Adaptive Learning Under Uncertainty and Its Relation to Symptoms of Mood Disorders

Fei Shang

St Anne's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity term, 2025

Supervised by Prof Michael Browning and Dr Susannah Murphy

Abstract

A key mechanism that may contribute to symptoms of anxiety and depression is how individuals learn under uncertainty. Understanding this process could inform the development of more effective treatments. This thesis focuses on two common types of uncertainty encountered in everyday decision-making: volatility, which involves systematic changes in the environment over time, and noise, which refers to random fluctuations in outcomes. These types of uncertainty demand opposite learning strategies: volatility requires rapid updating, while noise calls for slower learning to avoid overfitting. In addition, the thesis investigates momentum, a form of structured volatility characterized by gradual, directional change. While prior research suggests that individuals with higher anxiety symptoms show impaired adaptation to volatility, the underlying mechanisms remain unclear. One possibility is that they fail to detect volatility; another is that they confuse it with noise, leading to maladaptive learning.

To address these questions, this thesis presents a series of three studies. In the first study, we focused on developing a rigorous experimental and analytical framework. Specifically, we designed and validated a task in which both forms of uncertainty were independently manipulated and established a set of criteria for selecting outcome schedules. To characterize how learning adapted to these uncertainties, we compared model-free and model-based approaches for estimating learning rates. In the second study, we showed that human learners increased their learning rate with high volatility and reduced their learning rate with high noise. Individuals reporting greater anticipatory pleasure exhibited less adaptation to noise. Pupillometry revealed that higher anxiety symptoms correlated with reduced pupil response to volatility, with tonic pupil area tracking volatility only in low-anxiety participants. In the third study, we observed that momentum increased phasic

pupil dilation, while noise increased tonic pupil diameter. Behaviorally, participants showed higher learning rates for prediction errors in the same direction as the momentum, demonstrating adaptation. We extended the Rescorla-Wagner model to account for momentum effects, finding significantly higher model-estimated momentum when momentum was present.

These findings demonstrate how human learners adapt both behaviorally and physiologically to different types of uncertainty, with individual differences in mood disorder symptoms associated with variations in these adaptations. Future research should establish causal links between maladaptive learning and mood disorders to inform more effective interventions.

Acknowledgements

In the past few years at Oxford, I have been fortunate to meet so many wonderful people who have shaped both my work and my life.

First and foremost, I must thank my supervisor, Mike. He patiently listened to many of my ideas and examined more plots than I care to admit. He was Yoda to my Luke Skywalker, not just teaching me how to do good science but also reminding me to ask the right questions and to care about evidence that truly makes a difference. Susie also gave me valuable advice along the way and helped me greatly at every milestone, from settling into my DPhil to finding a postdoc. Together, they've been role models for what it means to be great scientists.

I've been lucky to work alongside brilliant colleagues in Mike's lab. Maggie, Verena, and Juliet—thank you for our countless discussions, your career advice, and for sitting through my many practice talks. Mary—my DPhil wouldn't be the same without your constant encouragement. Alex, Lisa, Victoria, and Kate—thank you for patiently answering all my questions about clinical trials. And thank you to Paula for brightening my days with cute cat videos. And to all of you, thank you for being such brilliant bakers and sharing your delicious creations!

Oxford has been a wonderfully collaborative and intellectually stimulating place. I was lucky to work with Michael on a pharmacological study, aand I'm grateful to be collaborating with Georgia, Grace, and Amy on a project about premenstrual syndromes. Very few people get the chance to explore so broadly and with such autonomy beyond their core DPhil topic. I consider myself truly lucky.

I love science because I love life, and my life at Oxford was deeply enriched by my friends. Thank you to Georgia, Zhen, Lisa, Charlotte, Alice, Blanca, and Merve for the welfare dinners, mental health walks, and countless coffee chats that kept me grounded and joyful.

And at the heart of it all, Edi has been my constant. I thank him for his intellectual clarity, emotional support, and unfailing humor. You're one of the rare people who actually know what their partner's doctoral thesis is about, and you're a mathematician!

Finally, to my parents: thank you for giving me a happy home. No matter how challenging things became, I could always find comfort and confidence in you. When it felt like every soul on Earth was busy building the next Facebook, you supported me in becoming a scientist and studying the human mind. Thank you for believing in me, and for being my best friends.

Contents

List of Figures

List of Tables

1 Introduction

Uncertainty plays a crucial role in shaping how individuals learn and adapt to new information. It has significant implications for both cognitive flexibility and emotional regulation. In the context of this thesis, uncertainty is defined as any condition in which the future state of the environment is not fully predictable. The ability to adaptively update one's learning rates in response to different forms of uncertainty has been shown to be critical for optimal decision-making and behavior. Anxiety disorders, characterized by an overestimation of threat or risk, are associated with an impaired ability to tolerate uncertainty, making this topic highly relevant to psychiatric research.

This thesis will focus specifically on associative learning, the process by which individuals form connections between stimuli, responses, and outcomes (Balleine & Dickinson, 1998). Associative learning is particularly suited to understanding how organisms adapt to unpredictable environments because it highlights the mechanisms underlying behavioral adjustments. Other forms of learning, such as observational, fall outside the scope of this work (Burke et al., 2010).

In this Chapter, I will first provide an overview of the different types of uncertainty that are relevant to cognitive processes and learning, including expected, unexpected, and estimation uncertainty, influence associative learning. Understanding these dynamics is pivotal for investigating how uncertainty affects learning

and, in turn, decision-making. Following this, I will explore the neuromodulatory mechanisms that facilitate learning under uncertain conditions, with a specific focus on the locus coeruleus (LC) and its involvement in regulating arousal and attention through norepinephrine (NE) release. Pupillometry, a non-invasive technique to measure pupil dilation, will be discussed as a proxy for LC activity, given recent studies linking fluctuations in pupil size to changes in environmental volatility. Furthermore, I will examine computational models such as reinforcement learning and Bayesian learning, which have been used to understand how humans and animals adjust their behavior in uncertain environments. This is particularly important as the current thesis employs a computational approach to investigate how symptoms of mood disorders such as anxiety and depression, affect learning under uncertainty. Lastly, I will explore the relationship between learning under uncertainty, learned helplessness, and intolerance of uncertainty (IU), both key concepts in understanding mood and anxiety disorders.

1.1 Different forms of Uncertainty

There are many different forms of uncertainties (Payzan-LeNestour & Bossaerts, 2011; Pulcu & Browning, 2019; Soltani & Izquierdo, 2019). In this chapter, we will discuss expected uncertainty, unexpected uncertainty, estimation uncertainty, higher order uncertainty and momentum, a linearly structured unexpected uncertainty.

1.1.1 Unexpected uncertainty (volatility)

Unexpected uncertainty, often referred to as volatility or second-order uncertainty, arises when there is a meaningful change in the underlying contingencies of the environment. In these situations, previously learned associations no longer hold, necessitating a rapid adjustment in behaviour. For example, an individual may have learned that a specific cue predicts a particular outcome; however, when the environment becomes volatile, this predictive relationship changes. If individuals fail to update their internal model of the world by integrating new and recent observations, they risk making decision based on an outdated representation

of the environment, which could lead to reduced chances of achieving desired outcomes. In contrast, adapting to volatility by increasing learning rates allows individuals to remain flexible and responsive to changing circumstances. Empirical evidence suggests that people do indeed exhibit increased learning rates in volatile environments, demonstrating their ability to adjust behaviours to align with new contingencies (Behrens et al., 2007; Nassar et al., 2012).

The locus coeruleus-norepinephrine (LC-NE) system has been implicated in reflecting the level of volatility. Research suggests that the LC signals uncertainty by modulating norepinephrine release, which in turn influences arousal and attention (Aston-Jones & Cohen, 2005). Pupillometry studies have provided compelling evidence for this connection, showing that pupil dilation correlates with volatility in the environment. In fact, Nassar et al. (2012) demonstrated that changes in pupil size reflect adjustments in learning rates in response to volatility, supporting the idea that the LC plays a crucial role in responding to unexpected uncertainty. A detailed review on LC-NE system can be found under Neurotransmitters signalling volatility, noise, and estimation uncertainty.

1.1.2 Expected Uncertainty (noise)

In contrast to volatility, expected uncertainty, also known as first-order uncertainty or noise, arises from the inherent probabilistic nature of certain environments. In these cases, individuals are aware that the environment contains noise or randomness, but the underlying contingencies remain stable. For example, even though one knows that flipping a fair coin results in heads or tails with equal probability, the outcome of any single flip is uncertain. In such noisy environments, it is adaptive to reduce the learning rate and rely more on long-term averages rather than over-adjusting to each individual outcome. Attempting to update mental representations of the environment based on noisy observations can lead to inaccurate estimations of the underlying contingencies. This is because the randomness inherent to the environment reflects variability as opposed to meaningful changes.

Acetylcholine (ACh) is thought to signal expected uncertainty or noise in environments where outcomes are probabilistic but follow a stable statistical pattern. Theoretical models by Dayan and Yu (2002) suggest that ACh helps allocate attention to cues that are noisy or invalid. Empirical studies support this role, showing that ACh facilitates faster attentional shifts in noisy environments and improves response times to invalid cues while also modulating attention when cue-outcome relationships change (Bucci et al., 1998; F. Murphy & Klein, 1998; Phillips et al., 2000; Stewart et al., 2001; Vossel et al., 2008; Witte et al., 1997). These studies will be discussed in more detail in Neurotransmitters signalling volatility, noise and estimation uncertainty.

1.1.3 Estimation uncertainty

Estimation uncertainty, or epistemic uncertainty, arises from an individual's limited knowledge about a particular environment. Unlike noise, which is an inherent characteristic of the environment, estimation uncertainty can be reduced through learning and the accumulation of additional data. This form of uncertainty is highly relevant in cases where individuals must infer the underlying structure of the environment based on incomplete or ambiguous information. Bayesian learning models have been particularly effective in capturing how individuals update their beliefs in the face of estimation uncertainty, as these models explicitly account for the trade-off between exploration (gathering more information) and exploitation (making decisions based on current knowledge) (Mathys et al., 2014).

1.1.4 Higher order uncertainty

Higher order uncertainty refers to meta level uncertainty—uncertainty about how volatility and noise in the environment themselves change over time. It goes beyond simply estimating the current level of volatility or noise, focusing instead on the underlying dynamics that govern these properties. For example, an individual may initially perceive the environment as stable and predictable, but as outcomes become more erratic, they may infer that volatility is increasing or fluctuating.

This introduces uncertainty not just about immediate outcomes, but also about the rules or patterns driving environmental change.

Such uncertainty complicates learning and decision making because it affects expectations about how quickly beliefs and the learning rate itself should adapt. In environments where uncertainty is dynamic, individuals must revise not only their beliefs about the present state but also their assumptions about the pace and nature of future change. This is especially relevant in cases of higher order structure, such as momentum, where changes in volatility follow a consistent trend or directional drift over time. This kind of linearly structured volatility suggests that recent increases or decreases in volatility may predict future changes in the same direction, offering a pattern that adaptive learners can potentially detect and exploit. I will expand on momentum and its implications in the next section.

1.1.5 Momentum – Linearly structured volatility

Volatility is often described as a condition in which the environment is changing in a way that disrupts previously learned associations. Some volatility exhibits an inherent structure to the way it changes. For example, seasonal changes follow a cyclical pattern, where the average temperature varies predictably throughout the year. Such structured changes offer opportunities for learners to identify patterns and adapt accordingly.

One unexplored form of structured volatility is momentum, which refers to a persistent, directional trend in environmental conditions over time. Unlike general volatility, where changes may lack a discernible pattern, momentum is characterized by a linear or progressive shifts in the variable being learned. Understanding how learners detect and adapt to momentum is crucial because it allows them to accelerate learning by anticipating where the environment is headed. This, in turn, leads to more accurate mental representations of the environment and better decision-making. There is limited evidence on how human learners adapt their learning strategies when there is momentum in the environment. Ritz et al. (2018) demonstrated that participants incorporated historical information when learning

under both momentum and a change-point schedule. However, their findings failed to reveal distinct behaviours specific to learning under momentum.

1.2 Neurotransmitters signalling volatility, noise and estimation uncertainty

A body of theoretical work by Dayan and Yu (2002, 2005, 2006) and Yu and Dayan (2003) proposed a dual-system model in which different forms of uncertainty are processed by specific neurotransmitters. NE is suggested to signal volatility, while ACh signals noise or expected uncertainty. This model highlights how these neurotransmitters dynamically adjust attentional resources in response to different types of environmental uncertainty.

Specifically, NE is theorized to encode a drop in confidence regarding the stability of cue-outcome relationships, increasing when an unexpected change (volatility) is detected in the environment. This surge in NE reflects the need for rapid learning or adaptation to a new environmental state. In contrast, ACh is responsible for reporting the validity of known cue-outcome pairings under conditions of noise, modulating attention allocation based on how reliable these pairings remain over time. Under noisy conditions, where variability does not reflect meaningful changes, ACh is thought to promote a reduction in learning rates, helping to preventing over-adjustment. Both neurotransmitters, therefore, are thought to be critical in adjusting learning rates and attention: NE facilitates higher learning rates in volatile environments to adapt to new environment statistics while ACh reduces learning rates in noisy environments to maintain focus on learned cues.

Though this thesis does not involve pharmacological manipulation of NE or ACh, it is important to review the empirical evidence because these neurotransmitters play pivotal roles in the arousal system, particularly through pathways like the locus coeruleus (LC), which we investigated using pupillometry.

1.2.1 Norepinephrine

The role of norepinephrine (NE) in signalling volatility is well-studied (Sara, 2009; Sara & Segal, 1991). In environments where sudden changes in contingencies occur, NE plays a central role in enhancing alertness and attention, allowing individuals to quickly adapt to unexpected changes. One early study demonstrated that injecting rats with idazoxan, a drug that increases NE release, enhanced the detection of unexpected switches in a visual maze task (Devauges & Sara, 1990). This finding provided early experimental support for NE's involvement in rapidly adjusting behavior in response to changing environmental cues.

Similarly, recent pharmacological interventions in rodents have reinforced the idea that NE is crucial for cognitive flexibility. For instance, the selective norepinephrine reuptake inhibitor (SNRI) atomoxetine, which boosts NE levels, has been shown to improve attentional set-shifting in rats with noradrenergic lesions. Interestingly, this effect was not observed in non-lesioned rats, suggesting that the role of NE maybe particularly pronounced when baseline noradrenergic function is impaired (Jahn et al., 2020). These findings build on earlier work by Newman et al. (2008), who also found that enhancing NE transmission through administration of atomoxetine improved flexibility in shifting attention between changing task rules in noradrenergically lesioned rats.

In humans, the role of NE in learning under uncertainty has been explored using both pharmacological manipulation and neuroimaging. Cook et al. (2019) found that enhancing NE and dopamine activity using methylphenidate (a blocker of NE and dopamine transporters) led to faster learning rates in volatile environments compared to stable environments. Similarly, Jepma et al. (2018) found that atomoxetine (a selective norepinephrine reuptake inhibitor) increased learning rates in participants with low baseline learning rates (defined as learning rate when on placebo), but decreased learning rates in participants with high baseline learning rates. Conversely, reducing NE activity impairs learning in volatile contexts. Lawson et al. (2017) found administering Propranolol (a β-adrenergic receptor antagonist) led to a reduction in learning rate on a sensory probabilistic associative learning task.

Moreover, neuroimaging studies have provided further evidence of NE's role in volatility signalling. Payzan-LeNestour et al. (2013) used blood oxygenation level dependent (BOLD) activity in the locus coeruleus (LC) to demonstrate that this brain region tracks volatility, with increased activation corresponding to changes in environmental states. However, it is important to note that BOLD signal is an indirect measure of neural activity and does not provide direct evidence of neurochemical processes.

A study by Marshall et al. (2016) highlighted the complexity of NE's role in processing uncertainty. They examined the effects of NE, ACh, and dopamine (DA) on learning under uncertainty and found that blocking NE with prazosin (an α_1-receptor antagonist) increased participants' estimates of volatility. At first glance, this finding appears to contrast with results from Cook et al. (2019), Lawson et al. (2017), and Payzan-LeNestour et al. (2013), which linked increased NE-related activity with greater sensitivity to environmental volatility. One potential explanation for this discrepancy is that NE regulates the rate at which individuals adjust their beliefs: when NE activity is reduced, individuals may become overly reactive to unexpected events, leading to faster but potentially less regulated adjustments in their learning processes. Alternatively, the effect of prazosin could also depend on participants' baseline arousal or learning rates (as suggested by Jepma et al. (2018)).

A recent review by Nassar (2024) argues that NE does not merely signal environmental changes but plays a more pivotal role in determining when the brain should update its internal models in response to latent state transitions. Latent states refer to hidden, underlying contexts or conditions that influence behaviour and learning but must be inferred from observable data (Cochran & Cisler, 2019; Gershman & Niv, 2012). For example, a latent state could represent whether the environment is stable or volatile, which affects how individuals adapt their predictions and actions (Ashwood et al., 2022; Zika et al., 2023). The review provides several key pieces of evidence supporting this role of NE in cognitive adaptation. First, humans demonstrate an ability to retain learned information

even when faced with changing environmental conditions, suggesting that NE does not simply trigger forgetting of previous knowledge. Instead, it helps signal when existing mental models should be adjusted. Second, the human brain is capable of discounting oddball situations (i.e., unexpected stimuli that are irrelevant to the larger context), and thus preventing overreaction to random, inconsequential changes in the environment. Lastly, while NE is widely believed to modulate brain responses to environmental shifts, the evidence regarding how environmental changes affect phasic NE activity remains inconsistent.

Empirical studies support the role of NE in signaling unexpected uncertainty, as proposed by Dayan and Yu (2002). However, emerging evidence suggests that the function of NE is more nuanced than initially proposed. Some studies suggest that the effects of NE may depend on noradrenergic lesions (Jahn et al., 2020; Newman et al., 2008) or baseline learning rates (e.g., (Jepma et al., 2018)). Nassar's view on NE signaling latent state transitions is an intriguing idea that could potentially explain many findings, but it lacks robust empirical evidence. Neither Dayan and Yu's framework nor Nassar's hypothesis fully accounts for the observation that increasing NE transmission improved flexibility in shifting attention only in noradrenergically lesioned rats, with effects varying by baseline learning rates. Furthermore, these theories overlook the potential contributions of tonic NE activity, a critical function of NE implicated in learning and attention (Aston-Jones & Cohen, 2005).

1.2.2 Acetylcholine

In contrast to NE, acetylcholine (ACh) is thought to signal expected uncertainty or noise. This type of uncertainty involves environments where outcomes are not fully predictable but follow a known statistical distribution. Theoretical models by Yu and Dayan (2003) postulate that ACh plays a key role in determining how much attention should be allocated to cues that are known to be noisy or invalid.

Empirical studies have supported this model. For instance, manipulations of ACh levels using spatial cueing paradigms, where subjects respond to visual cues with varying levels of validity, have demonstrated the importance of ACh in

modulating attention. In such paradigms, reaction times (RT) to valid cues (those that accurately predict outcomes) are typically shorter, while invalid cues (those that mislead) result in longer RTs as attention needs to be shifted back to the correct target. Enhancing ACh transmission with nicotine, an agonist of nicotinic cholinergic receptors, has been shown to reduce RTs, particularly for invalid cues, indicating that nicotinic receptor activation facilitates faster attentional shifts in noisy environments (F. Murphy & Klein, 1998; Stewart et al., 2001; Witte et al., 1997). However, Vossel et al. (2008) found that nicotine reduced reaction times only under conditions of high cue validity (90%) but not under lower validity (60%) during a probabilistic spatial cueing task. This finding implies that nicotinic receptors may modulate attention more effectively in less noisey environment, where cues are relatively reliable. Conversely, in noisier environments, where greater integration of probabilistic information is required, the benefits of nicotinic activation appear to diminish.

Conversely, reducing ACh levels through administration of scopolamine, an antagonist of muscarinic cholinergic receptors, increases RTs, as shown in studies where rats were slower to react to invalidly cued targets (Phillips et al., 2000). Similar effects have been observed in other paradigms; rats with cholinergic lesions to the posterior parietal cortex failed to increase attention when conditioned relationships between cues and outcomes changed (Bucci et al., 1998).

While these empirical findings align with Yu and Dayan (2005)'s theoretical model, they also highlight complexities in ACh's role in signalling expected uncertainty. Evidence that nicotinic effects are modulated by environmental noise levels, suggests that ACh's role in allocating attention is context-dependent.

1.2.3 Interaction Between Norepinephrine and Acetylcholine

An intriguing aspect of the Yu and Dayan model is the proposed interaction between NE and ACh. The authors suggest that ACh modulates or "gates" the effectiveness of NE in signalling volatility. In this context, "NE" refers not just to the level of norepinephrine but to its functional response in signalling changes in environmental

contingencies. According to their model, ACh levels can regulate how strongly NE responds to environmental shifts, with higher ACh levels attenuating NE's ability to signal volatility. This relationship is formalized in their equation:

$$NE \propto \frac{ACh}{0.5 + ACh} \tag{1.2.3.1}$$

This equation implies that the functional response of NE is influenced by the level of ACh, with the denominator 0.5+ACh acting to dampen the impact of increasing ACh levels. Specifically, as ACh levels rise, the output of NE diminishes, reflecting the gating mechanism where ACh modulates the sensitivity of NE to changes in environmental volatility.

Thus, as ACh increases, the ability of NE to signal volatility decreases, which could lead to a delayed or less frequent detection of changes in the environment. Although empirical evidence for this interaction is limited, the proposed relationship highlights the complexity of how neurotransmitter systems cooperate in the brain to manage different forms of uncertainty.

1.3 Pupil diameter as an indirect signal of locus coeruleus activity

The locus coeruleus (LC), a small nucleus located in the brainstem, is the principal site of NE production in the brain and plays a critical role in modulating arousal, attention, and adaptive behavior under varying environmental conditions. Rajkowski et al. (1994) provided early evidence for this by recording neuronal activity in the LC of monkeys performing goal-directed tasks. Their findings revealed that LC activity increased not only during task engagement but also in response to unexpected stimuli, such as an infrequent visual cue in a visual oddball discrimination task or unconditioned sensory stimuli (e.g., the sound of the experimenter entering the lab). This led them to conclude that the LC might regulate both tonic levels of vigilance, which maintain a general state of alertness, and phasic responses that drive heightened attention in response to specific, salient stimuli.

