participant responds]

The group were correct this time. The tick in the middle tells you that they picked the correct choice. There will now be a short quiz. Pick one more shape and then we'll head to the real game! [Participant responds]

Social-primary group

Welcome. You have a choice between going with, or against advice from a group. Below you can see a blue and green frame, one frame is filled with a red box: this indicates the most popular choice selected by a group of 4 people who previously played this task. One frame is correct. You can pick the same frame as the group have picked or choose to go against the group's advice. When the question mark appears, make your selection by pressing the left or right keyboard buttons. [Participant responds]

Feedback: After you make a choice, a tick or cross will appear in the middle. This tells you if the group of previous players were correct or incorrect.

This time they were correct! This means that the frame filled with the red square was the correct frame.

Here they think the blue frame (filled with red) will be correct. Try picking a frame now. [Participant responds]

The group were correct! This means that this time the others got it right and picked the correct colour.

Things happen in phases in this game. The game could be in a phase where the group are more likely to be correct. Have another go. [Participant responds]

The group were correct again! The blue frame was right again. It certainly looks as though you are in a phase where the group are correct but make sure you pay attention to the feedback because the phase that you are in can change at any time. Blue and green can also

go through phases: it looks like you might be in a phase where the blue frame is more likely to be correct. Try again. [Participant responds]

The others got it right again. It looks like, right now, you could be in a phase where the group's information is pretty useful. Perhaps these are trials from the end of their experiment, when they had developed a pretty good idea of what was going on. Be careful though because we have mixed up the order of the other people's trials so that their choices will follow phases. Try again. [Participant responds]

The group were incorrect this time. This time the green frame was correct. The chance of each frame being right or wrong will change as you play, so pay attention! Remember that sometimes you will see less useful information from the group – for example from the beginning of their experiment where they didn't have a very good idea of what was going on. Have another go... [Participant responds]

The group were correct this time. The chance of each frame being right or wrong will change as you play, so pay attention. Try to be as accurate as possible. Getting it right, gives you points. Get enough points and you could earn a silver or even a gold prize! Have another go... [Participant responds]

Things happen in phases in this game. Remember, the tick or cross in the middle tells you if the group were correct or incorrect. That means that the frame filled with the red was the correct choice. Have another go... [Participant responds]

The group were correct this time. The tick in the middle tells you that they picked the correct choice. There will now be a short quiz. Pick one more time and then we'll head to the real game! [Participant responds]

Feedback Questionnaire

Participants competed a short feedback questionnaire after the behavioural task. 100% of participants said that they understood the task instructions and what they were supposed to do. Participants were then asked to rate on a 5-point Likert scale how often they i) used the group's suggestions (red shape) to help make their decision, comprising the social rating score, and ii) if they paid attention to the colour of the shape (blue/green) that was correct when making their decision (the individual rating score). Social and individual ratings were submitted to separate one-sample t-tests, to ensure that participants in both the individual-primary and social-primary groups were paying attention to both sources of information. Both social ($t(42) = 30.765$, $p < 0.001$) and individual ratings ($t(42) = 29.565$, $p < 0.001$) were significantly greater than zero.