However, these early recordings did not specifically differentiate between NE-releasing neurons and other types of neurons within the LC, leaving open the possibility that non-NE neurons could have been contributing to the observed effects. A more recent study by Varazzani et al. (2015) provided more direct evidence by showing that pupil dilation, a physiological response often used as an indicator of arousal, correlated specifically with the firing rates of NE neurons in the LC. In this study, monkeys engaged in a reward-effort trade-off task, and the researchers observed that pupil dilation increased in parallel with the heightened activity of NA neurons. While this study did not directly explore how changes in environmental conditions influenced this relationship, it nonetheless supports the hypothesis that increased activity in LC-NE neurons leads to pupil dilation, further establishing pupil diameter as a useful proxy for LC activity.

In humans, similar findings have emerged. P. R. Murphy et al. (2014) observed that pupil diameter covaried with BOLD activity in the LC region, both at rest and during an oddball task. While BOLD signals do not directly measure NE release, the correlation with pupil diameter supports the idea that pupil size reflects underlying changes in noradrenergic activity. By linking pupil size to BOLD signals in the LC, these studies extend the idea that pupil diameter is a reliable indicator of LC function, not only during active tasks but also in baseline arousal states.

Additionally, Reimer et al. (2016) demonstrated that changes in pupil size are closely tied to fluctuations in cortical state during wakefulness, further suggesting that pupil dynamics are a reflection of broader neuromodulatory changes. Using two-photon microscopy to monitor the activity of noradrenergic and cholinergic projections in mice, the study revealed that rapid pupil dilations were temporally aligned with bursts of NE activity, whereas sustained dilations correlated with cholinergic activity during behaviours like locomotion. This research provides a more nuanced understanding of how neuromodulatory systems, including both norepinephrine and acetylcholine, influence pupil size. The findings suggest that pupil size may serve as a non-invasive biomarker for tracking the dynamic interaction between these neuromodulatory systems across different behavioral states.

Building on this, Joshi et al. (2016) offered compelling evidence for the link between LC-NE neuron activity and pupil dilation by conducting simultaneous recordings of pupil diameter and neural activity in the LC, inferior colliculus, superior colliculus, and cortical areas such as the anterior and posterior cingulate cortices in alert, fixating rhesus monkeys. The study utilized a variety of tasks, including passive fixation and auditory oddball paradigms, where an unpredictable stimulus was played at irregular intervals. Across these tasks, fluctuations in LC neuron firing rates were consistently associated with changes in pupil size. Notably, macrostimulation of the LC, but not other regions like the colliculi, reliably evoked pupil dilation, providing strong evidence that the LC is a key driver of these changes. This work not only highlighted the role of the LC in regulating arousal and attention through pupil dynamics but also emphasized the specificity of the LC in modulating pupil diameter compared to other neural regions.

Together, these studies underscore the utility of pupil diameter as a robust, non-invasive indicator of LC activity. This relationship between pupil size and LC activity has significant implications for understanding how the brain orchestrates arousal and adaptive responses to environmental uncertainty, particularly through the modulation of tonic and phasic noradrenergic systems. Given the role of the LC-NE system in regulating attention and learning under uncertainty, pupil dilation serves as an important physiological marker for studying the interplay between neuromodulatory activity and cognitive processes in both animal models and humans. However, given the involvement of other neuromodulators, such as ACh (Reimer et al., 2016) and serotonin (Cazettes et al., 2021), in pupil regulation, it is important to interpret pupil dilation as being consistent with increased LC activity rather than as direct evidence of it.

1.4 Pupil dilation tracks volatility

Several studies have demonstrated that pupil diameter not only reflects levels of arousal but also provides insight into how individuals perceive and adapt to environmental volatility. Urai et al. (2017), using a random dot coherence

discrimination task, found that pupil dilation occurring after a decision but before receiving feedback was indicative of decision uncertainty. This suggests that the degree of post-choice pupil dilation may reflect the brain's assessment of the confidence or uncertainty regarding the decision made, even in the absence of immediate feedback.

In another study, Nassar et al. (2012) employed a predictive-inference task in which participants were asked to predict numbers drawn from a Gaussian distribution. The authors manipulated volatility using a changepoint design, where the mean of the distribution would unpredictably shift at various intervals, simulating an unstable environment. Their findings indicated that pupil dilation increased more significantly during moments of high surprise but low noise, suggesting that changes in pupil diameter may be sensitive to perceived environmental instabilities, specifically during unexpected shifts in context (i.e., changepoints). This supports the notion that pupil dilation tracks an individual's sensitivity to volatility, aligning with the idea that noradrenergic activity in the LC mediates these responses to sudden environmental shifts.

Browning et al. (2015) further explored this dynamic by examining how pupil diameter fluctuated during an aversive probabilistic reversal learning task, where participants had to adapt their behavior in response to changing contingencies under conditions of both high and low volatility. The study revealed that pupil dilation was modulated by both surprise (unexpected outcomes) and volatility (rate of environmental change). Interestingly, individuals with high trait anxiety exhibited a diminished ability to adjust their learning rates in response to volatility, which corresponded with altered pupillary responses. This finding suggests that individuals with anxiety disorders, where the LC-NE system is often dysregulated (Sullivan et al., 1999), may show impaired sensitivity to volatility, potentially due to a maladaptive response to environmental uncertainty. This underscores the role of noradrenergic neurotransmission in modulating pupil dynamics, especially in the context of emotional and cognitive responses to unpredictable environments.

Taken together, these findings indicate that pupil dilation, modulated by LC-NE activity, reflects an individual's sensitivity to environmental volatility and may serve as a reliable, non-invasive measure for assessing how the brain computes and responds to changes in environmental stability.

1.5 Pupil dilation may track noise

Unlike the well-established link between pupil dilation and volatility, the relationship between pupil diameter and noise (or irreducible uncertainty) is more ambiguous and has generated mixed findings. de Berker et al. (2016) explored how subjective estimates of uncertainty influence both psychological and physiological stress responses. In the study, participants predicted the presence of snakes under rocks, receiving mild shocks when correct. The probability of a snake being under a given rock varied throughout the experiment. Using a hierarchical Bayesian learning model, they dissociated noise from volatility in a probabilistic learning task. The researchers found that pupil dilation dynamically tracked levels of irreducible uncertainty during the task. Specifically, as irreducible uncertainty increased, participants exhibited greater pupil dilation. This suggests that pupil dilation may serve as a physiological marker of stress and uncertainty in environments where outcomes are inherently unpredictable or noisy.

Similarly, Preuschoff et al. (2011) investigated how different factors, such as expected reward, noise (which they referred to as risk), and surprise (or risk prediction error), affected pupil diameter during an auditory gambling task. In their study, participants bet on whether the first or second card drawn from a deck numbered 1–10 would have a higher value. The first card functioned as a probabilistic cue, and the second card revealed the outcome. Irreducible uncertainty (or noise) was highest when the first card had a middle value (5 or 6), creating greater ambiguity about the outcome, and was minimal with extreme values (1 or 10), where the outcome was more predictable. Pupil dilation was more pronounced when participants faced sure outcomes (i.e., when noise was low and the first card was 1 or 10), while higher uncertainty (cards 4, 5, 6, or 7) led to smaller

pupil dilations. Trials with medium uncertainty (cards 2, 3, 8, and 9) fell between these two extremes. These findings suggest that pupil dilation might vary based on the level of uncertainty in the task, but the exact relationship between pupil responses and noise remains complex.

A recent study by Pajkossy et al. (2023) sought to disentangle the effects of volatility and noise on pupil dilation using a probabilistic reversal learning task. In a probabilistic reversal learning task, participants were asked to guess which of a fictional actor's hands hid a stone. The preferred hand switched with a fixed probability of 0.06 per trial. To manipulate the levels of volatility and noise across trials, the researchers employed three different task schedules (0.75/0.25, 0.85/0.15, and 0.65/0.35). Their findings revealed that both volatility and noise independently led to significant increases in pupil dilation. This suggests that pupil responses reflect a generalized sensitivity to uncertainty, whether that uncertainty stems from volatile changes in the environment or the inherent unpredictability of outcomes (noise). The ability to track both forms of uncertainty through pupil dilation highlights the potential of using pupil size as a non-invasive measure for studying how the brain processes different types of uncertainty during decision-making.

1.6 Computational Models of Learning under Uncertainties

In uncertainty research, four computational models are frequently employed: reinforcement learning (RL) models, Bayesian learning models, Bayesian reinforcement learning models, and hidden state inferences. Each model offers a different approach to how agents learn and adapt under uncertainty, with varying degrees of complexity and computational demands. In this section, I will review these models in greater detail, comparing their strengths, weaknesses, and suitability for different types of tasks.

1.6.1 Reinforcement learning

Reinforcement learning (RL) posits that individuals learn to maximize rewards in an environment primarily by adapting through trial and error. Specifically, people form expectations about outcomes or the likelihood that certain events will occur, and learning happens when these expectations deviate from the actual results. When there is a mismatch between the expected reward from an action or stimulus and the actual outcome, this discrepancy, known as prediction error (PE), drives the learning process. The prediction error is mathematically defined as:

$$PE = \text{Outcome} - \text{Prediction} \tag{1.6.1.1}$$

The extent to which one learns from this prediction error is determined by the learning rate (LR). The learning rate reflects how much weight or importance is given to recent observations in updating future expectations, with higher learning rates indicating that individuals are learning rapidly from new experiences.

A foundational model in RL is the Rescorla-Wagner model, which describes the process of belief updating. In this model, beliefs about the expected value of a stimulus or action (V_t) are updated according to the prediction error:

$$V_{t+1} = V_t + LR \cdot (O_{t+1} - V_t) \tag{1.6.1.2}$$

As previously discussed, when the environment is uncertain, whether due to estimation uncertainty or unexpected uncertainty, new outcomes carry more informational value. In such environments, individuals are more likely to adjust their beliefs and learning rates to adapt to changing conditions. However, when expected uncertainty is high, where variations are caused by random fluctuations, individuals may reduce their learning rates, as randomness provides less information and thus fewer incentives for further adaptation.

There is abundant evidence that reinforcement learning operates at the neural level. For instance, single-neuron recordings have demonstrated that the phasic activity of dopamine (DA) neurons correlates with the signaling of prediction errors, which are central to RL processes (Montague et al., 1996; Sutton, 1988). DA

neurons fire in response to unexpected rewards, and this firing pattern serves as a prediction error signal that guides future behavior. Similarly, electrophysiological and fMRI studies have shown that activity in the striatum is closely associated with prediction errors during reward-based learning tasks (D'Ardenne et al., 2008; Daw & Doya, 2006; Hart et al., 2014; Montague et al., 1996; O'Doherty et al., 2003; Rutledge et al., 2014, 2015; Schultz et al., 1997).

Further research suggests that the computation of learning rates is influenced by activity in the anterior cingulate cortex (ACC). Lesion studies in monkeys show that damage to the ACC impairs their ability to maintain behaviors associated with reward, although it does not affect their immediate ability to correct errors (Kennerley et al., 2006). In experiments where monkeys were exposed to probabilistic environments (e.g., where actions led to rewards 75% of the time), ACC-lesioned monkeys struggled to sustain rewarded behaviors, whereas their performance in deterministic settings (where outcomes were certain) was unaffected. This suggests that the ACC is critical for adapting behavior in uncertain or volatile environments. Neuroimaging studies in humans similarly show that BOLD responses in the ACC correlate with the brain's estimation of volatility during learning tasks (Behrens et al., 2007; Yoshida & Ishii, 2006).

1.6.2 The pros and cons of Reinforcement learning model

The RL model stands out as a computationally simple and efficient approach, grounded in the principles of prediction error in learning. It avoids making assumptions about how individuals solve tasks involving uncertainties or requiring knowledge of the underlying environment distribution, such as a normal distribution. However, it is important to note that traditional reinforcement learning models do not offer an explicit representation of uncertainty. RL models typically update beliefs based on single-point estimates (i.e., the expected value of an action), but they do not incorporate probability distributions to represent uncertainty around those beliefs. This is a significant limitation, as representing beliefs probabilistically, with a distribution that includes both mean and variance, inherently captures the

uncertainty present in the environment (Griffiths & Tenenbaum, 2006; Müller et al., 2019; Pouget et al., 2013). Despite this limitations, RL primary strength lies in its ability to describe the learning rate of individual participants effectively.

1.6.3 Pearce-Hall model

The Pearce-Hall model introduces a key distinction from the Rescorla-Wagner model by emphasizing the dynamic nature of associability. In contrast to the Rescorla-Wagner model, where the learning rate (LR) is fixed, the Pearce-Hall model proposes that the associability (α) of a conditioned stimulus (CS) changes over time based on how surprising or predictable the stimulus is. Associability in this model serves a function similar to a learning rate, but it dynamically adjusts in response to the prediction error—the difference between expected and actual outcomes.

The core concept of the Pearce-Hall model is the idea of 'associability', which quantifies how effective a conditioned stimulus (CS) is at influencing learning at any given moment. The model suggests that the associability of a CS, or its ability to drive learning, decreases as the consequences of the CS become accurately predicted.

Formally, the change in associability ($\Delta\alpha$) at time t is expressed as:

$$\Delta\alpha_t = \lambda - \alpha_t \left|\Delta V_t\right| \tag{1.6.3.1}$$

Where:

- λ is a constant representing the maximum possible change in associability.
- α_t is the associability of the CS at time t, which decreases as the unsigned magnitude of the prediction error ($|\Delta V_t|$) diminishes.
- ΔV_t is the absolute value of the change in associative strength of the CS at time t, which reflects the magnitude of the prediction error.

The associative strength of the CS (V_t) represents how strongly the CS is linked to the US, and it is updated according to:

$$\Delta V_t = \alpha_t \left(R_t - V_{t-1}\right) \tag{1.6.3.2}$$

Here, associability (α_t) influences how much learning occurs on each trial, similar to a learning rate, but it adjusts based on how surprising the CS's outcomes are. When the prediction error is large, associability increases, leading to greater learning, whereas when the prediction error is small, associability decreases, and the CS becomes less impactful for learning.

The critical difference between the Pearce-Hall model and the Rescorla-Wagner model is this adjustment of associability. In the Rescorla-Wagner model, the learning rate is constant, and the magnitude of belief updating depends on the size of the prediction error. This means that larger prediction errors—indicating more surprising outcomes—lead to greater belief updates. However, the model does not explicitly track uncertainty, as it does not adjust the learning rate dynamically based on changes in the environment. In contrast, the Pearce-Hall model takes account of uncertainty by tracking the expected magnitude of prediction errors over time. Associability, a key feature of this model, adjusts based on the predictability of outcomes. When outcomes are highly unexpected, associability increases, prioritizing learning in uncertain environments.

The Pearce-Hall model can be applied to human learning tasks such as Probabilistic Reversal Learning (Li et al., 2011). In these tasks, participants initially experience a stable period where one action yields more favourable outcomes. During the learning phase, participants have high estimation uncertainty, resulting in high associability for both actions. As they gather more data, they learn the relationship between stimuli and rewards, leading to an increase in associative strength (V_t) for the rewarding action and a decrease for the less rewarding one. Over time, associability (α_t) decreases as prediction errors become smaller, reflecting a reduction in uncertainty. However, during the reversal phase—when action-outcome contingencies are suddenly switched—the prediction error increases sharply because participants continue to choose the previously rewarding action, expecting it to still yield rewards. The increased prediction error boosts the associability of both actions, enabling participants to rapidly adjust their choices based on the new reward contingencies.

1.6.4 The pros and cons of Pearce-Hall model

The Pearce-Hall model introduces a dynamic learning rate, or associability, which adjusts based on the magnitude of prediction errors. This allows the model to account of uncertainty by increasing learning rate when outcomes are unexpected . Similar to limitations of RL model, Pearce-Hall model does not explicitly represent uncertainty using a probabilistic framework.

1.6.5 Bayesian Learner

Behrens et al. (2007) proposed a Bayesian learner model to explain how individuals adapt to changing environments by updating their beliefs about reward probabilities and volatility (the rate at which these probabilities change). The model assumes that the observed outcome (y_i) at trial i depends on the current reward rate (r_i), but the reward rate itself is subject to fluctuations due to volatility (v_i).

The change in reward probability is modelled using a beta distribution, which governs how likely the reward is to occur. Formally,

$$\Pr(r_{i+1} \mid r_i, v) \sim \text{Beta}(r_i, V) \tag{1.6.5.1}$$

where:

- r_i is the mean of the distribution.
- $V = \exp(v)$, and the volatility parameter v controls how much r_i is subject to change.

A high volatility parameter (v) means that the reward rate is more likely to change significantly from one trial to the next, requiring the learner to adjust expectations frequently.

The model assumes that the change in volatility is governed by a Gaussian distribution, reflecting how quickly the environment is shifting. The volatility at the next trial (v_{i+1}) is drawn from a distribution based on the current volatility (v_i):

$$\Pr(v_{i+1} \mid k, v_i) \sim \mathcal{N}(v_i, K) \tag{1.6.5.2}$$

where:

- k is a parameter controlling the rate of change in volatility.
- $K = \exp(k)$ represents the variance in volatility. A large k indicates a high rate of change in volatility, while a small k indicates a stable environment.

Unlike the Rescorla-Wagner and Pearce-Hall models, which adjust learning based on a single point estimate (e.g., associative strength or associability), the Bayesian learner uses probability distributions to represent uncertainty. Specifically, the Bayesian learner updates its beliefs about r, v, and k based on the observed data. This allows the model to explicitly reason about how uncertain the learner is about reward probabilities and how much those probabilities might change in the future.

With a Markovian approach, where the next outcome probability $r_(i+1)$ depends only on the current probability r_i and not on the full history of previous probability, the posterior probability after observing an outcome y_i:

$$\Pr(r_{i+1}, v_{i+1}, k \mid y_{\leq i+1}) = \frac{\Pr(y_{i+1} \mid r_{i+1}, v_{i+1}, k)\ \Pr(r_{i+1}, v_{i+1}, k)}{\Pr(y_{i+1})} \tag{1.6.5.3}$$

where:

- $y_{\leq i+1}$ represents all observed data up to trial $i + 1$.
- $\Pr(y_{i+1} \mid r_{i+1}, v_{i+1}, k)$ is the likelihood of observing y_{i+1} given the reward rate r_{i+1} and volatility v_{i+1}.
- $\Pr(r_{i+1}, v_{i+1}, k)$ is the prior belief about reward probability, volatility, and rate of change before observing the new outcome.
- $\Pr(y_{i+1})$ is the marginal likelihood of observing outcome y_{i+1} under all possible parameter values.

Neuroimaging studies (e.g., (Behrens et al., 2007; Payzan-LeNestour et al., 2013)) show that the ACC and LC play a key role in monitoring environmental volatility, with greater activity observed in the ACC when individuals are adapting to uncertain or rapidly changing conditions. This suggests that the brain implements Bayesian-like computations to optimize decision-making under uncertainty.

Piray and Daw (2021) extended the model by estimating both noise (which they call stochasticity) and volatility simultaneously. Their approach suggests that the amygdala is involved in tracking volatility, though empirical evidence is lacking.

1.6.6 The pros and cons of Bayesian Learner

Unlike Rescorla-Wagner model and Pearce-Hall model, the Bayesian learner incorporates probabilistic learning, where the learner's beliefs are modelled as probability distributions. Furthermore, the model explicitly represents volatility and change of volatility using probability distributions, tracking both reward probabilities and volatility, offering additional insights into learning in uncertain environment.

However, the Bayesian learner depends on initial assumptions about the reward probabilities (r_i), volatility (v), and rate of change in volatility (k). For instance, the prior for reward probabilities (r_i) is determined by a beta distribution parameterised by previous reward probability and volatility, while volatility (v) assumes to follow a Gaussian distribution. Unlike reinforcement learning models, Bayesian learners typically do not have free parameters that can be fit to participant choice; rather, they describe an idealized, near-optimal learning process. This makes them less flexible in capturing individual differences in behaviours.

Additionally, Bayesian models have significant computational demands, as they require greater processing power and more parameters for belief updates. These computational requirements suggest that Bayesian learning models, in their full form, are unlikely to be directly implemented in the brain. Instead, the brain may approximate Bayesian inference through heuristic or simplified mechanisms that retain some of its benefits without the full computational complexity.

1.6.7 Hierarchical Bayesian framework

Mathys et al. (2014) introduced the Hierarchical Gaussian Filter (HGF) as a framework to model how individuals learn under different forms of uncertainty. The HGF operates at three levels, capturing a hierarchy of beliefs that govern predictions. Each level reflects different types of uncertainty, ranging from immediate predictions to abstract representations of volatility. At the first level, the model describes the participant's predictions on each trial (e.g., predicting specific outcomes). The second level captures the individual's beliefs about the probabilities underlying those predictions. At the third level, the model tracks the estimated volatility of these probabilities—essentially, how quickly the environment is changing.

We can express the three levels of the HGF in terms of the environmental state x_i^k

At Level 3,

$$\Pr(x_3^{(k+1)} \mid x_3^k, \theta) \sim \mathcal{N}(x_3^k, \theta) \tag{1.6.7.1}$$

At Level 2,

$$\Pr(x_2^{(k+1)} \mid x_2^k, x_3^{(k+1)}) \sim \mathcal{N}(x_2^k, \exp(kx_3^k + \omega)) \tag{1.6.7.2}$$

At Level 1,

$$\Pr(x_1^{(k+1)} \mid x_2^{(k+1)}) \sim \text{Bernoulli}(\text{sigmoid}(x_2^{(k+1)})) \tag{1.6.7.3}$$

k, ω could be fitted to participants behavioural data with k showing how much Level 2 is affected by Level 3 and ω as a constant.

At each trial, the model generates an initial estimate for each level before the outcome is known and then updates this estimate based on the actual outcome. The model employs Gaussian distributions to represent predictions at each level, characterized by a mean (μ_i) and variance (σ_i^2), where i indicates the level (1, 2, or 3 in this model). The variance, σ_i^2, quantifies the uncertainty of the estimate at each respective level. Specifically, the variance at the first, second, and third levels (σ_1^2, σ_2^2, and σ_3^2) reflects irreducible, estimation, and volatility uncertainties, respectively.

At each trial, the model generates an initial estimate for each level before the outcome is known and updates this estimate based on the actual outcome. Gaussian distributions represent the predictions at each level, characterized by a mean (μ_i) and variance (σ_i^2), where i denotes the level. The variances σ_1^2, σ_2^2, and σ_3^2 reflect different types of uncertainty:

- σ_1^2 reflects irreducible uncertainties, i.e., noise
- σ_2^2 reflects estimation uncertainties
- σ_3^2 reflects volatility

Belief updating at each level and at each trial is driven by PEs modulated by precision ratios. Precision is inverse of variances at each level i and each trial k denoted by π_i^k.

$$\pi_i^k = \frac{1}{(\sigma_i^k)^2} \tag{1.6.7.4}$$

The precision ratio between two levels dictates how much weight is given to information from the lower level relative to the current level. This ratio is expressed as:

$$\frac{\hat{\pi}_{i-1}^{(k+1)}}{\pi_i^{(k+1)}} \tag{1.6.7.5}$$

Here, $\hat{\pi}_{i-1}^{(k+1)}$ represents the precision of the prediction at trial $k+1$ at the level below, where $\pi_i^{(k+1)}$ represents the precision of the current level. In short, it compares how much the individual is learning to how much the individual already knows.

$$\mu_i^{(k+1)} = \mu_i^k + \varepsilon_i, \quad \varepsilon_i = \frac{\hat{\pi}_{i-1}^{(k+1)}}{\pi_i^{(k+1)}} \cdot \mathrm{PE}_{i-1}^{(k+1)} \tag{1.6.7.6}$$

The term $\frac{\hat{\pi}_{i-1}^{(k+1)}}{\pi_i^{(k+1)}}$ functions as the learning rate (LR), determining how much weight is given to the prediction error at each trial.

Iglesias et al. (2013) explored the relationship between hierarchical prediction errors from the HGF and brain activity. They focused on sensory learning tasks,

where participants predicted the association between auditory and visual stimuli (tones and images). The study found that low-level prediction errors correlated with activity in the ventral tegmental area (VTA) and substantia nigra (SN), regions associated with dopaminergic neurotransmission. These errors related to immediate, concrete predictions (e.g., visual stimulus outcomes) independent of reward. In contrast, high-level prediction errors were associated with activity in the basal forebrain, a region involved in cholinergic activity. This supports the idea that higher levels of the brain's predictive hierarchy, which deal with abstract, probabilistic information (like volatility), may rely on different neuromodulatory systems—specifically, acetylcholine for higher-level PEs and dopamine for lower-level PEs.

Other empirical support for HGF includes pupil dilation correlating with irreducible uncertainties (de Berker et al., 2016), pharmacological manipulating NE increases learning rate derived from HGF (Marshall et al., 2016). The HGF has also been successfully applied in several investigations of learning and action in volatile environments (De Berker et al., 2016; Diaconescu et al., 2014, 2017; Hauser et al., 2015; Iglesias et al., 2013; Vossel, Bauer et al., 2014; Vossel, Mathys et al., 2014; Vossel et al., 2015). Furthermore, HGF seems to be an effective tools in studying a range of psychiatric symptoms, including autism (Lawson et al., 2017) and anxiety (Hein et al., 2023).

1.6.8 The pros and cons of HGF

Overall, the HGF is a valuable framework for estimating learning adaptations to uncertainty by modelling learning at multiple levels, from immediate predictions to abstract representations of volatility. Its hierarchical structure enables nuanced probabilistic reasoning, making it particularly effective in exploring complex decision-making process. Additionally, the HGF has free parameters that can be fit to choice behaviour. However, there lacks strong evidence for parameter recoverability, and its frequent application to simple tasks limit its interpretability.

1.6.9 Hidden state inference

Another computational model that describes how the brain learns under uncertainty is state-dependent learning. Unlike the gradual learning models discussed earlier, where beliefs are updated slowly and continuously with each new piece of evidence, state-dependent learning involves recognizing and switching between different states when a significant change in the environment is detected (Gershman & Niv, 2012). This model assumes that the brain learns by inferring distinct states or contexts and adjusts its beliefs when it detects a change in the environment that indicates a switch to a new state. This class of model is beyond the scope of this thesis, so it will only be briefly discussed here.

In state-dependent learning, the learner's belief about the probability of an outcome (e.g., receiving a shock) is represented by a Beta distribution with parameters α and β. These parameters are updated based on the trial outcomes and controlled by a decay parameter λ.

For a positive outcome (e.g., receiving a shock when expected):

$$\alpha_{(t+1,s)} = \lambda(\alpha_{(t,s)} + \tau^{+}) \tag{1.6.9.1}$$

$$\beta_{(t+1,s)} = \lambda\beta_{(t,s)} \tag{1.6.9.2}$$

For a negative outcome (e.g., not receiving a shock when expected):

$$\alpha_{(t+1,s)} = \lambda\alpha_{(t,s)} \tag{1.6.9.3}$$

$$\beta_{(t+1,s)} = \lambda(\beta_{(t,s)} + \tau^{-}) \tag{1.6.9.4}$$

Here, τ^{+} and τ^{-} represent the weight given to each outcome (positive or negative), reflecting how much attention is paid to those outcomes.

In a multi-state model, individuals can infer and switch between different underlying states based on their experiences. Each state is characterized by a Beta distribution with parameters α and β. The model keeps track of a running average of surprise, which is used to detect changes in the environment. Specifically, surprise $S_{(t,s)}$ is calculated as:

$$S_{(t,s)} = (1-\pi)S_{(t-1,s)} + \pi \left| O_t - \hat{P}_{(t,s)} \right| \tag{1.6.9.5}$$

Where π weights current and past surprise keeping surprise values between 0 and 1. $P_{(t,s)}$ represents the expected probability of receiving the positive outcome, which in this case is the mode of the beta distribution at trial t,

$$P_{(t,s)} = \frac{\alpha - 1}{\alpha + \beta - 2} \tag{1.6.9.6}$$

To distinguish between the switching between different states and creating a new state, the model compares surprise to uncertainty of the current state, which is defined as the standard deviation of the beta distribution,

$$\sigma = \sqrt{\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}} \tag{1.6.9.7}$$

If surprise $S_{(t,s)}$ exceeds the product of current uncertainty (σ) and the threshold parameter η, this indicates the current state may no longer adequately represent the environment. The model then examines all existing states against the inferred probability ($P_{(t,s)}$) to determine whether to remain in the current state or transition onto another existing state. However, if

$$S_{(t,s)} > \sigma\eta q \tag{1.6.9.8}$$

where q represents the difficulty of creating a new state (derived from the Chinese restaurant process), the model concludes that no existing state represents the environment and creates a new state.

Zika et al. (2023) tested this concept in a probabilistic reversal learning task with different reward schedules (e.g., 90/10, 75/25, 60/40). The study found that participants were more likely to switch from gradual learning (in conditions with small changes, such as 60/40) to state-inference learning (in conditions with larger changes, such as 90/10). This suggests that human learners use state inference when there are large, noticeable shifts in the environment's reward contingencies.

1.6.10 The pros and cons of Hidden state inference

Hidden state inference models adapt by inferring and switching between distinct states when significant environmental changes occur, making them particularly suitable to apply for situations with large and noticeable shifts in contingencies. Unlike models that continuously update beliefs based on recent experiences, hidden state inference models handle uncertainty by tracking the likelihood of being in different latent states and shifting a new state when sufficient evidence suggests a change has occurred. This allows them to efficiently manage abrupt changes in the environment.

However, these models rely on the assumption that individuals perceive the world as governed by multiple discrete states. While this approach effectively captures an environment with change points, such as abrupt shifts in rewards contingencies, it may struggle with more gradual or fluctuating forms of uncertainty (e.g., a drift in reward probabilities). Additionally, the process of state creation is inherently subjective and lacks robust empirical evidence to support that it is indeed how state creation is carried out.

In the analysis section of this thesis, we predominantly employed RL models and their modifications due to their simplicity and practicality. By focusing on describing individual learning rate and avoiding assumptions about how people are solving the tasks (as required by Bayesian learner and hidden state inference), RL offers an empirically grounded framework for analysing behaviour in uncertain environment. Additionally, their parameter recoverability is higher compared to more complex models like the HGF, making it more suitable choice of tasks designed in this thesis.

1.7 Computational Models of learning under Momentum

An underexplored question in the study of volatility is how different types of volatility influence learning. Volatility is generally described as the extent to which the thing being learned is changing, but not all changes are the same. One such

form is momentum, which refers to a persistent trend or directional movement in the environment over time. While momentum shares similarities with volatility – both involve changes in the mean of a variable – it has a crucial distinction: momentum implies a sustained directional trend rather than random fluctuations. Understanding how learners detect and adapt to momentum is crucial because it allows them to accelerate learning by anticipating where the environment is headed. This, in turn, leads to more accurate mental representations of the environment and better decision-making.

1.7.1 Proportional-Integral-Derivative Controller

To date, only one study, Ritz et al. (2018), has specifically investigated whether humans can detect these trends and adjust their beliefs accordingly. The study employed a Proportional-Integral-Derivative (PID) controller, a model commonly used in control theory, to explain how participants integrate information from previous trials and immediate prediction errors to update their expectations about the environment. Ritz et al. found that participants relied on historical PEs from as far back as 10 trials when updating their beliefs. This pattern was observed both in a change-point paradigm, where abrupt shifts occurred, and in an experimental condition where the schedule followed a trend. However, they did not find evidence of behaviour uniquely adapted to learning momentum – participants in the momentum environment did not exclusively rely on long-term history (e.g., the last 10 trials), nor did participants in the change-point condition limit themselves to short-term history (e.g., the last 3 trials). Instead, their use of past information was relatively consistent across conditions.

The PID model consists of three core components:

- Proportional (P) Term: This term generates an output directly proportional to the current prediction error. A larger prediction error leads to a stronger correction in future beliefs, making this component similar to how the Rescorla-Wagner model adjusts beliefs based on prediction errors.

- Integral (I) Term: The integral term focuses on the cumulative sum of all past prediction errors. By accounting for accumulated error over time, it helps to address persistent biases or trends that the proportional term alone might miss.

- Derivative (D) Term: This term accounts for the rate of change of the prediction error, allowing the model to adjust more rapidly when the environment is changing at an accelerating or decelerating pace.

While a model using only the P term closely resembles the Rescorla-Wagner learning model, the PID controller incorporates past information (via the I term) and anticipates future changes (via the D term), allowing it to model dynamic environments more flexibly. Formally, the model is expressed as:

$$\text{Belief Update} = \text{intercept} + \beta_1 \cdot \text{PE}_t + \beta_2 \cdot \sum_{n=1}^{t} \lambda^{t-n} \, \text{PE}_n + \beta_3 \cdot (\text{PE}_n - \text{PE}_{n-1})$$

Where:

- β_1, β_2, β_3: represents the beta weights

- λ is a memory persistence term

1.7.2 Alpha-Beta Filter

In addition to the PID model, the alpha-beta filter is often used to estimate moving targets in environments with constant acceleration. Unlike the PID controller, the alpha-beta filter does not require participants to recall historical prediction errors. Instead, it uses two parameters:

- Alpha (similar to the learning rate in the Rescorla-Wagner model), which controls how much the belief estimate is adjusted based on the prediction error.

- Beta, which adjusts the velocity estimate of the change based on the current prediction error.

$$V_t^0 = V_{t-1}^0 + V_{t-1}^1 + \mathrm{LR}^0 \cdot \mathrm{PE}_{t-1} \tag{1.7.2.1}$$

$$V_t^1 = V_{t-1}^1 + V_{t-1}^1 + \mathrm{LR}^1 \cdot \mathrm{PE}_{t-1} \tag{1.7.2.2}$$

Where:

$$\mathrm{PE}_{t-1} = O_{t-1} - \mathrm{Belief}_{t-1} \tag{1.7.2.3}$$

$$= O_t - (V_{t-1}^0 + V_{t-1}^1) \tag{1.7.2.4}$$

The updated belief is a combination of the previous estimate, the prediction error, and the velocity estimate, allowing for more refined adjustments in a dynamic environment.

$$\mathrm{Belief}_t = V_t^0 + V_t^1 \tag{1.7.2.5}$$

Eldar et al. (2016) introduced a variant of the alpha-beta filter to model how humans track the moving average reward rate in volatile environments. However, the efficacy of both the alpha-beta filter and the PID model in capturing actual human behaviour remains uncertain and has yet to be empirically confirmed.

In this thesis, we tested the PID controller and the alpha-beta filter, alongside a novel RL model, which we introduce in later chapters. We selected PID controller and alpha-beta filter because they present two distinct approaches to learning in environments with structured changes: the PID controller assumes that individuals integrate feedbacks from historical trials, while the alpha-beta filter assumes that individuals estimate the rate at which the environment changes. The novel RL model is designed to further capture adaptive learning under momentum by providing an alternative framework for understanding how individuals adjust to change. Importantly, all three models are relatively simple and make minimal assumptions about how human learners solve the task, increasing the likelihood of successful parameter recovery and making them well-suited for over experimental design.

1.8 Clinical Significance

It has been hypothesized that maladaptive learning under uncertainty is a key cognitive process that may drive the development of anxiety disorder and depression (Pulcu & Browning, 2019). Research suggests that when individuals struggle to accurately represent uncertainty in the environment, their internal models fail to adapt properly to changing conditions, leading to suboptimal decision-making (Browning et al., 2015; de Berker et al., 2016; Gagne et al., 2020; Huang et al., 2017). This impairment in learning is thought to reduce the ability to predict negative events, which in turn may contribute to a sense of diminished agency and control, both of which are central features of anxiety and depression. Difficulties in adapting to uncertainty may also underlie the rigid cognitive and behavioral patterns observed in these mood disorders, reinforcing maladaptive responses to changing environment. Further understanding of how individuals with anxiety and depression learn under uncertainty could provide valuable insights into their cognitive processes and inform targeted interventions.

One symptom closely linked to maladaptive learning under uncertainty is intolerance of uncertainty (IU), which is a core feature of anxiety disorder. It refers to the extent of which one would react negatively on emotional, cognitive and behavioural level to uncertain situations and events (Buhr & Dugas, 2009). Reducing IU has been linked to post-treatment reductions in anxiety and depression symptoms (Boswell et al., 2013), suggesting that IU may play a causal role in the development of these disorders. However, it remains unclear whether IU, as measured by questionnaires, directly relates to processes like adaptation of learning rates in response to environmental uncertainties.

Another process that exemplifies maladaptive learning under uncertainty is learned helplessness, first identified by Overmier and Seligman (1967) through experiments with dogs. It refers to a phenomenon where a subject (whether human or animal), after being repeatedly exposed to an unavoidable stressor, develops a behavioural pattern characterized by a diminished ability to escape the same stressor, even when escape is later made possible. The key issue in learned helplessness

may be an inability to maintain a belief that the environment can change. In other words, once individuals or animals have learned that their actions have no effect on outcomes, they may fail to update their beliefs when contingencies shift. This rigid expectation that control is permanently absent could be viewed as a failure to detect or respond to volatility, leading to sustained passivity even when opportunities for change arise.

Although both intolerance of uncertainty and learned helplessness could be driven by maladaptive learning processes in anxiety and depression, the causal relationship between these processes and development of these disorders remains unclear as most evidence is correlational (which will be discussed in detail in the following section). Further research is needed to determine the causal role of these processes, potentially through longitudinal studies or targeted interventions designed to improve learning under uncertainty and assess their impact on symptom reduction.

1.8.1 Negative affectivity associated with learning rate adaption

Higher anxiety is often associated with reduced adaptation to environmental volatility (de Berker et al., 2016; Gagne et al., 2020; Huang et al., 2017). Two competing hypotheses have been proposed to explain the mechanisms underlying this association. The first one hypothesizes that anxious individuals misattribute volatility to noise. According to Pulcu and Browning (2017), anxious individuals may not distinguish between predictable changes in the environment (volatility) and random fluctuations (noise), leading to maladaptive decision-making. Alternatively, individuals with higher anxiety could have an elevated learning rate in all conditions. Piray and Daw (2021) proposed that this could cause anxious individuals to show less flexibility when transitioning from stable to volatile environments. In this view, anxious individuals fail to adapt their learning rates, leading to poor adjustment when environmental conditions shift. However, empirical support for this hypothesis remains inconsistent, as some studies have failed to find a strong link between baseline learning rates and anxiety (Browning et al., 2015; Satti et al., 2024).

To test the first hypothesis (misattribution of volatility to noise), there would need to be an experimental paradigm where both volatility and noise are manipulated in ways that are difficult to distinguish from one another. However, most studies so far have not sufficiently manipulated noise levels. For instance, Huang et al. (2017) found that high-anxiety individuals tended to employ a sub-optimal decision-making strategy characterized by an elevated lose-shift rate, meaning they were more likely to switch away from a chosen option after experiencing a loss, regardless of whether the loss was statistically expected or due to random chance. Similarly, de Berker et al. (2016) did not manipulate noise levels directly but used the HFG to quantify participants' subjective estimates of noise. Their findings showed that higher subjective estimates of irreducible uncertainty were strongly associated with elevated stress response.

On the other hand, the second hypothesis, that anxious individuals maintain higher learning rates across all environments, has received mixed support. Some studies show increased punishment learning rates in participants with mood and anxiety disorders (Aylward et al., 2019), but others do not find a clear relationship between baseline learning rates and anxiety (Browning et al., 2015).

Recent studies have suggested that the relationship between anxiety and learning may involve differences in explorative behavior (Aberg et al., 2022; Fan et al., 2023). For example, Fan et al. (2023) demonstrated that individuals with high somatic anxiety symptoms showed reduced directed exploration, meaning that they were less likely to explore uncertain options in decision-making tasks, opting instead for familiar, known options regardless of the uncertainty level. This behavioral tendency may contribute to the misadaptation to volatility seen in anxious individuals, as they may not adequately explore new possibilities in changing environments.

Interestingly, a study by Zika et al. (2023) found the opposite effect of trait anxiety on learning rate adaptation. In a probabilistic reversal learning task, Zika et al. tested participants using different experimental conditions (90/10, 75/25, and 60/40 reward schedules) to assess learning rate changes across environments. Participants with higher trait anxiety showed increased learning after reversals in

the 90/10 condition (where environmental changes were large and obvious) but reduced learning from oddball events (where changes were less meaningful or more random), although the authors did not find this effect at other probability levels. This suggests that high-anxiety individuals are more attuned to significant changes (reversals) but may struggle with more subtle or random fluctuations (oddballs).

Fitting a hidden-state model to the data, Zika et al. (2023) found that the n-state model, which allows for switching between distinct environmental states, was a better fit for participants with high trait anxiety. This suggests that anxious individuals may be more inclined to assume that the environment has shifted to a new state rather than gradually updating their beliefs within the same state. In particular, anxious participants may infer that large, salient changes indicate a new state of the environment, while smaller, random changes (oddballs) do not prompt as much learning or belief updating.

There are some notable differences in experiment designs between Zika et al. (2023) and Browning et al. (2015). First, Zika et al. (2023) asked participant to estimate the probability of receiving a shock, while Browning et al. (2015) instructed participants to choose an option that was least likely to deliver a shock. Furthermore, a significant difference in the change of estimation in high trait anxiety is only observed in the 90/10 condition, while there was no group difference between low and high anxiety in the 60/40 an 75/25 conditions.

Zika et al. (2023)'s findings raise the interesting hypothesis that state-inference (switching between different mental models of the environment) might be more common in individuals with high trait anxiety, compared to gradual learning (updating beliefs slowly based on new evidence). However, this thesis does not directly test this hypothesis. Instead, it is better suited examining how volatility and noise influence adaptation of learning. The outstanding questions my work aims to address relate to how individuals adjust their learning strategies in response to different types of environmental uncertainty.

Specifically, two key points from Zika et al. (2023)'s findings could be interesting for future research:

- State-switching vs. Gradual Learning: Individuals may switch between state-inference and gradual learning based on how noisy the environment is. When noise is low and changes are clear (e.g., in the 90/10 condition), anxious individuals rely more on state-inference. When noise is high, gradual learning might dominate.

- Anxiety and Environmental Structure: Higher trait anxiety may be associated with better recognition of environmental structure, such that anxious individuals are better at differentiating between reversals and oddballs when the environment is less noisy. This suggests that anxious individuals may overestimate the likelihood of large, meaningful changes, leading to state-switching rather than gradual adaptation.

In summary, the relationship between anxiety, learning rate adaptation, and state inference remains complex. Zika et al. (2023) suggests that anxious individuals might show more state-inference as opposed to fast, gradual learning. In other cases, misattribution of volatility to noise (Pulcu & Browning, 2019) could explain why anxious individuals struggle to adjust to environmental changes.

Building on this work, this thesis will further develop a task that incorporates both volatility and noise to examine the possibility of such misattribution in anxious individuals. This approach will help to clarify whether misattributions contribute to maladaptation of learning rate and to untangle the mechanisms of these processes more rigorously.

1.8.2 Symptom Specificity in Learning Rate Adaptation: Exploring the Overlap Between Anxiety and Depression

Most of the evidence reviewed thus far relates to the adaptation of learning rate and trait anxiety (de Berker et al., 2016; Gagne et al., 2020; Huang et al., 2017; Zika et al., 2023). However, trait anxiety is highly correlated with depression symptoms, making it difficult to disentangle whether these findings are specific to anxiety or reflect broader internalizing symptoms. Fewer studies have examined

how depressive symptoms relate to adaption of learning rate under uncertainty. Mukherjee et al. (2020) showed that patients with depression exhibited lower learning rates across both reward and punishment conditions compared to healthy controls in a probabilistic reversal learning task. However, the task did not include volatile blocks (more frequent reversals) and stable blocks (less frequent reversals), which would allow for an assessment of how each individual adapt their learning rates to environmental changes. Despite this, the authors did report that both depressed individuals learned the new contingencies more slowly due to lower learning rates after a reversal.

The relationship between depressive symptoms, anxiety symptoms, and learning adaptation in volatile environments remains unclear, as trait anxiety strongly correlates with depression symptoms. To study the shared mechanism between anxiety and depression and distinguishes these from disorder-specific symptoms, Gagne et al. (2020) recruited one group of participants with generalized anxiety disorder, one with major depressive disorder, and one healthy control group. Participants were asked to complete two versions of the probabilistic reversal task: one with outcome as a financial rewards (reward version) and one with outcome as electrical shock (aversive version). Through bifactor modelling, the authors separated the symptom variance to components common to anxiety and depression, i.e., general factor and those specific to each disorder. They found impaired learning rate adaption to volatility was linked to general factor. However, the study did not identify any specific depressive or anxious symptoms that were uniquely linked to learning adaptation difficulties in volatile contexts.

These results suggest that shared symptoms between depression and anxiety are central to maladaptive learning processes under uncertainty. Building on this work, this thesis specifically recruited participants based on depression symptoms to further investigate these shared mechanisms.

1.9 Outstanding Questions & Experimental Approach to Answering Them

Learning under uncertainty has been studied in both animals and humans. However, there remain some outstanding questions. This thesis addresses these gaps in the following chapters through a series of experiments.

1.9.1 How well do human learners differentiate between noise and volatility?

While strong evidence indicates that humans adapt their learning rates normatively in volatile environments, there is limited understanding of how they adjust learning rates in response to environmental noise. Additionally, it remains unclear how effectively humans can navigate environments that are both volatile and noisy, a process likely dependent on their ability to distinguish between the two.

Experimental approach:

To investigate this, we designed an experiment where noise and volatility were manipulated simultaneously. Chapter 2 details the experimental design and methodology used to vary these parameters systematically. By identifying conditions where distinguishing between noise and volatility becomes particularly challenging, we explored how participants adapted to volatile and noisy environment.

1.9.2 What are the neural processes underlying the attribution and signalling of volatility and noise?

A recent review by Nassar proposed an expanded role for NE, suggesting it signals transitions between latent states. The theoretical framework by Dayan and Yu posits that NE signals volatility, while ACh signals noise. Here, 'signals' means that we that fluctuations in NE and ACh activity encode or track environmental properties, volatility and expected noise respectively, in a form that other neural systems can use to adjust learning. In this model, ACh "gates" NE's effectiveness in signalling environmental volatility. Although there is some evidence supporting NE's role in signalling volatility and regulating learning rate adaptation, empirical

evidence for ACh's involvement including the gating mechanism itself remains lacking. The precise neural mechanisms by which the brain differentiates between noise and volatility remain an open question.

Experimental approach:

Although NE is not the only neurotransmitter affecting pupil dilation, LC-NE activity is strongly correlated with changes in pupil size. To explore the neural basis of noise and volatility attribution, we recorded pupillary data while participants completed the task developed in Chapter 1. In Chapter 3 and 4, we examined how volatility, linearly structured volatility, and noise influence pupil dilation. These analyses provide insight into the neuromodulatory processes underlying uncertainty attribution.

1.9.3 How do human learners adapt under linearly structured volatility?

Volatility can exhibit structured patterns, such as linear trends where environmental changes shift predictably in one direction (i.e., momentum). Recognizing and leveraging such structures enables more efficient learning and accurate predictions. However, human learning under momentum has not been systematically studied. Although some models address these types of environmental changes (e.g., Eldar et al. (2016) and Ritz et al. (2018)), empirical evidence remains sparse.

Experimental approach:

To examine how human learners adapt to momentum, we modified the task introduced in Chapter 1. In Chapter 4, we investigated whether individuals could recognize and integrate higher-order uncertainty into their predictions. We also proposed a novel reinforcement learning model that may explain the cognitive mechanisms underlying adapting to structured volatility.

1.9.4 Are adaptations of learning to noise and linearly structured volatility associated with depressive and anxiety symptoms?

Maladaptive learning under uncertainty may cause to anxiety and depression, though there is causal evidence. Difficulty interpreting uncertainty impairs internal models, leading to suboptimal decisions, reduced ability to predict negative events, and diminished sense of control. These factors may contribute to symptoms such as intolerance of uncertainty and learned helplessness. Understanding whether adaptations to noise and linearly structured volatility are associated with depressive and anxiety symptoms could provide critical insights into the cognitive processes underlying these conditions.

Experimental approach:

To address whether learning adaptation under uncertainty are associated with anxiety and depressive symptoms, we recruited a sample with spectrum of mood disorder symptoms. Since maladaptive responses to volatility are not uniquely characteristic of either anxiety or depression, we purposively sampled participants based on their depressive symptoms to capture a range of learning behaviours under uncertainty.

2

Designing a Behavioral and Computational Framework to Dissociate Volatility and Noise in Learning

2.1 Introduction

Learning in an uncertain environment is a fundamental part of cognition. Various types of uncertainties influence learning in distinct ways. In particular, in a volatile environment, one should increase how much one learns from recent observations and learn less from recent observations when the environment is noisy (Yu & Dayan, 2005). However, Individuals with high levels of anxiety have difficulty adapting to volatility (Browning et al., 2015), although this association has never been tested in situations where both the volatility and noise vary. Going beyond mood disorders, it is worth noting that the impairment in volatility adjustment also extends to other disorders and psychiatric symptoms, including autism (Lawson et al., 2017), compulsivity (Vaghi et al., 2017), and paranoia (Nour et al., 2018).

As discussed in Chapter 1 – Clinical Significance, when individuals fail to accurately represent environmental uncertainty, their internal model cannot adapt effectively, leading to poor decision-making. This maladaptive learning under uncertainty may contribute to developing some of the symptoms in mood disorders,

such as learned helplessness and intolerance of uncertainty. Understanding why this impaired adaption to volatility occurs could be key to understand why mood disorders develop in the first place.

One hypothesis for such observation relates to participants misattributing noise as volatility (Pulcu & Browning, 2019). An alternative hypothesis is that anxious/depressed individuals have a generally higher baseline LR which reduces the degree to which they adapt to both volatility and noise (Piray & Daw, 2021). Alternatively, people with mood disorder could simply have trouble detecting both noise and volatility, where participants have a generalized impairment in using variations and structure in the observed data to learn causal statistics in the environment. To test these hypotheses, we need to design an experiment where both volatility and noise vary.

Previous research has mainly focused on studying learning under variable levels of volatility either using a change-point detection paradigm or a gaussian random walk in various contexts (e.g., auditory, sensory). Only a few studies have looked at how participants adapt their LR to varying levels of noise and these studies have only investigated adaptation under specific, and quite restrictive, conditions. For instance, Nassar et al. (2010) studied noise with a change-point paradigm, where the 'jump' (i.e. the volatility) was around 100 units in magnitude whereas the standard deviation of the outcome (which determines the noise) was set at 5, 15, or 35. In this case therefore, the noise was very much smaller than the volatility and thus the two could be easily distinguished by the size of the variation. Similarly, Diederen and Schultz (2015) used a learning paradigm in which an onscreen cue explicitly indicated the level of noise used to generate the outcomes. In other words, how people adapt to changes in noise has only been described for situations in which the noise is clearly distinguishable from volatility, or when it is explicitly labelled. It remains unclear to what extent human participants estimate and adapt to noise when it produces similar changes in an outcome as volatility and when it is not labelled.

In this study, we investigated how humans adapt to volatility and noise when both factors were present and variable. The focus of this research was to examine the

normative adjustments in human learning rates in response to these uncertainties. Although our ultimate goal was to explore the relationship between this adaptation process and anxiety and depressive symptoms, this study focused on unselected participants to uncover the underlying mechanisms of adaptation. These findings = serve as a foundation for future in-person studies recruiting participants based on their symptom profiles.

To achieve this, we designed and validated an experiment paradigm involving a coin catching task, where the coin's landing positions follow a von mises random walk. Volatility and noise were manipulated by varying the step size of changes in the mean and the concentration parameter (analogous to standard deviation), respectively. We tested a range of volatility and noise levels to see how well human learners could distinguish between the two. Additionally, we compared two designs: one with an adjustable bucket size and one with a fixed bucket size. The former one allow us to determine whether participants' estimates of noise levels could be inferred from their bucket adjustment. These manipulations of uncertainties was critical for assessing how well humans distinguish between noise and volatility. We hypothesized that participants would increase their learning rates under high volatility but decrease it under high noise. While we did not specifically select participants based on anxiety or depressive symptoms, we evaluated previous theories suggesting that anxious individuals may struggle with adapting to volatile situations by analyzing participants behaviors in this task.

2.2 Methods

We recruited participants online through Prolific (https://www.prolific.co/), and tasks and consent procedures were implemented using Gorilla (https://gorilla.sc/). After signing up for the study on Prolific, participants were provided with a link to download the Participant Information Sheet (PIS). Participants were then asked to complete an online consent form. Participants who either did not complete the consent form or declined to give consent were redirected back to Prolific.

Upon providing consent, participants received instructions for the coin-catching task. The task blocks were presented in random order, and the task itself lasted approximately 25 minutes. Following the coin-catching task, participants completed the Automated Operation Span Task (AOSPAN) to assess working memory capacity and a series of questionnaires, including the trait anxiety subscale of the STAI, WSAS, CES-D, and TEPS. The study was approved by the Oxford University Research Ethics Committee and the entire experimental session lasted approximately one hour.

2.2.1 Sample

In total, we recruited 404 participants. The inclusion criteria were being 18 years or older, living in the UK or USA, fluent in English, and having an approval rating of 90 – 100 on the Prolific platform. We aimed to recruit 50 participants per experiment. However, due to multiple participants signing up simultaneously, some experiments occasionally exceeded the target of 50 participants (Table 2.1).

Experiment	Participants	Age (SD)	Females (% sample)
Experiment 1	50	28.87 (10.98)	24 (48%)
Experiment 2	51	25.96 (7.49)	21 (41%)
Experiment 3	50	31.3 (11.41)	17 (34%)
Experiment 4	50	24.5 (4.23)	16 (32%)
Experiment 5	50	32.5 (8.75)	22 (44%)
Experiment 6	50	36.6 (12.67)	24 (48%)
Experiment 7	50	35.9 (11.02)	19 (38%)
Experiment 8	53	39.2 (12.06)	27 (51%)
Total	**404**	**31.8 (11.22)**	**170 (42%)**

Table 2.1: Demographics of sample recruited for Experiment 1–8.
We aimed to recruit 50 participants per experiment with inclusion criteria as being 18 years or older, living in the UK or USA, fluent in English, and having an approval rating of 90–100 on the Prolific platform.

2.2.2 Data Quality Assurance

We identified and excluded participants who failed to demonstrate engagement with the choice task (Table 2.2). Specifically, we excluded data of participants who failed to move the bucket more than 25% of the trials (i.e., 50 out of 200). We also excluded data from participants who did not the responded to attention check

questions and who exceeded maximum time to complete the task (90 mins). Finally, some participants completed the same block twice, most likely due to internet connection issues. We excluded those participants from our analysis too.

Experiment	Recruited	Lack of Engagement	Repeated Block	Failed Attention Check	Analyzed
Experiment 1	50	0	1	2	47 (94%)
Experiment 2	51	0	0	0	51 (100%)
Experiment 3	50	0	3	0	47 (94%)
Experiment 4	50	4	0	0	46 (92%)
Experiment 5	50	4	3	1	42 (84%)
Experiment 6	50	5	2	0	43 (86%)
Experiment 7	50	6	0	0	44 (88%)
Experiment 8	53	3	1	0	49 (92%)
Total	**404**	**22**	**10**	**3**	**369 (91%)**

Table 2.2: Data quality assurance.
We initially recruited 404 participants for Experiments 1–8. Participants were excluded if they failed to move their bucket in more than 25% of trials, failed the attention check, or exceeded the maximum time allowed to complete the task ($n = 35$). This resulted in a final sample of 369 participants included in the analysis.

2.2.3 Coin Catching Task

We modified the Vaghi et al. (2017) coin catching task, where a coin is released from the center of a circle at every trial. In this task, a coin was released from the center of a large circle. A yellow bucket, positioned at the edge of the circle with a central angle of 30 degrees, was used by participants to catch the coin (see Figure 1). The task consisted of four blocks, each containing 50 trials: high volatility/low noise, high volatility/high noise, low volatility/low noise, and low volatility/high noise.

We chose the Vaghi et al. (2017) experimental framework because the circular design is suitable for noise exploration. Unlike linear distributions, the circular distribution is unbounded, meaning there are no artificial minimum and maximum values, which could otherwise constrain the task. Additionally, this design minimizes reliance on numerical strategies, such as computing a rough running average across trials, thereby ensuring participants focus on predictive learning as opposed to mathematical estimation. In contrast, some studies (e.g., (Nassar et al., 2010; Wilson et al., 2013) asked participants to explicitly predict the next number in a

series. This approach may inadvertently emphasize mathematical estimation over how people adapt their learning under uncertainty.

Experiments 1 through 3 used a resizable bucket, while in Experiments 4 through 8, the bucket size was fixed. The resizable bucket was designed as a proxy for participants' confidence in their predictions. In particular, higher confidence was reflected in smaller bucket size. This allowed participants to adjust their bucket size based on how certain they were about the coin's landing position.

Each coin was worth 18 points, and the cost of the bucket was determined by its size (measured in degrees). Increasing the bucket size resulted in a deduction of points, creating a trade-off between accuracy and cost. Each degree deducted 0.05 points, such that the final score was calculated as $18 - (\text{bucket size in degrees}) \times 0.05$. This trade-off was introduced to prevent participants from simply using the largest possible bucket size to guarantee catching the coin. For instance, using a 180-degree bucket (covering half the circle) results in receiving 0 points, even if the coin was successfully caught. This mechanism incentivized participants to balance their confidence in the coin's predicted position with the potential cost of increasing the bucket size. Participants could adjust the bucket size using a green slider on the left, and a bar next to the slider displayed the bucket size used in the previous trial (Figure 2.1b).

In Experiment 4 to 8, we kept the bucket size constant because the adjusting the bucket size could be used as a competing strategy to adapt to noise and thus, affect our ability to fully and and consistently observed the noise effect under different schedules. Additionally, interpreting the bucket size data would require making assumptions about the optimal cost-to-size trade-off, which added complexity to both the task design and data analysis. In the fixed bucket size version of the experiment, coins were no longer assigned point values. Instead, catching a coin was simply counted as "1 coin," with the total count displayed as an incrementing counter (Figure 2.1a).

In this online experiment, participants controlled the bucket using a mouse, allowing them to adjust its position. After dragging the bucket to their predicted

location, they submitted their predictions by clicking the "Play!" button. There was no time limit per trial. During the instruction phase, participants were informed that their task was to catch as many coins as possible and that the coin's landing position would change with each trial.In the version of the experiment in which bucket size was adjustable, participants were shown how to adjust bucket size using a vertical slider. In this version of the task, each coin was worth 18 points, and the cost of the bucket was calculated as its final size multiplied by 0.05. Participants were also shown how bucket size affects both cost and the net points they could gain. After the instructions, participants completed 10 practice trials. Before beginning the main task, we asked participants to answer two multiple choice questions to check their understanding of the task: "What is the purpose of the task?" and "How do I decide where to put the bucket?"

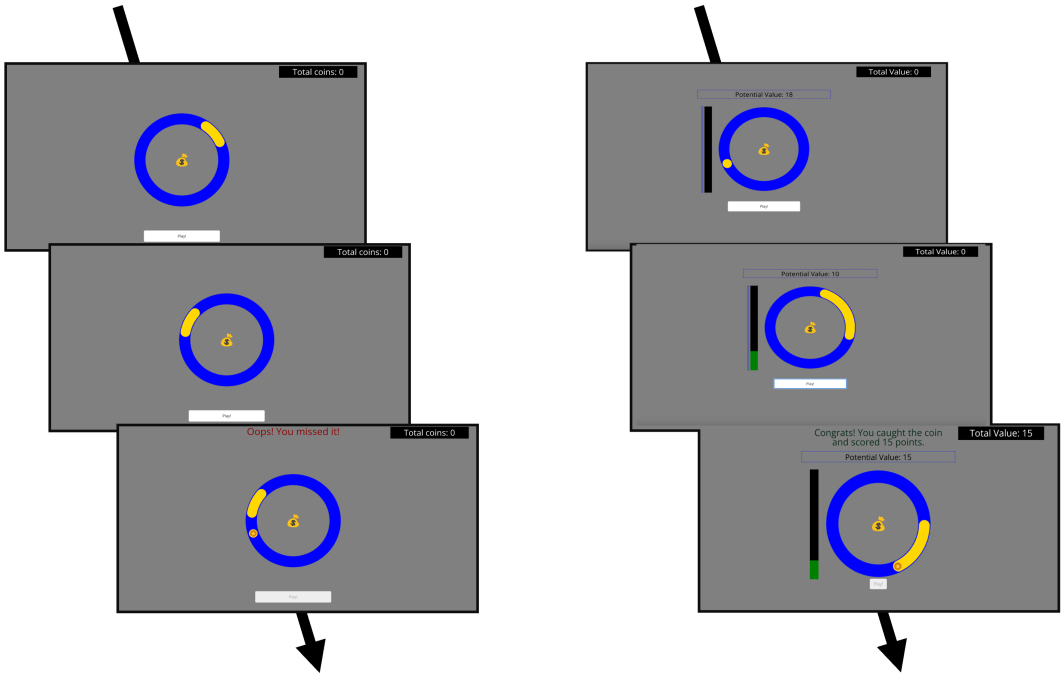
At the start of each trial, the bucket was positioned at a random location to maintain participant engagement. Once participants chose a location, they submitted their prediction by pressing the controller, which triggered the coin's release on the screen. After the participant's response, the coin was released from the center of the circle (where the bag is located; see Figure 2.1a) toward the outcome location on the blue track. The coin took 100 ms to reach the blue track, where it remained visible for 2 second.

2.2.4 Schedule Generation

The experimental design, across all experiments, was a 2x2 factorial, varying both volatility (high/low) and noise (high/low), resulting in four blocks containing 50 trials each. Block order was randomized for each participant. The coin's location for each trial was generated using a von Mises random walk ($\theta_i \sim \text{von Mises}(\mu_i, \kappa_{\text{noise}})$), with μ_i changing on each trial. The update of this parameter was determined by:

$$\mu_i = \mu_{i-1} + v_t,$$

where v_t, the step taken by the mean at each trial, was drawn from a second von Mises distribution $v_t \sim \text{von Mises}(0, \kappa_{\text{volatility}})$.



(a) **(b)**

Figure 2.1: Task Structure: Coin catching task.
(a) Example trial with a fixed bucket. Participants were instructed to place the yellow bucket at the location where they believed the coin would most likely land. After positioning the bucket, a coin was released from the centre of the circle (taking 20 ms) and landed on the blue track. If the coin did not land in the yellow bucket (as shown in this example), a red message reading "Oops! You missed it" appeared at the top of the screen, and the total coin counter (located in the top right corner of the screen) remained unchanged. The outcome was displayed for a variable duration between 2 and 7 seconds. (b) Example trial with a resizable bucket. Participants could adjust the size of the bucket using the green vertical slider. Potential values with different sizes of the bucket were highlighted in green and shown on top of the circle.

As a result, the coin's locations were influenced by both the mean (μ_i) and the concentration (κ_{noise}) of a von Mises distribution, analogous to 1/standard deviation of a standardised normal distribution. This procedure introduced both expected and unexpected uncertainty into the environment. Each time the mean changed, participants needed to form a new belief about the mean of the distribution and distinguish these changes from noise.

2.2.5 Schedule selection criteria developed in Experiments 1 – Experiment 3

The noise and volatility parameters are critical to test the hypotheses. To this end, a total of 37,921 schedules were generated for Experiments 1 through 3, using various permutations of noise and volatility levels. Specifically, 15 values were sampled from a high noise parameter range of 5 to 20 ($k_{\text{noise}} = 5$ to 20) and 25 values were sampled for the low noise parameter from a range of 20 to 45 ($k_{\text{noise}} = 20$ to 45), incrementing by one unit. For the high volatility parameter, we set the spectrum to be between 2 and 10 ($\kappa_{\text{volatility}} = 2$ to 10), while the low volatility parameter was defined within a span of 80 to 95 ($\kappa_{\text{volatility}} = 80$ to 95), also incrementing by one unit.

To design the experimental schedules, we conducted initial plotting and exploration to understand how different range of the concentration parameters ($\kappa_{\text{volatility}}$ and k_{noise}) affect the random walk. These analyses allowed us to identify parameter ranges for volatility and noise that met our criteria.

The schedules were selected based on two primary criteria:

1. For high volatility blocks, it was imperative that the random walk's mean traversed back and forth across the circle rather than in a singular direction. This ensures that participants perceive volatility instead of a unidirectional trend or bias.

2. We aimed for the maximum angular difference within high volatility blocks to approximate those within high noise blocks. To achieve this, we calculated the angular difference between the coin's landing position at time t and $t + 1$ within each block. This comparison was crucial because it ensured that participants' differentiation between volatility and noise was not solely dependent on the magnitude of change.

High Volatility Range (2 to 10)

The high volatility range was chosen to be between 2 and 10 to satisfy criteria 1 above. Initial exploration showed that setting volatility parameter $\kappa_{\text{volatility}}$ below 2

caused the mean to traverse the entire circle multiple times during a single block (Figure 2.2a). Conversely, setting the parameter above 10 resulted in minimal changes to the mean, causing the coin to land predominately on one side of the circle and reducing the desired variability (Figure 2.2a). The selection range of 2 to 10 balanced these extremes.

Low Volatility Range (80 to 95)

The low volatility range was set between 80 and 85 to ensure the mean of the random walk remained relatively stable across trials. Achieving this stability is challenging because step sizes accumulate over successive trials, which resulted in noticeable changes in the mean. Through our initial exploration, we found that setting the volatility parameter below 80 produced obvious changes in the mean (Figure 2.2b). This range is significantly smaller than that of low noise, reflecting the need to suppress cumulative shifts in the mean over time.

High Noise Range (5 to 20)

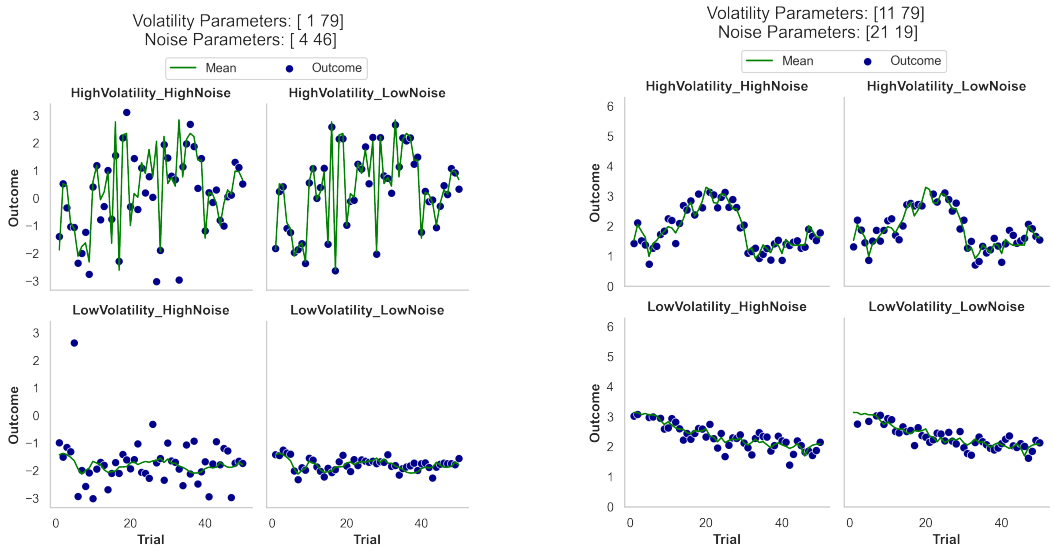
The high noise range was chosen to control variability around the mean within each trial and avoid extreme deviations, which make the schedule look like an "odd-ball" design. Values below 5 were excluded because they produced a "jump" in the coin's landing position (Figure 2.2a).

Low Noise Range (20 to 45)

The low noise range was set between 20 and 45 to produce a subtle variation around the mean. However, low noise blocks should not include landing positions directly on the mean. We thus capped the upper limit of the low noise range to 45 (Figure 2.2a).

2.2.6 Additional schedule selection criteria developed in Experiments 4 – Experiment 8

From Experiment 4 onwards, we introduced an additional constraint to schedule selection. Participants were asked to maximize the number of coins caught, rather than attributing the observed variations to their underlying sources. Therefore, we investigated whether adjusting LRs—decreasing them in high noise conditions and



(a) **(b)**

Figure 2.2: Schedule selection criteria developed in Experiments 1 – Experiment 3.
We conducted initial plotting and exploration to understand how different ranges of the concentration parameters ($\kappa_{\text{volatility}}$ and k_{noise}) affected the random walk.

increasing them in high volatility conditions, which we deemed normative—would indeed enhance the success rate of coin catching for a given schedule.

To test this, we ran simulations based on the assumption that participants initially place their bucket randomly on the first trial and maintain a constant LR throughout the block. Specifically, we used the Rescorla-Wagner learning model to generate predictions and then calculated the number of coins collected. The model updates beliefs on each trial using the following formula:

$$\text{Belief}_t = \text{Belief}_{t-1} + \text{LR} \times \text{PE}_t \tag{2.2.6.1}$$

where PE_t is the prediction error, defined as the difference between the predicted and observed outcome on trial t. Since the difference was computed on a circular scale, the minimum distance between the prediction and the outcome from either direction was taken:

$$\text{PE}_t = \min\left(\left|\text{Outcome}_t - \text{Prediction}_t\right|,\ 2\pi - \left|\text{Outcome}_t - \text{Prediction}_t\right|\right) \tag{2.2.6.2}$$

2. Designing a Behavioral and Computational Framework to Dissociate Volatility and Noise in Learning

To systematically explore how different learning rates influenced performance, we varied the LR incrementally from 0 to 1 in steps of 0.01, and simulated outcomes for each value. The updated belief, Belief_t, was used as the prediction at that trial. We then evaluated whether Belief_t resulted in a successful coin catch. For Experiment 4 through 8, the bucket size was fixed at 30 degrees, meaning that if the coin landed within the range $\text{Belief}_t \pm 15$ degrees, it was recorded as a successful catch. This allowed us to generate a distribution of coin catches for each learning rate across all blocks, which enables us to determine how performance varied with different LRs.

To evaluate how closely the optimal LRs for each block aligned with normative behaviour, we introduced a metric based on weighted average of the LRs. The weighted average was calculated as:

$$\text{Weighted Average LR} = \frac{\sum_i (\text{LR}_i \times \text{Points}_i)}{\sum_i \text{Points}_i} \tag{2.2.6.3}$$

where LR_i represents each possible learning rate, and Points_i is the number of points earned at that specific LR. This metric provides an insight into the typical effectiveness of a learning rate when considering how often it leads to a successful prediction.

We then computed the expected volatility effect by calculating the difference between the weighted average LR in high volatility blocks versus low volatility blocks. Similarly, the expected noise effect was obtained by calculating the difference between the weighted LR in low noise blocks versus high noise blocks. The larger these effects, the more suitable the schedule for testing whether human learners adapt normatively to volatility and noise.

In addition, we calculated the weighted dispersion of LRs to assess the variability of LRs around the weighted average. This measure was important because participants were instructed to focus solely on maximizing the number of coins caught, rather than demonstrating their ability to adapt to volatility and noise as rational agents. A smaller dispersion indicates that a narrow range of LRs leads to higher number of coins caught, encouraging participants to adopt an LR within that optional range. This consistency in behaviours makes observing the

hypothesized volatility and noise effects easier. Conversely, a larger dispersion suggests that a wide range of LRs result in similar outcomes, allowing participants to select any LR within that range, which could obscure the hypothesized effects and make them harder to detect.

Weight dispersion was calculated as:

$$\text{Weighted Dispersion} = \sqrt{\frac{\sum_i \left(\text{Points}_i \times (\text{LR}_i - \text{Weighted Average LR})^2\right)}{\sum_i \text{Points}_i}} \quad (2.2.6.4)$$

By combining the metrics of expected volatility and noise effect, obtained via weighted average, and weighted dispersion, we were able to systematically assess how participants' performance aligned with normative behaviour and how volatility and noise affected learning.

Furthermore, we examined how noise parameters could affect participants' behaviour throughout the experiment. As discussed in the Introduction to this chapter, several studies have tested whether humans adapt to noise in a normative manner when the noise level is low (Diederen & Schultz, 2015; Nassar et al., 2010). Building on this, we aimed to identify the threshold at which participants begin to perceive environmental changes as driven by volatility rather than noise.

In Experiments 4 through 7, we systematically manipulated the noise parameters to investigate how participants adapt to varying levels of noise while keeping the volatility parameters the same. In particular, we incrementally increased the absolute average noise across experiments by adjusting the high and low noise parameters:

$$\frac{\text{high noise parameter} + \text{low noise parameter}}{2} \quad (2.2.6.5)$$

while maintaining the difference between them constant:

$$|\text{high noise parameter} - \text{low noise parameter}| = 36 \quad (2.2.6.6)$$

This approach ensured that the relative contrast between high and low noise blocks remained consistent, minimizing the possibility that observed changes in behaviours were driven by shifts in in the relative difficulty of distinguishing between the two noise levels. By gradually raising the average noise, we aimed to identify the

threshold at which participants shift from interpreting changes in the environment as noise to perceiving them as driven by volatility.

For Experiment 8, we replicated the schedule of Experiment 2 with the modification of using a fixed bucket size. This iteration was conducted as a preparatory step for Project 2, which aims to investigate the physiology of learning under environmental uncertainty with an eye-tracker. The decision to model Experiment 8 after Experiment 2 was informed by the Experiment 2's significant volatility effect (details in the Results section).

The overarching goal of this process was to identify task parameters that enabled participants to detect and respond appropriately to changes in both volatility and noise. At the same time, we aimed to avoid making the task overly simplistic, ensuring it could reveal meaningful individual differences in performance. This was particularly crucial for our planned in-person studies, where we investigated these individual differences in adaptive behaviour, specifically by selecting participants based on their depressive symptoms (Table 2.3).

Experiment	High Noise Parameter	Low Noise Parameter	High Volatility Parameter	Low Volatility Parameter	Bucket
Experiment 1	3	35.00	8	85	Resizable
Experiment 2	9	42.00	5	84	Resizable
Experiment 3	8	38.00	3	88	Resizable
Experiment 4	9	45.00	3	94	Fixed
Experiment 5	8.2	41.08	3	94	Fixed
Experiment 6	8.451	42.238	3	94	Fixed
Experiment 7	8.696	43.462	3	94	Fixed
Experiment 8	9	42.00	5	84	Fixed

Table 2.3: Experimental Parameters for Noise and Volatility Manipulation Across Experiments.
This table outlines the key parameter values used in eight experiments designed to investigate the effects of noise and volatility on participants' behaviour.

2.2.7 Working memory capacity measure

Previous research has demonstrated that both working memory and reinforcement learning contribute to instrumental learning (Collins & Frank, 2012). Additionally,

higher working memory capacity has been shown to protect model-based reinforcement learning from the effects of stress (Otto et al., 2013). To assess participants' working memory capacity, we used the automated Operation Span Task (OSPAN). Participants memorized sequences of letters (length 4 to 7) while solving interleaved arithmetic problems. Each arithmetic problem was time limited to 6 seconds to impose concurrent processing demands. We did not collect stress measures during OSPAN, and the task was not intended as a stress induction but as an index of working memory capacity. OSPAN scores were calculated by summing the number of letters selected for all correctly recalled sets.

2.2.8 Questionnaire

Although the sample in this study was not selected or pre-screened based on anxiety or depressive symptoms, we were still interested in examining whether participants' symptoms of depression and anxiety were linked to their behaviours in the task. To assess participants' symptoms of depression and anxiety, we used four questionnaires. We picked four questionnaires to understand participants' symptoms of depression and anxiety. The Trait Anxiety Inventory (STAI-T) was used to assess enduring quality of trait anxiety. To measure symptoms of depression, we used the Centre for Epidemiologic Studies Depression Scale (CES-D), which emphasize the frequency and intensity of depressive episodes within the past two weeks. The Work and Social Adjustment Scale (WSAS) focuses on the functional aspects of an individual's life, explicitly assessing the impairment in daily functioning attributable to psychological distress. Lastly, the Temporal Experience of Pleasure Scale (TEPS) is deployed to examine hedonic capacity, that is, the ability to experience pleasure. TEPS distinguishes between anticipatory pleasure, the enjoyment derived from the expectation of a rewarding experience, and consummatory pleasure, the satisfaction felt during the experience itself.

2.2.9 Model-free Analysis

All angles were measured in radians, ranging from 0 to 2π. The position of the bucket represents participants' best prediction of where the coin was most likely to land. Specifically, we took the midpoint of the bucket as participants' predictions of each trial (Figure 2.3). We then calculated the signed prediction error (PE) as the difference between where participants predicted the coin would land (midpoint of the bucket) and the actual position of the coin (outcome). The signed prediction error for trial t was calculated as:

$$\mathrm{PE}_t = \mathrm{Outcome}_t - \mathrm{Prediction}_t \tag{2.2.9.1}$$

Since the difference was computed on a circular scale, the minimum distance between the prediction and the outcome from either direction was taken:

$$\mathrm{PE}_t = \min\left(\left|\mathrm{Outcome}_t - \mathrm{Prediction}_t\right|,\ 2\pi - \left|\mathrm{Outcome}_t - \mathrm{Prediction}_t\right|\right) \tag{2.2.9.2}$$

Similarly, the belief update (BU) was defined as the change in participants' prediction from trial t to trial $t+1$, calculated as:

$$\mathrm{BU}_t = \min\left(\left|\mathrm{Prediction}_{t+1} - \mathrm{Prediction}_t\right|,\ 2\pi - \left|\mathrm{Prediction}_{t+1} - \mathrm{Prediction}_t\right|\right) \tag{2.2.9.3}$$

Normally, BU is calculated by finding the shortest path between the current and previous predictions. Yet, when BU and PE pointed in opposite directions, and the subsequent prediction was significantly different from the actual outcome (more than half a circle) but still within a reasonable range of the current prediction (within half a circle), BU was adjusted to match the direction of the PE. In other words, when BU_t and PE_t are in opposite direction and $|\mathrm{Prediction}_{t+1} - \mathrm{Outcome}_{t+1}| > 0.5\pi$ and $|\mathrm{Prediction}_{t+1} - \mathrm{Prediction}_t| < 0.5\pi$, BU_t was assigned to be in the same direction as PE_t.

This was a crucial adjustment for cases where a moderate PE was accompanied by a significantly larger BU in a different direction, suggesting that participants were not actually adjusting their predictions as dramatically as the raw data

suggested (i.e., making the next prediction almost half a circle away from the previous prediction). Instead, it seemed that they were anticipating the next position of the coin beyond the feedback indicated by PE.

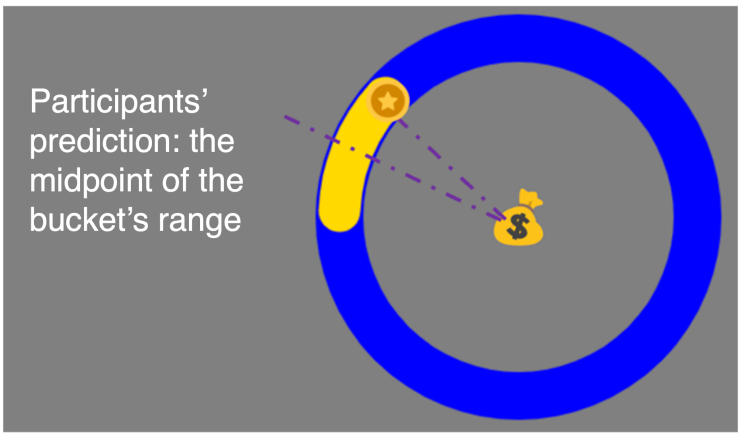


Figure 2.3: Outcome and prediction.
We set the midpoint of the bucket as the prediction made in a trial. The outcome is measured by the angle where the coin landed.

Estimating changes in learning rate

The primary outcome variable was learning rate in each block and learning rate adjustment between high volatile and low volatile blocks and learning rate adjustment between low noise and high noise blocks. We estimated learning rate using three distinct methods: (1) trial-wise learning rate; (2) a linear regression model; and (3) fitting a simple learning rate model. Each of these three methods has its own pros and cons, which are discussed below.

Trial-wise Learning Rate

For learning rate at each trial, we calculated the ratio between BU_t and PE_t. In particular,

$$\text{Trial-wise LR} = \frac{\mathrm{BU}_t}{\mathrm{PE}_t} \tag{2.2.9.4}$$

For block-wise LR, we averaged all trial-wise learning rates for each block and each participant.

We subsequently evaluated the distribution of LRs each block across eight experiments. A notable proportion of extremely large and negative outliers (Figure 2.4), which can be partly attributed to the study design involving a continuous variable. This design occasionally produced very small prediction errors (PE), resulting in disproportionately large learning rates (LRs). Furthermore, negative learning rates arise in instances where PE and belief updates (BU) point in opposite directions. Despite modifications to the BU calculation to mitigate this issue, such discrepancies persist, highlighting the complexity of accurately modeling belief updates in these scenarios.

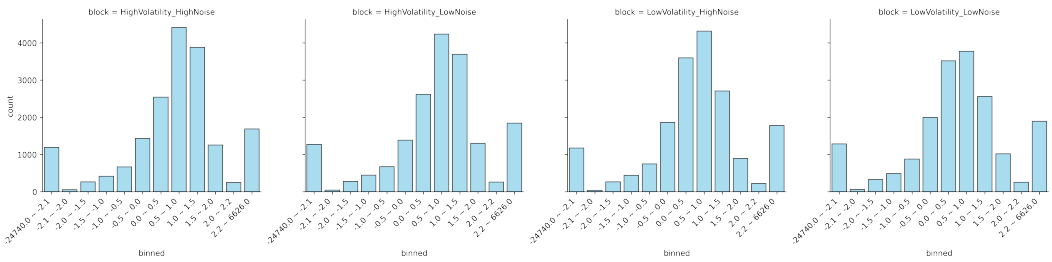


Figure 2.4: Histogram of trial-wise learning rates across eight experiments. This figure illustrates the distribution of trial-wise learning rates for all eight experiments, grouped by block. Each subplot represents a separate block, with the x-axis showing binned learning rates and the y-axis indidcating the frequency of occurrences within each bin. Learning rates between -2 and 2 were divided into bins of width 0.5, while extreme values of -2.1 and above 2.2 were grouped into separate bins to prevent distortion in visualization.

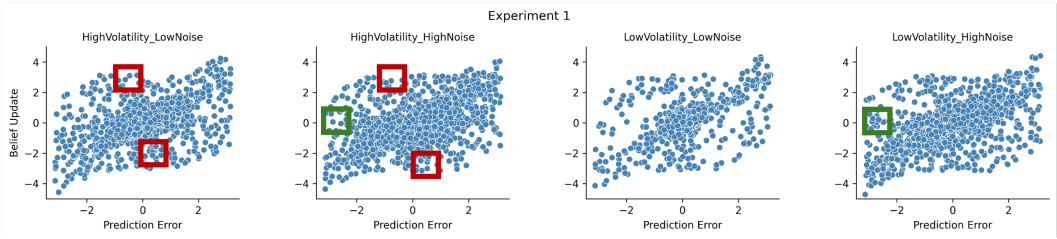
Linear Regression

We conducted a linear regression for each participant within each experimental block. In this regression model, BU was set as the dependent variable, with the prediction error PE as the independent variable. The regression included an intercept term. The coefficient of prediction errors was obtained from each regression and served as the computed estimate of the learning rate (LR) for the block in question, providing a block-specific measure of how participants' beliefs were adjusted in response to errors in their predictions:

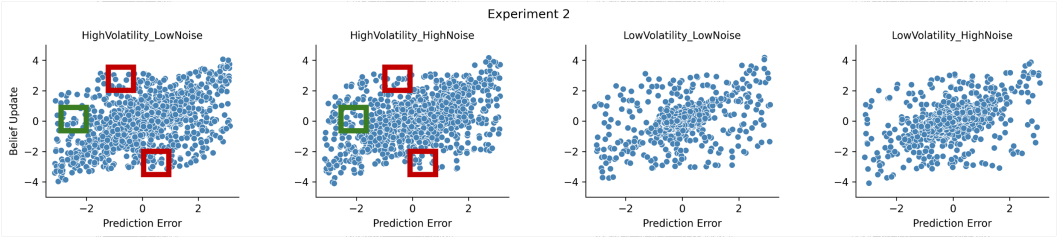
$$\mathrm{BU}_t = \beta_0 + \beta_1 \times \mathrm{PE}_t \tag{2.2.9.5}$$

We first checked linearity by inspecting scatterplots between PE and BU across all eight experiments (See Figure 2.5). In these scatterplots, the x-axis represents PE, and the y-axis represents BU, dividing the plot into four quadrants: (1) top-right (positive PE, positive BU), (2) top-left (negative PE, positive BU), (3) bottom-left (negative PE, negative BU), and (4) bottom-right (positive PE, negative BU). While BU and PE appeared linear in many cases, certain trials revealed measurement issues, particularly in blocks with a high proportion of data falling into the second quadrant (negative PE, positive BU) and fourth quadrant (positive PE, negative BU) (e.g., Experiment 1 High Volatility Low Noise block). These issues arise from two key scenarios: (1) trials with a very small PE in the opposite direction of a very large BU (e.g., red boxes in Figure 2.5); (2) trials with extreme PE values on the x-axis but near-zero BU values on the y-axis (e.g., green boxes in Figure 2.5).

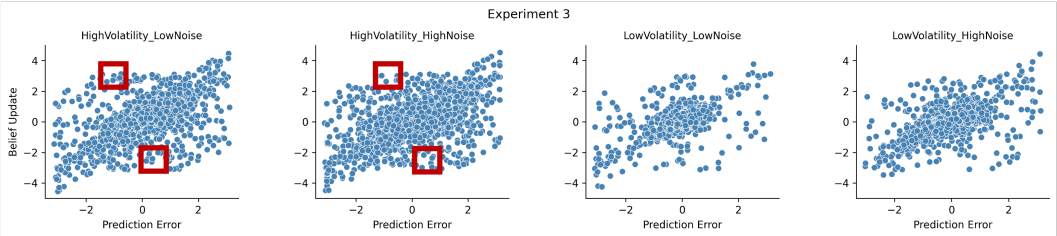
Despite modifications to how BU was calculated, the adjustments did not completely resolve the issue. Specifically, the decision to use half a circle (π) as the threshold for correcting BU was somewhat arbitrary, raising questions about whether this approach genuinely reflects participants' intentions. It remains uncertain whether participants were making predictive jumps of this nature or if these patterns arose as artifacts of the circular task structure and the methods used to calculate the minimum distance.However, in finidng the best fit of the line across all trials in a block, the regression approach is able to account for these outliers to some extent, mitigating their influence on the overall relationship between BU and PE.


Experiment 1
HighVolatility_LowNoise
HighVolatility_HighNoise
LowVolatility_LowNoise
LowVolatility_HighNoise
Belief Update
Prediction Error

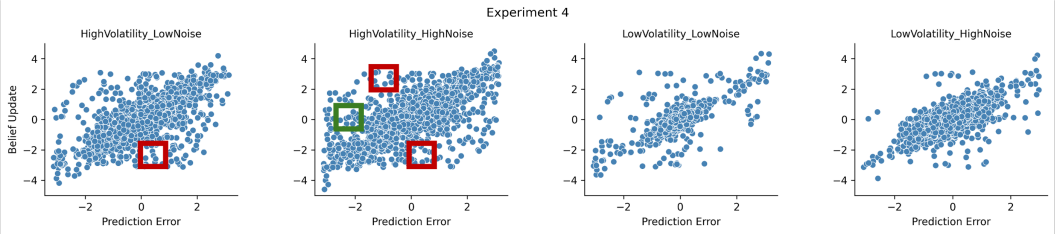
(a)


Experiment 2
HighVolatility_LowNoise
HighVolatility_HighNoise
LowVolatility_LowNoise
LowVolatility_HighNoise
Belief Update
Prediction Error

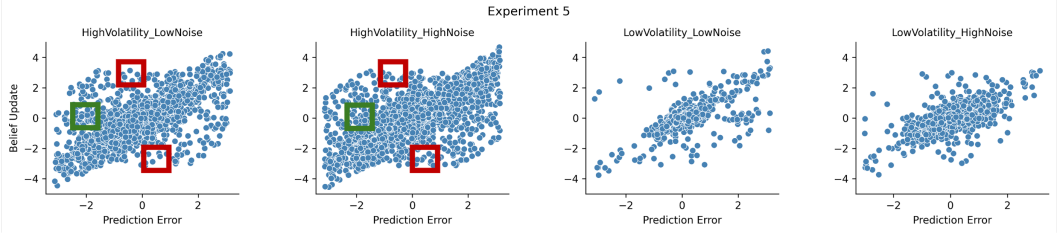
(b)


Experiment 3
HighVolatility_LowNoise
HighVolatility_HighNoise
LowVolatility_LowNoise
LowVolatility_HighNoise
Belief Update
Prediction Error

(c)


Experiment 4
HighVolatility_LowNoise
HighVolatility_HighNoise
LowVolatility_LowNoise
LowVolatility_HighNoise
Belief Update
Prediction Error

(d)


Experiment 5
HighVolatility_LowNoise
HighVolatility_HighNoise
LowVolatility_LowNoise
LowVolatility_HighNoise
Belief Update
Prediction Error

(e)

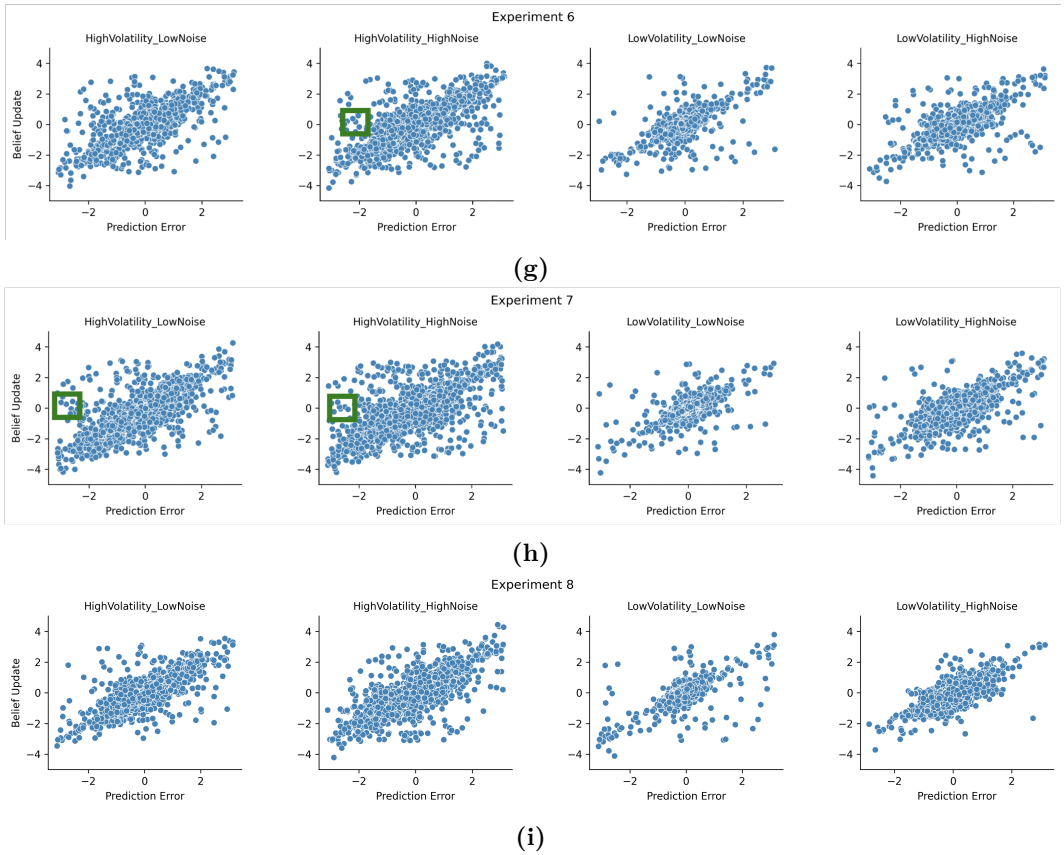


Figure 2.5: Scatterplots showing the relationship between PE and BU across experiments.
We examined the relationship between prediction error (PE) and belief update (BU) by block for each experiment separately. Each row (a-i) shows one experiment; each column shows one block. Red boxes highlight trials where a very small PE, opposite in sign to a very large BU, produces noticeable deviations. The green box highlights trials with extreme PE on the x-axis but near-zero BU on the y-axis. These patterns reveal challenges in computing PE and BU for circular variables; despite adjustments to the BU calculation, these issues persist.

2.2.10 Model-based Analysis

To see how learning rate changes under volatility and noise, we estimated the changes in learning rate between volatile and noisy blocks by fitting a modified Rescorla-Wagner learning model to participants' choice data. The belief update process for trial t was defined as follows:

$$\text{Belief}_{t+1} = \text{Belief}_t + \text{LR} \times \text{PE}_t \tag{2.2.10.1}$$

The von Mises probability density function transforms these belief estimates into probabilities by calculating the likelihood of observing participants' choice data, conditional upon their updated belief and the precision with which they are expected to place the bucket.

$$\begin{aligned} \text{Prob}(\text{choice} \mid \text{LR}, \text{prec}) &= \text{Prob}(\text{choice} \mid \text{belief}, \text{prec}) \\ &= \frac{\exp(\text{precision} \cdot \cos(\text{choice} - \text{belief}))}{2\pi I_0(\text{precision})} \end{aligned} \tag{2.2.10.2}$$

Here *prec* is a precision parameter and controls how accurately participants place the bucket where they believe the coin is most likely to land Belief_{t+1}.

Parameter Estimation

In total, the model has 2 free parameters (*LR* and *prec*). These parameters were estimated separately for each block and for each participant. LR was represented using 100 points spaced equally from logit (0.01) to logit (1) and 99 points for prec ranging from log (0.1) to log (100). For this model, every parameter combination was tested, resulting in 9900 unique combinations.

Each parameter combination was applied to simulate model behaviour on a trial-by-trial basis. The likelihood across all trials was calculated by multiplying the trial-wise likelihoods of participants' observed choices. For each parameter, the best estimate was determined by marginalizing over all other parameters to calculate the expected value of its probability distribution. This was achieved by summing the discrete parameter values weighted by their marginalized probabilities (Behrens et al., 2007; Browning et al., 2015). All trials were used when estimating the values of LR and prec.

We subsequently compared how well our model captures participant behavior. We overlayed the choice behavior of participants with model behavior. The model output was produced by running the behavioral model once for each participant using individually fitted parameter values (Figure 2.6).

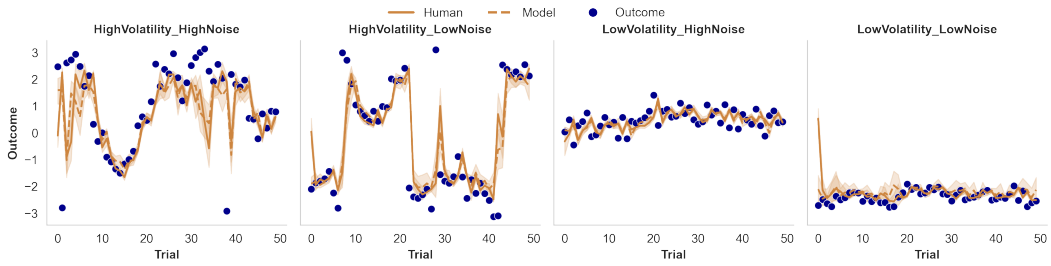
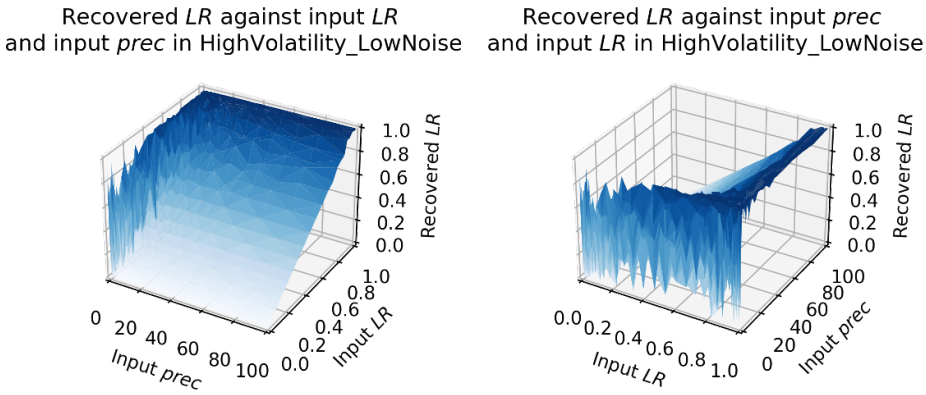


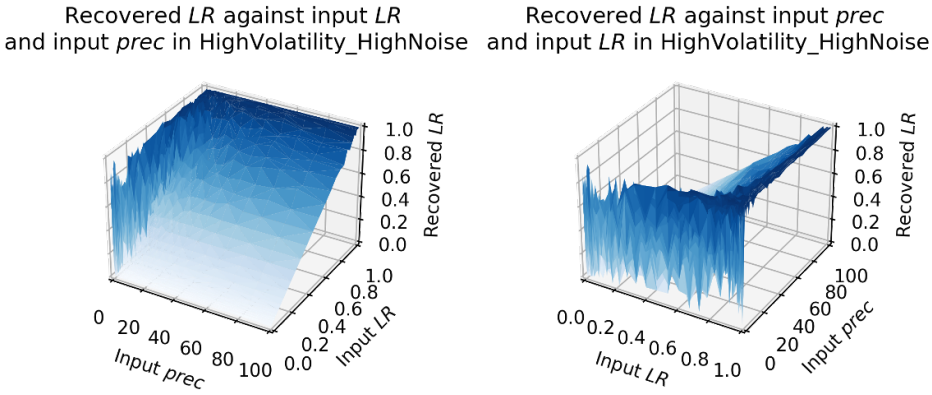
Figure 2.6: Comparison of Model-Simulated and Participant Behavioural Data.
The plot shown here overlays the choice behaviour of participants (brown solid line) with the behaviour generated using model-fitted parameters (brown dashed line) and the output (blue dots). The model output was generated by running the Rescorla-Wagner learning model for each participant using their individually fitted parameter values on the task trial sequences. As shown, the model accurately reflects participant behaviours during the task.

Parameter Recovery

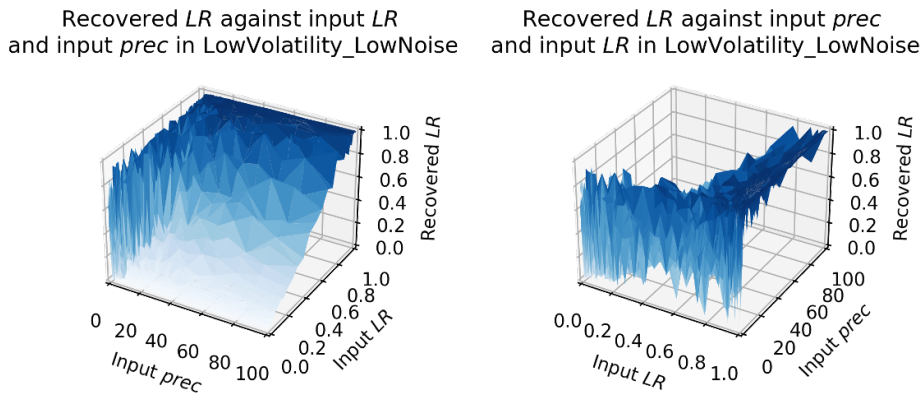
To evaluate whether our parameter estimation procedure reliably reproduces parameter values, we first defined the parameter ranges for the modified Rescorla-Wagner model. The learning rate was divided into 50 equally spaced values, and the precision parameter was divided into 49 equally spaced values, covering their entire respective ranges. Using these parameters, predictions were simulated for 2,450 combinations per block, resulting in a total of 9,800 simulations across all blocks. These simulated choices were then used to assess the reliability of parameter recovery. The Rescorla-Wagner model was then fitted to the simulated choices, and the resulting parameter estimates were compared to the input parameters used for simulating those choices. The results indicate that the learning rate (LR) parameter was reasonably well recovered, closely matching its input values. However, under low input precision ($prec$), the input LR did not correlate with the recovered LR (see Figure 2.7).



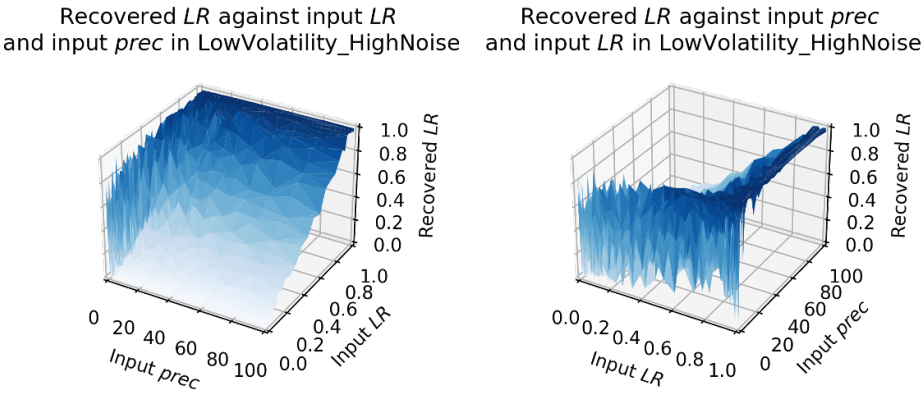
(a)Recoverability of LR in High Volatility Low Noise
The recovered LR (z-axis) is assessed as a function of the input LR and input $prec$. Right panel: x = input $prec$, y = input LR Left panel: x = input LR, y = input $prec$ Recovery of LR was reasonably good. We observed a clear linear relationship between input LR and recovered LR across all values of $prec$, except when $prec$ was low. In contrast, there was no relationship between recovered LR and input $prec$.



(b)Recoverability of LR in High Volatility High Noise



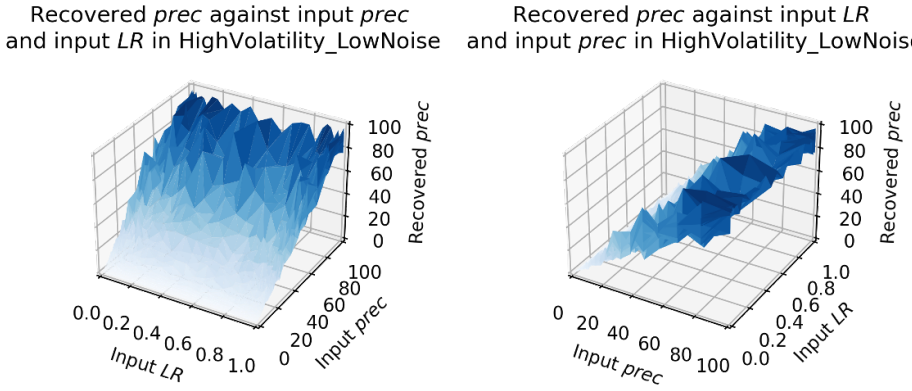
(c)Recoverability of LR in Low Volatility Low Noise



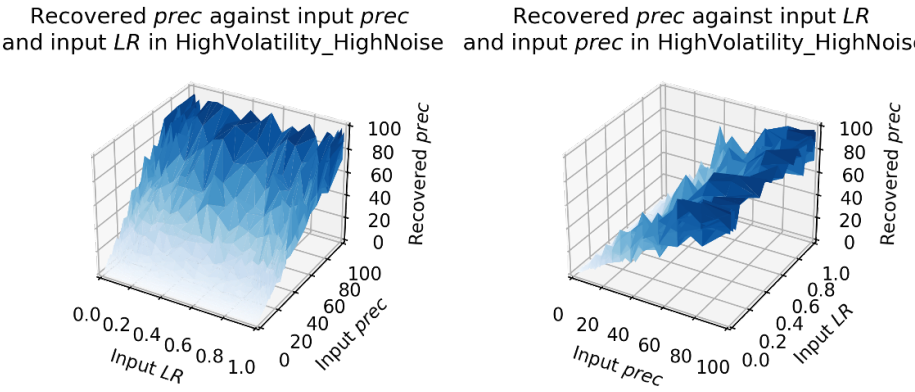
(d) Recoverability of LR in Low Volatility High Noise

Figure 2.7: Parameter recovery for *LR*.
Parameter recovery for the learning rate (LR) was conducted separately for each block. **(a–d)** show recovery surfaces by block: **(a)** High Volatility–Low Noise, **(b)** High Volatility–High Noise, **(c)** Low Volatility–Low Noise, and **(d)** Low Volatility–High Noise. In each panel, the recovered LR (z-axis) is plotted as a function of input LR and input precision (prec); the two subplots swap the x–y axes for readability. For example, in the left subplot of **(a)** (x-axis = prec, y-axis = input LR), recovered LR increases roughly linearly with input LR across most precision values, except when precision is very low (prec less than 10). In the right subplot of **(a)** (x-axis = input LR, y-axis = prec), recovered LR shows no systematic dependence on precision. Across all blocks, recovered LR increases approximately linearly with input LR, with weaker recovery when precision is very low (prec less than 10). No systematic dependence is observed between recovered LR and precision alone. Color shading from light to dark blue encodes low to high recovered LR.

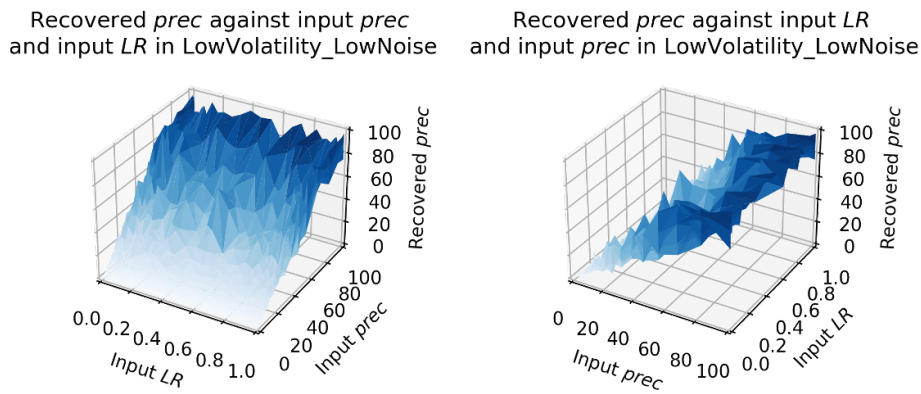
Recovery of precision is close to input values and is independent from the values of input LR (see Figure 2.8). In particular, across all volatility–noise conditions, recovered precision increases approximately linearly with the input precision parameter, demonstrating reliable parameter recovery. By contrast, there is no systematic dependence of recovered precision on the input learning rate, as indicated by the flat recovery surfaces along the LR dimension. This pattern holds consistently across blocks, confirming that our estimation procedure can successfully recover decision precision irrespective of environmental condition or input LR.



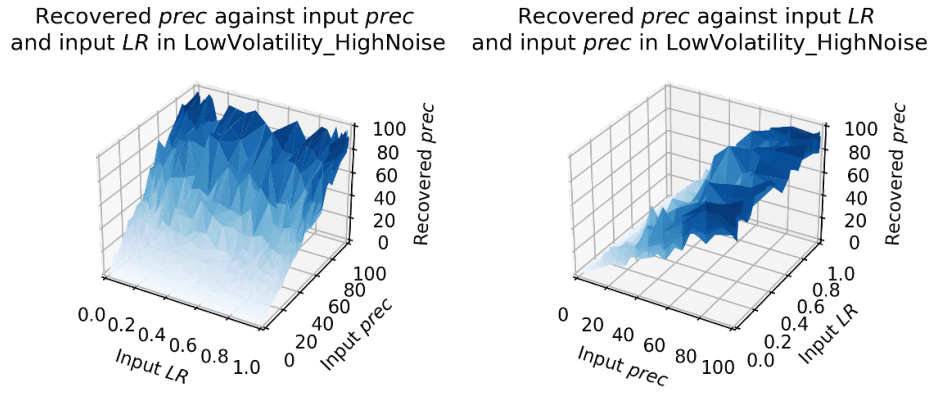
(a)Recoverability of *prec* in High Volatility Low Noise
The recovered *prec* (z-axis) is assessed as a function of the input *prec* and input *LR*. Right panel: x = input *LR*, y = input *prec* Left panel: x = input *prec*, y = input *LR* Recovered precision tracks input precision closely (approximately linear), with no observable relationship between recovered precision and input LR. Color shading from light to dark blue encodes low to high recovered precision.



(b)Recoverability of *prec* in High Volatility High Noise



(c)Recoverability of *prec* in Low Volatility Low Noise



(d) Recoverability of *prec* in Low Volatility High Noise

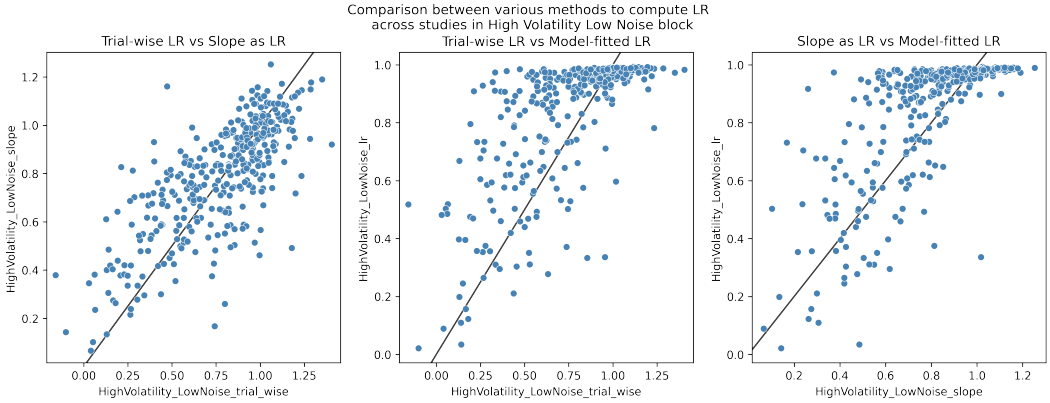
Figure 2.8: Parameter recovery for *prec*.
Parameter recovery for decision precision (prec) was conducted separately for each block. **(a–d)** show recovery surfaces by block: **(a)** High Volatility–Low Noise, **(b)** High Volatility–High Noise, **(c)** Low Volatility–Low Noise, and **(d)** Low Volatility–High Noise. In each panel, the recovered precision (z-axis) is plotted as a function of input precision (*prec*) and input learning rate (*LR*); the two subplots swap the x–y axes for readability. For example, in the left subplot of **(a)** (x-axis = inpurt LR, y-axis = input prec), recovered precision increases approximately linearly with input prec across LR values. In the right subplot (x-axis = input prec, y-axis = inpurt LR), the surface is essentially flat along the x-axis, indicating no systematic dependence of recovered precision on LR. Across all four blocks, recovered precision tracks input precision closely (approximately linear), with no observable relationship between recovered precision and input LR. Color shading from light to dark blue encodes low to high recovered precision.

2.2.11 Comparison between Model-based and Model-free Analysis

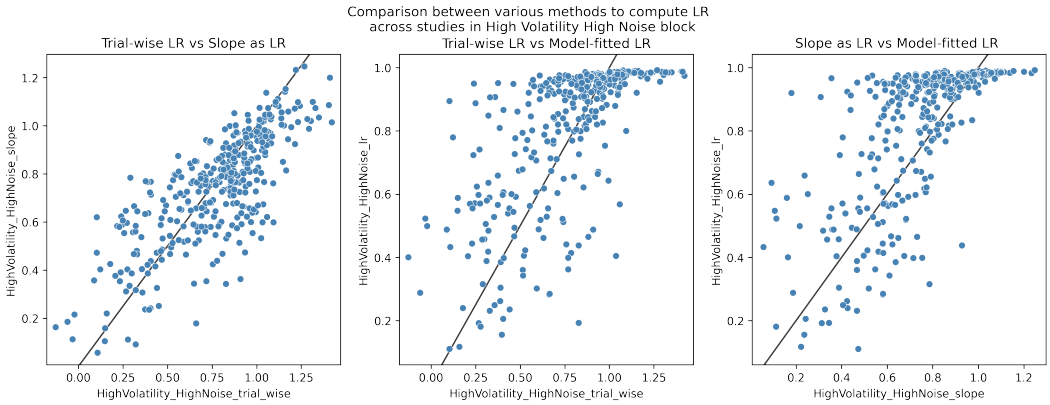
The analysis conducted here aims to elucidate the most robust and accurate method for computing learning rates (*LRs*) by comparing three distinct approaches: median of trial-wise *LRs* by participants by block, extracting slopes as *LRs*, and employing model-fitted *LRs*. The comparison was based on data aggregated across all blocks from eight experiments. The median of trial-wise *LRs* was chosen to mitigate the influence of extreme values, which could bias the results if an average were used instead.

We observed that trial-wise LRs demonstrated the strongest correlation with the slopes interpreted as *LRs* across all blocks. In contrast, model-fitted LRs did

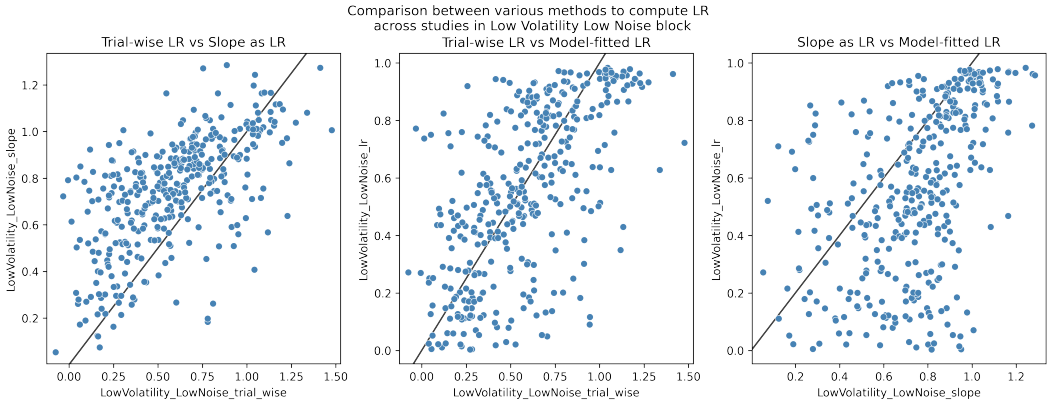
not exhibit a strong correlation with either trial-wise LRs or slopes as LRs mostly as result of having more variability (Figure 2.9).



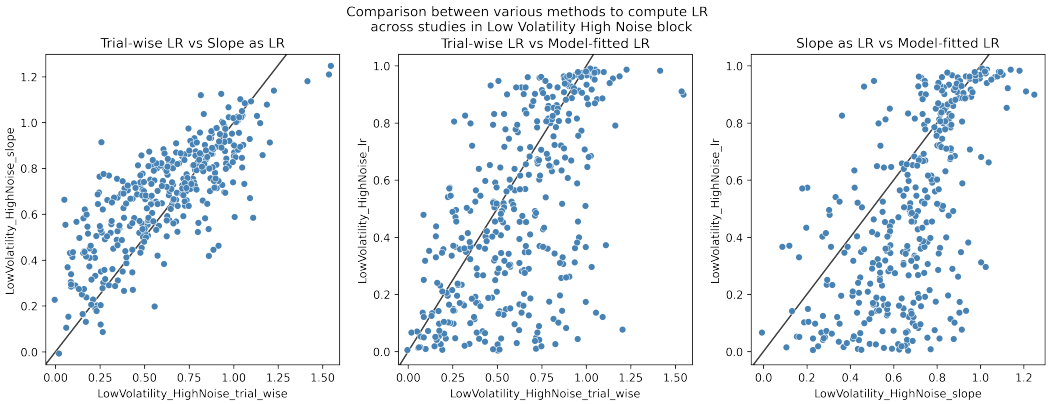
(a) High Volatility Low Noise



(b) High Volatility High Noise



(c) Low Volatility Low Noise



(d) Low Volatility High Noise

Figure 2.9: Comparison of *LRs*. These scatterplots visualize the relationships and variance between different learning rate computation methods, highlighting the distinct characteristics of model-fitted learning rates. Each point represents a participant's learning rate across experimental conditions. The $y = x$ line indicates perfect correlation between the two learning rate methods being compared.

The decision to use model-fitted learning rates (LRs) in this study is based on several key considerations. Trial-wise LRs often include extreme values that bias block-wise estimates when averaged, while using the median reduces such bias but disregards variability in the data. Slopes as LRs, estimated via linear regression, are also influenced by outliers, such as trials with very small prediction errors (PE) in the opposite direction of a large belief update (BU) or trials with extreme PE values paired with near-zero BUs. Both trial-wise LRs and slopes as LRs rely on predefined assumptions about the direction of belief updates. In contrast, model-fitted LRs are advantageous because they avoid averaging or median-based distortions caused by outliers and incorporate the precision parameter, a crucial aspect of this learning task that explicitly captures the randomness of choices (Table 2.4).

Given that the model-fitted LRs present a more holistic view of learning by integrating key task parameters and the input variables are raw, unmodified data collected from the task (i.e., prediction and outcome), we decided that model-fitted LRs would provide the most accurate and reliable measure for statistical analysis in this study.

Methods	Input Data	Pros	Cons
Trial-wise LR	PE and BU	Simple and straightforward	Outliers bias the average LR per block
Slope as LR	PE and BU	Simple and straightforward; less affected by outliers	Non-linearity and modification used to present BU
Model-fitted LR	Prediction (choice data) Outcome (schedule)	Avoids using PE and BU; relatively simple; accounts for precision when placing the bucket	Assumes LR is static within a block

Table 2.4: Pros and cons of each analysis method of LR.
Three methods were developed to estimate learning rates (LRs): trial-wise LRs, slope-based LRs, and model-fitted LRs. Each method is evaluated based on its input data, strengths, and limitations. Trial-wise LRs are simple but sensitive to outliers. Slope-based LRs mitigate this but introduce non-linearity. Model-fitted LRs rely on raw predictions and outcomes and incorporate task precision, but assume static LRs within blocks.

2.2.12 Statistical Analysis

For the statistical analysis, a logit transformation was applied on the model-based learning rates to ensure comparability and interpretability across methods. Unlike trial-wise learning rates and learning rates estimated from linear regression slopes, which are not constrained and can exceed 1 and go below 0, the model-derived methods only fits learning rate values within the range of 0 to 1. The logit transformation is particularly suitable as it effectively stretches the scale around lower learning rates (e.g., between 0 and 0.1), enhancing the discriminability of small but significant variations in learning. Simultaneously, it compresses the differences in the higher range, such as between 0.5 and 0.6. Additionally, we applied log transformation to the precision to approximate a normal distribution.

We conducted repeated measures Analysis of Variance (ANOVA) to evaluate the learning rate estimates for each experimental block as the dependent variable. The model included within-subject factors, specifically volatility (categorized as high vs. low) and noise (categorized as low vs. high), to discern their respective impacts on the learning rate.

We also conducted the same ANOVA analysis for the two model-free approaches of estimating LRs. To estimate LR in a block with trial-wise LRs, we took the median as opposed to mean to avoid extreme values biasing the results.

2.3 Results

We concluded that model-fitted *LR* provide the most robust analytical approach. Nevertheless, for completeness, we report results from all three analytical approaches. Summary plots of learning rate adaptation across experiments for each method can be found in Appendix A (Figures A.1, A.2, and A.3).

2.3.1 Normative adaptation of learning rate to volatility

Across Experiments 1 through 8, participants consistently demonstrated normative adaptation of their LRs in response to volatility (Table 2.5). In Experiment 1, there was a significant main effect of volatility ($F(1, 46) = 31.187$, $p < 0.001$), as hypothesized. This effect was replicated in Experiment 2 – Experiment 8.

To further examine the volatility effect, we conducted t-tests to determine if participants increased their LRs in volatile environments under both high noise and low noise conditions. In Experiment 1, the volatility effect was significant in the low noise condition ($t(46) = 5.720$, $p < 0.001$), but not under high noise ($t(46) = 1.570$, $p = 0.060$). In Experiment 2, participants significantly increased their LRs in both high noise ($t(50) = 9.560$, $p < 0.001$) and low noise ($t(50) = 7.590$, $p < 0.001$). Similarly, significant increases were observed in Experiment 3 (high noise: $t(46) = 7.160$, $p < 0.001$; low noise: $t(46) = 7.640$, $p < 0.001$), Experiment 4 (high noise: $t(45) = 12.110$, $p < 0.001$; low noise: $t(45) = 7.290$, $p < 0.001$), Experiment 5 (high noise: $t(41) = 8.030$, $p < 0.001$; low noise: $t(41) = 8.260$, $p < 0.001$), Experiment 6 (high noise: $t(42) = 7.430$, $p < 0.001$; low noise: $t(42) = 10.920$, $p < 0.001$), Experiment 7 (high noise: $t(43) = 9.080$, $p < 0.001$; low noise: $t(43) = 10.290$, $p < 0.001$), and Experiment 8 (high noise: $t(48) = 9.730$, $p < 0.001$; low noise: $t(48) = 10.470$, $p < 0.001$).

There were also consistent effects on median trial-wise learning rate across experiments, while the effect of volatility on learning rate estimated with slope from linear regression was less consistent.

Experiment	R-W model	Trial-wise LR	Slope as LR
Experiment 1	$F(1,46) = 31.187$ $p < 0.001$	$F(1,46) = 10.418$ $p = 0.002$	$F(1,46) = 0.845$ $p = 0.363$
Experiment 2	$F(1,50) = 126.226$ $p < 0.001$	$F(1,50) = 11.839$ $p = 0.001$	$F(1,50) = 3.740$ $p = 0.059$
Experiment 3	$F(1,46) = 87.577$ $p < 0.001$	$F(1,46) = 44.405$ $p < 0.001$	$F(1,46) = 8.569$ $p = 0.005$
Experiment 4	$F(1,45) = 173.239$ $p < 0.001$	$F(1,45) = 34.128$ $p < 0.001$	$F(1,45) = 3.512$ $p = 0.067$
Experiment 5	$F(1,41) = 115.293$ $p < 0.001$	$F(1,41) = 49.674$ $p < 0.001$	$F(1,41) = 0.130$ $p = 0.720$
Experiment 6	$F(1,42) = 153.591$ $p < 0.001$	$F(1,42) = 22.081$ $p < 0.001$	$F(1,42) = 6.950$ $p = 0.012$
Experiment 7	$F(1,43) = 164.260$ $p < 0.001$	$F(1,43) = 38.909$ $p < 0.001$	$F(1,43) = 3.838$ $p = 0.057$
Experiment 8	$F(1,48) = 180.673$ $p < 0.001$	$F(1,48) = 31.518$ $p < 0.001$	$F(1,48) = 17.601$ $p < 0.001$

Table 2.5: Volatility effect across experiments and analysis methods.
This table summarizes the main effects of volatility on LRs across eight experiments, analysed using three methods: Rescorla-Wagner (R-W) model, median trial-wise LRs, and slope-based LRs. Participants consistently demonstrated a significant volatility effect in learning rates, as estimated using the R-W model across all experiments ($p < 0.001$). The volatility effect was also evident in trial-wise LRs in most experiments ($p < 0.002$). In contrast, the slope-based LRs showed less consistent effects, with significant results in fewer experiments.

2.3.2 Normative adaptation of learning rate to noise

Contrary to the consistent volatility effect, the noise effect showed more variation across experiments (Table 2.6). Specifically, no significant noise effect was observed in Experiment 1 ($F(1,46) = 2.801$, $p = 0.101$), Experiment 2 ($F(1,50) = 2.397$, $p = 0.128$), Experiment 6 ($F(1,42) = 0.256$, $p = 0.615$), or in Experiment 7 ($F(1,43) = 1.953$, $p = 0.169$). However, a significant main effect of noise was observed in Experiment 3 ($F(1,46) = 4.211$, $p = 0.046$), Experiment 4 ($F(1,45) = 21.897$, $p < 0.001$), Experiment 5 ($F(1,41) = 7.448$, $p = 0.009$), and Experiment 8 ($F(1,48) = 13.089$, $p = 0.001$).

In Experiment 3, the significant noise effect was primarily driven by adaptation to noise under high volatility condition ($t(46) = 2.200$, $p = 0.020$), while adaptation under low volatility condition was not significant ($t(46) = 1.000$, $p = 0.160$). In Experiment 4, the adaptation of LRs is driven by low volatility condition ($t(45) = 5.640$, $p < 0.001$) but not high volatility condition ($t(45) = -1.490$, $p = 0.930$). In Experiment 5, normative adaptation to noise only observed in

low volatility condition ($t(41) = 2.040$, $p = 0.020$) and was near significant in high volatility condition ($t(41) = 1.490$, $p = 0.070$). In Experiment 8, significant adaptation was observed under both high volatility ($t(48) = 3.100$, $p < 0.001$) and low volatility ($t(48) = 2.400$, $p = 0.010$) conditions.

Although no main effects of noise were observed in Experiment 6, participants normatively adapted to noise under high volatility condition ($t(42) = 2.220$, $p = 0.020$) but not under low volatility condition ($t(42) = -0.970$, $p = 0.830$). Similarly in Experiment 7, participants adapted to noise under high volatility condition ($t(43) = 2.470$, $p = 0.010$) but not under low volatility condition ($t(43) = 0.320$, $p = 0.380$).

We did not observe any significant effect of noise on normative adaptation to noise for trial-wise LRs, except in Experiment 1 ($F(1, 46) = 10.305$, $p = 0.002$). T-tests revealed that LRs were higher in high noise environments compared to low noise environments under both high volatility ($t(46) = -1.910$, $p = 0.030$) and low volatility ($t(46) = -2.330$, $p = 0.010$) conditions. Results from slope-based LRs, however, showed greater variability across experiments.

Experiment	R-W model	Trial-wise LR	Slope as LR
Experiment 1	$F(1, 46) = 2.801$ $p = 0.101$	$F(1, 46) = 10.305$ $p = 0.002$	$F(1, 46) = 5.977$ $p = 0.018$
Experiment 2	$F(1, 50) = 2.397$ $p = 0.128$	$F(1, 50) = 0.147$ $p = 0.704$	$F(1, 50) = 6.647$ $p = 0.013$
Experiment 3	$F(1, 46) = 4.211$ $p = 0.046$	$F(1, 46) = 0.609$ $p = 0.439$	$F(1, 46) = 0.752$ $p = 0.390$
Experiment 4	$F(1, 45) = 21.897$ $p < 0.001$	$F(1, 45) = 0.126$ $p = 0.725$	$F(1, 45) = 5.969$ $p = 0.019$
Experiment 5	$F(1, 41) = 7.448$ $p = 0.009$	$F(1, 41) = 0.000$ $p = 0.983$	$F(1, 41) = 3.806$ $p = 0.058$
Experiment 6	$F(1, 42) = 0.256$ $p = 0.615$	$F(1, 42) = 0.018$ $p = 0.894$	$F(1, 42) = 0.795$ $p = 0.378$
Experiment 7	$F(1, 43) = 1.953$ $p = 0.169$	$F(1, 43) = 0.024$ $p = 0.879$	$F(1, 43) = 7.421$ $p = 0.009$
Experiment 8	$F(1, 48) = 13.089$ $p = 0.001$	$F(1, 48) = 2.682$ $p = 0.108$	$F(1, 48) = 3.463$ $p = 0.069$

Table 2.6: Noise effect across experiments and analysis methods.
This table summarizes the main effects of noise on learning rate (LR) across eight experiments using three different analytical approaches: the R-W model, median trial-wise LRs, and slope-based LRs. The R-W model consistently captures significant effects of noise in Experiments 3, 4, 5, and 8 ($p < 0.046$). Median trial-wise LRs reveal a significant noise effect only in Experiment 1 ($F(1, 46) = 10.305$, $p = 0.002$), where LRs were higher in high noise conditions, contrary to the hypothesized direction. Slope-based LRs demonstrate significant effects in Experiments 1, 2, 4, and 7 ($p < 0.013$).

2.3.3 Precision decreases as noise and volatility increase

Overall, participants have the highest precision in low volatility, low noise blocks. Precision decreases with increasing uncertainties in the environment. In Experiment 1, we observe a significant effect of volatility on precision ($F(1, 46) = 19.756$, $p < 0.001$). This effect is replicated in Experiment 2 ($F(1, 50) = 72.387$, $p < 0.001$), Experiment 3 ($F(1, 46) = 118.658$, $p < 0.001$), Experiment 4 ($F(1, 45) = 82.796$, $p < 0.001$), Experiment 5 ($F(1, 41) = 319.877$, $p < 0.001$), Experiment 6 ($F(1, 42) = 75.490$, $p < 0.001$), Experiment 7 ($F(1, 43) = 141.884$, $p < 0.001$), and Experiment 8 ($F(1, 48) = 54.130$, $p < 0.001$).

High noise significantly reduces precision as well, with significant effects observed in Experiment 1 ($F(1, 46) = 41.459$, $p < 0.001$), Experiment 4 ($F(1, 45) = 11.142$, $p = 0.002$), Experiment 5 ($F(1, 41) = 7.266$, $p = 0.010$), Experiment 6 ($F(1, 42) = 10.663$, $p = 0.002$), Experiment 7 ($F(1, 43) = 51.642$, $p < 0.001$), and Experiment 8 ($F(1, 48) = 4.849$, $p = 0.032$). However, no significant effect of noise on precision was observed in Experiment 3 ($F(1, 46) = 3.359$, $p = 0.073$) and Experiment 2 ($F(1, 50) = 1.999$, $p = 0.164$).

2.3.4 Bucket size in Experiment 1 – 3 is driven by noise and volatility

In Experiment 1 – 3, we designed a resizable bucket such that we could measure how confident participants were about their predictions. Volatility has a significant main effect on bucket size adjustment in Experiment 1 ($F(1, 37) = 11.793$, $p = 0.001$), Experiment 2 ($F(1, 50) = 37.506$, $p < 0.001$) and Experiment 3 ($F(1, 49) = 59.605$, $p < 0.001$). Similarly, noise also has a significant main effect across Experiment 1 ($F(1, 37) = 27.982$, $p < 0.001$), Experiment 2 ($F(1, 50) = 6.991$, $p = 0.011$) and Experiment 3 ($F(1, 49) = 7.727$, $p = 0.008$), suggesting participants increase their bucket sizes in high noise environment. These sets of results suggest participants become less confident in high volatility and noise conditions.

2.3.5 Psychiatric symptoms are not associated with learning rate adaptation to Volatility

Overall, psychiatric symptoms were not consistently associated with R-W learning rate adaptation to volatility (Table 2.7). Across Experiments 1 through 8, we did not observe any significant correlation between LR adaptation to volatility and depressive symptoms ($p > 0.260$) or trait anxiety ($p > 0.197$). Functional impairment caused by mood-related symptoms, measured by WSAS, were not related to volatility adaptation ($p > 0.256$). However, a significant correlation was observed between anticipatory TEPS and LR adaptation to volatility under low noise conditions ($r_s(44) = -0.356, p = 0.015$) in Experiment 4. Specifically, participants who reported a higher ability to derive pleasure from future rewards exhibited less adaptation to volatility under low noise condition. This effect, however, was not found in the rest of the experiments ($p > 0.085$) and did not pass multiple comparison correction. Additionally, no significant correlations were found between LR adaptation to volatility and consummatory TEPS ($p > 0.168$).

Experiment	Noise	CES-D	WSAS	STAI	Ant-TEPS	Con-TEPS
1	high	$r = 0.157$, $p = 0.292$	$r = 0.155$, $p = 0.298$	$r = 0.090$, $p = 0.572$	$r = 0.010$, $p = 0.947$	$r = -0.281$, $p = 0.720$
	low	$r = -0.060$, $p = 0.670$	$r = -0.150$, $p = 0.321$	$r = -0.110$, $p = 0.471$	$r = 0.157$, $p = 0.293$	$r = 0.179$, $p = 0.229$
2	high	$r = -0.056$, $p = 0.695$	$r = 0.117$, $p = 0.415$	$r = 0.017$, $p = 0.905$	$r = 0.137$, $p = 0.337$	$r = 0.040$, $p = 0.782$
	low	$r = -0.072$, $p = 0.613$	$r = 0.141$, $p = 0.325$	$r = 0.163$, $p = 0.254$	$r = 0.022$, $p = 0.881$	$r = -0.020$, $p = 0.887$
3	high	$r = -0.068$, $p = 0.651$	$r = 0.042$, $p = 0.779$	$r = -0.023$, $p = 0.876$	$r = -0.199$, $p = 0.180$	$r = -0.070$, $p = 0.641$
	low	$r = -0.079$, $p = 0.599$	$r = 0.088$, $p = 0.559$	$r = 0.192$, $p = 0.197$	$r = -0.254$, $p = 0.085$	$r = -0.086$, $p = 0.567$
4	high	$r = -0.170$, $p = 0.260$	$r = -0.171$, $p = 0.256$	$r = -0.045$, $p = 0.767$	$r = -0.199$, $p = 0.180$	$r = -0.207$, $p = 0.168$
	low	$r = 0.058$, $p = 0.702$	$r = 0.033$, $p = 0.825$	$r = 0.066$, $p = 0.664$	$r = -0.356$, $p = 0.015$	$r = -0.031$, $p = 0.837$
5	high	$r = -0.075$, $p = 0.638$	$r = -0.107$, $p = 0.502$	$r = -0.168$, $p = 0.288$	$r = -0.131$, $p = 0.410$	$r = -0.112$, $p = 0.481$
	low	$r = 0.049$, $p = 0.760$	$r = -0.064$, $p = 0.687$	$r = -0.082$, $p = 0.604$	$r = -0.107$, $p = 0.501$	$r = -0.115$, $p = 0.468$
6	high	$r = -0.020$, $p = 0.896$	$r = -0.080$, $p = 0.611$	$r = 0.026$, $p = 0.870$	$r = -0.072$, $p = 0.647$	$r = 0.021$, $p = 0.895$
	low	$r = 0.001$, $p = 0.993$	$r = 0.107$, $p = 0.494$	$r = 0.145$, $p = 0.354$	$r = 0.082$, $p = 0.602$	$r = 0.131$, $p = 0.403$
7	high	$r = -0.091$, $p = 0.557$	$r = -0.076$, $p = 0.623$	$r = 0.075$, $p = 0.630$	$r = -0.089$, $p = 0.566$	$r = 0.090$, $p = 0.560$
	low	$r = -0.118$, $p = 0.446$	$r = 0.021$, $p = 0.892$	$r = -0.094$, $p = 0.546$	$r = 0.080$, $p = 0.605$	$r = -0.041$, $p = 0.792$
8	high	$r = 0.049$, $p = 0.738$	$r = 0.029$, $p = 0.844$	$r = 0.104$, $p = 0.477$	$r = -0.151$, $p = 0.300$	$r = -0.181$, $p = 0.213$
	low	$r = -0.059$, $p = 0.688$	$r = 0.110$, $p = 0.452$	$r = -0.097$, $p = 0.506$	$r = -0.011$, $p = 0.939$	$r = -0.050$, $p = 0.735$

Table 2.7: Psychiatric symptoms are not associated with learning rate adaptation to volatility.
This table summarizes the correlations between the adaptation of learning rates (ΔLR_{vol}) to volatility and psychiatric measures across high and low noise conditions in Experiments 1 through 8. Correlations are presented for depressive symptoms (CES-D), functional impairment (WSAS), trait anxiety (STAI), anticipatory TEPS (Ant-TEPS), and consummatory TEPS (Con-TEPS). Significant correlations were rare and primarily observed under specific conditions. For example, in Experiment 4 under low noise conditions, a significant negative correlation between anticipatory TEPS and ΔLR_{vol} was found ($r_s(44) = -0.356$, $p = .015$), indicating that participants with higher anticipatory pleasure reported less adaptation to volatility. No other significant correlations were consistently observed across experiments for the remaining measures or noise levels ($p > 0.085$).

2.3.6 Psychiatric symptoms are not associated with learning rate adaptation to noise

Overall, psychiatric symptoms were not consistently associated with R-W learning rate adaptation to noise (Table 2.8). In Experiment 2, there was a significant positive relationship between LR adaptation to noise and depressive symptoms ($r_s(48) = 0.281$, $p = 0.046$), as well as with trait anxiety ($r_s(48) = 0.305$, $p = 0.030$). These results suggest that individuals with higher trait anxiety and depressive symptoms adapt more to noise under high-noise conditions.

However, these findings were not replicated in the remaining experiments. Specifically, there was no significant correlation between LR adaptation to noise and trait anxiety ($p > 0.083$) or depressive symptoms ($p > 0.105$). Similarly, we did not observe any significant relationships between LR adaptation to noise and functional impairment ($p > 0.114$), anticipatory TEPS ($p > 0.064$), or

consummatory TEPS ($p > 0.057$).

Furthermore, there was no significant main effect of noise on LR adaptation at the group level in Experiment 2, which complicates the interpretation of individual differences. Without a significant group-level effect, individual differences may reflect noise in the data rather than true variability in how individuals adapt to noise. For example, observed correlations with psychiatric measures might be driven by chance or confounded by other unmeasured variables.

Experiment	Volatility	CES-D	WSAS	STAI	Ant-TEPS	Con-TEPS
1	high	$r = -0.080$, $p = 0.594$	$r = -0.043$, $p = 0.774$	$r = -0.129$, $p = 0.389$	$r = 0.128$, $p = 0.392$	$r = 0.272$, $p = 0.064$
	low	$r = 0.122$, $p = 0.413$	$r = 0.125$, $p = 0.401$	$r = -0.026$, $p = 0.864$	$r = -0.037$, $p = 0.807$	$r = -0.279$, $p = 0.057$
2	high	$r = 0.281$, $p = 0.046$	$r = 0.133$, $p = 0.353$	$r = 0.305$, $p = 0.030$	$r = 0.049$, $p = 0.735$	$r = 0.069$, $p = 0.631$
	low	$r = 0.041$, $p = 0.776$	$r = -0.024$, $p = 0.868$	$r = -0.005$, $p = 0.972$	$r = 0.127$, $p = 0.373$	$r = 0.130$, $p = 0.363$
3	high	$r = -0.055$, $p = 0.712$	$r = 0.021$, $p = 0.887$	$r = -0.039$, $p = 0.796$	$r = 0.009$, $p = 0.951$	$r = 0.057$, $p = 0.705$
	low	$r = -0.154$, $p = 0.301$	$r = 0.053$, $p = 0.724$	$r = -0.199$, $p = 0.180$	$r = -0.097$, $p = 0.518$	$r = -0.030$, $p = 0.843$
4	high	$r = -0.056$, $p = 0.710$	$r = -0.151$, $p = 0.316$	$r = 0.213$, $p = 0.155$	$r = 0.052$, $p = 0.731$	$r = 0.132$, $p = 0.383$
	low	$r = -0.142$, $p = 0.345$	$r = -0.196$, $p = 0.193$	$r = 0.034$, $p = 0.822$	$r = -0.275$, $p = 0.064$	$r = -0.159$, $p = 0.292$
5	high	$r = -0.254$, $p = 0.105$	$r = -0.160$, $p = 0.312$	$r = 0.271$, $p = 0.083$	$r = -0.119$, $p = 0.453$	$r = -0.060$, $p = 0.708$
	low	$r = 0.110$, $p = 0.488$	$r = 0.125$, $p = 0.431$	$r = 0.170$, $p = 0.283$	$r = -0.158$, $p = 0.317$	$r = -0.021$, $p = 0.897$
6	high	$r = -0.010$, $p = 0.950$	$r = 0.244$, $p = 0.114$	$r = -0.016$, $p = 0.919$	$r = -0.230$, $p = 0.138$	$r = 0.058$, $p = 0.710$
	low	$r = 0.035$, $p = 0.821$	$r = 0.023$, $p = 0.885$	$r = -0.045$, $p = 0.776$	$r = -0.278$, $p = 0.071$	$r = -0.154$, $p = 0.323$
7	high	$r = 0.038$, $p = 0.805$	$r = 0.273$, $p = 0.073$	$r = 0.100$, $p = 0.518$	$r = -0.098$, $p = 0.528$	$r = 0.171$, $p = 0.268$
	low	$r = -0.005$, $p = 0.975$	$r = 0.088$, $p = 0.571$	$r = 0.170$, $p = 0.270$	$r = -0.126$, $p = 0.413$	$r = 0.178$, $p = 0.249$
8	high	$r = 0.086$, $p = 0.557$	$r = 0.156$, $p = 0.285$	$r = -0.068$, $p = 0.641$	$r = 0.123$, $p = 0.398$	$r = -0.224$, $p = 0.121$
	low	$r = 0.083$, $p = 0.568$	$r = 0.018$, $p = 0.900$	$r = 0.085$, $p = 0.560$	$r = -0.031$, $p = 0.834$	$r = -0.266$, $p = 0.064$

Table 2.8: Psychiatric symptoms are not associated with learning rate adaptation to noise.
This table summarizes the correlation between learning rate (ΔLR_{noise}) adaptation to noise and various psychiatric measures, including depressive symptoms (CES-D), functional impairment (WSAS), trait anxiety (STAI), and anticipatory and consummatory components of the Temporal Experience of Pleasure Scale (Ant-TEPS and Con-TEPS, respectively).

2.3.7 Working memory capacity is not related to learning under uncertainty

We did not observe any significant effect of working memory capacity on learning rate adaptations to volatility ($p > 0.112$) or noise ($p > 0.217$) in Experiment 1 through 8 (see Table 2.9).

Experiment	Condition	$\Delta LR_{\mathbf{vol}}$ & OSPAN (r, p)	$\Delta LR_{\mathbf{noise}}$ & OSPAN (r, p)
Experiment 1	high	$r = -0.039$, $p = 0.795$	$r = -0.018$, $p = 0.907$
	low	$r = -0.105$, $p = 0.481$	$r = -0.004$, $p = 0.977$
Experiment 2	high	$r = 0.130$, $p = 0.365$	$r = -0.001$, $p = 0.996$
	low	$r = 0.225$, $p = 0.112$	$r = -0.084$, $p = 0.556$
Experiment 3	high	$r = -0.054$, $p = 0.718$	$r = -0.054$, $p = 0.718$
	low	$r = -0.060$, $p = 0.691$	$r = -0.039$, $p = 0.794$
Experiment 4	high	$r = -0.026$, $p = 0.862$	$r = -0.085$, $p = 0.576$
	low	$r = -0.012$, $p = 0.934$	$r = -0.041$, $p = 0.787$
Experiment 5	high	$r = 0.164$, $p = 0.298$	$r = -0.194$, $p = 0.217$
	low	$r = -0.132$, $p = 0.404$	$r = 0.176$, $p = 0.264$
Experiment 6	high	$r = -0.028$, $p = 0.857$	$r = -0.000$, $p = 0.999$
	low	$r = -0.022$, $p = 0.890$	$r = 0.032$, $p = 0.837$
Experiment 7	high	$r = 0.002$, $p = 0.990$	$r = 0.015$, $p = 0.922$
	low	$r = 0.065$, $p = 0.673$	$r = -0.099$, $p = 0.523$
Experiment 8	high	$r = 0.056$, $p = 0.701$	$r = 0.097$, $p = 0.506$
	low	$r = 0.029$, $p = 0.842$	$r = 0.045$, $p = 0.759$

Table 2.9: Working memory and learning rate adaptation.
This table shows Spearman correlation coefficients (r) and corresponding p-values for the association between OSPAN scores and learning rate adaptation to volatility and noise across high and low uncertainty conditions. No consistent patterns of association were observed.

2.4 Discussion

Learning under uncertainty is a fundamental cognitive process, yet the interaction between different sources of uncertainty and their effect on adaptation remains understudied. This study aimed to develop and validate an experimental paradigm capable of manipulating both noise and volatility simultaneously. To achieve this, we designed systematic rules for generating experimental schedules and tested these schedules across eight experiments. We successfully identified a schedule that demonstrated robust effects of both volatility and noise, providing a suitable framework for subsequent studies investigating learning under uncertainty and its relationship with mood symptoms.

We compared three methods for estimating participants' learning rates: trial by trial estimates, slopes from linear regression, and estimates from model fitting. We found that model fitting was the most suitable approach for this task. Continuous

predictions and outcomes in this task led to extreme values in certain cases, such as when prediction errors were very small (resulting in very large learning rates; see Figure 2.4) or when prediction errors and changes in belief occurred in opposite directions (see Figure 2.5). These scenarios created challenges for trial-wise and slope-based approaches.

Previous studies addressed similar issues by restricting learning rates to values between 0 and 1, excluding trials with predictions outside the range of the previous prediction and the most recent outcome (Nassar et al., 2010, 2012). Alternatively, some studies excluded trials with extreme learning rates (above the 99th percentile) or zero spatial prediction error (Vaghi et al., 2017). However, these "atypical" values could reflect meaningful cognitive strategies. For example, learning rates surpassing 1 could indicate proactive strategies where participants anticipate the coin's landing site. Similarly, learning rates below zero might suggest that participants attribute the most recent observation to noise, adjusting their predictions accordingly for subsequent trials.

We thus decided to mitigate the aforementioned issues without excluding trials with learning rates below 0 or above 1, which accounted for approximately 30% of our data. For trial-wise learning rates, we used medians instead of means to reduce bias caused by extreme values when calculating block-wise learning rates. Additionally, we adjusted the calculation of belief updates to ensure changes in belief aligned with the direction of prediction errors when subsequent errors exceeded half a circle while belief updates remained below half a circle. However, these adjustments required assumptions about the direction of belief updates, which introduced a layer of subjectivity. Unlike trial-wise or slope-based approaches, model-fitted learning rates avoid biases from outliers and incorporate the precision parameter, a key parameter when modelling human behavior.

Throughout eight experiments, learning rates were normatively adapted to volatility, suggesting a robust and consistent adaptability in humans. However, normative adaptation to noise was only observed under some schedules. These observations lead to the question of why it is so much harder to detect and adapt

to noise. One explanation proposed by Williams and Griffiths (2013) argued that randomness (in our case we call it noise) in the observations can be reasonably explained away by a systematic process, and thus, people are less likely to see random outcomes to be generated by a random process. Furthermore, there is a local representativeness effect when judging a binary sequence where participants assume there needs to be equal proportions of Outcome A and Outcome B (Reimers et al., 2018), suggesting people are looking for specific features (e.g., an equal proportion of heads and tails, and little internal structure) in the data when judging randomness. However, in the current study, outcomes were not binary but continuous, involving variability in learning rates. Detecting noise when estimating the mean of a continuous variable may be inherently more challenging because there are no clear features to guide judgments of randomness.

Contrary to expectations based on prior literature linking mood disorders to cognitive adaptability, we found no significant associations between psychiatric symptoms (e.g., anxiety, depression, or anhedonia) and learning rate adaptation. Similarly, working memory capacity, often associated with cognitive performance under uncertainty, showed no relationship with learning rate adaptation. These findings suggest that adaptability to volatility and noise may operate independently of working memory.

One potential explanation for these null findings lies in the design of the experimental schedule. If noise and volatility levels were too high, the task may have been overly difficult, preventing most participants from adapting effectively. Conversely, if the ratio was too low, the task may have been too easy, leading to universal adaptation. Additionally, the study did not specifically select participants based on depressive or anxiety symptoms, as these hypotheses were secondary, potentially limiting the sensitivity of our analyses to detect individual differences related to mood symptoms.

2.5 Conclusion

This study provides a novel experimental framework for investigating learning under uncertainty by manipulating both noise and volatility. Our findings underscore the robustness of model-fitted learning rates as the most suitable approach for estimating learning rates in tasks involving continuous predictions and outcomes. While normative adaptation to volatility was consistent, adaptation to noise proved more elusive, likely due to the inherent challenges of detecting randomness in continuous data. Future research could refine the experimental schedule and participant selection criteria to explore individual differences in learning rate adaptation more effectively.

3

Anxiety symptoms are associated with reduced pupil response to volatility

3.1 Introduction

In our previous study, we found that individuals adapt to environmental volatility in a normative manner, irrespective of the level of uncertainty present. However, adaptation to noise was found to be contingent on the noise level within the environment. To date, no pupillometry studies have specifically examined the interaction between noise and volatility. Additionally, there has been limited research into how depressive symptoms relate to adaptations to volatility and noise, despite the common co-morbidity between anxiety disorders and depression.

The objectives of this study were to investigate: (1) whether humans respond to volatility and noise in a normative manner; (2) pupil dilation in response to volatility and noise; and (3) whether depressive symptoms are associated with reduced adaptation to these factors. We expected to replicate previous findings from Chapter 2, indicating that learning rates would increase in volatile environments and decrease in noisy environments. As reviewed in Chapter 1, there is robust evidence that pupil dilation increases in response to high volatility (Browning et al., 2015; Nassar et al., 2012). We thus anticipated increased pupil dilation in response to volatility.

While de Berker et al. (2016) reported increased pupil dilation with model-derived noise estimates, this finding contrasts with the theoretical framework proposed by Dayan and Yu (2002). According to Dayan and Yu (2002), acetylcholine (ACh) signals the presence of noise, indicating that noise should lead to decreased pupil dilation, as it may signify situations in which attention allocation is suppressed due to the irrelevance of random, uninformative variations. Given that participants must distinguish between volatility (structured uncertainty) and noise (unstructured, random changes), we argue that the Dayan and Yu (2002) framework is more applicable. We thus hypothesized that pupil dilation would decrease in response to noise.

While de Berker et al. (2016) reported increased pupil dilation with model-derived noise estimates, this finding contrasts with the framework proposed by Yu and Dayan (2003), who argue that ACh signals expected uncertainty, or noise. In this view, elevated ACh reduces the influence of top-down expectations and enhances the precision of bottom-up input (Moran et al., 2013; Pinto et al., 2013). Empirical evidence is consistent with this account (reviewed in Chapter 1): nicotine facilitates faster reorienting to invalid cues (F. Murphy & Klein, 1998; Stewart et al., 2001; Witte et al., 1997), scopolamine slows attentional shifts (Phillips et al., 2000), and cholinergic lesions impair adaptive reallocation of attention (Bucci et al., 1998). Taken together, these findings suggest that ACh "signals noise" by biasing attentional allocation in contexts where outcomes are best treated as random. Because participants in our task had to distinguish between volatility (structured change) and noise (unstructured randomness), we considered the Dayan and Yu framework most applicable and hypothesised that pupil dilation would decrease in response to noise.

Previous research suggests that higher trait anxiety is linked to reduced adaptation in both learning rate and pupillary response to environmental volatility (Browning et al., 2015). Furthermore, Gagne et al. (2020) showed that this effect is not specific to anxiety but is better explained by a shared internalizing dimension. In their framework, this dimension is a latent factor that captures variance common to anxiety and depression symptoms, distinct from variance unique to either disorder.

In particular, they found that higher scores on the general internalizing factor predicted reduced adjustment of the learning rate to volatility. On this basis, we hypothesized that depressive symptoms would also be associated with reduced adaptation to both volatility and noise.

In this study, 50 participants with varying levels of depressive symptoms completed the coin-catching task, which manipulates both volatility and noise simultaneously. Consistent with our first hypothesis, participants adjusted their learning rates normatively in response to both volatility and noise. However, we did not observe significant effects of either volatility or noise on pupil dilation. Notably, pupil dilation tracked volatility only in participants with low levels of anxiety. Additionally, there was no significant relationship between the adaptation of learning rate to volatility or noise and measures of depressive symptoms or trait anxiety. Interestingly, a higher anticipatory experience of pleasure was associated with reduced adaptation of learning rate to both volatility and noise.

3.2 Methods

Participants first completed an online survey that included the CES-D, demographic information, and questions about their mental health. Those meeting the selection criteria were invited to Warneford Hospital for a 1-hour testing session. After providing informed consent, participants received instructions for the coin-catching task. The task blocks were presented in random order, and the task itself lasted approximately 25 minutes. Following the coin-catching task, participants completed the Automated Operation Span Task (AOSPAN) to assess working memory capacity and a series of questionnaires, including the trait anxiety subscale of the STAI, WSAS, and TEPS. The entire experimental session lasted about 1 hour.

3.2.1 Sample

Fifty healthy, English-speaking adults (31 female; age 20–63 years, M = 31, SD = 10) were recruited from the Oxford, UK area (Table 3.1). Recruitment was conducted via Facebook Ads targeted by geolocation to users whose registered location fell

within a 15 km radius of the Oxford Botanic Garden. The advertisement appeared in users' news feeds and directed interested individuals to the prescreening survey. Exclusion criteria included current use of psychoactive medication, participation in psychological therapy, and any neurological illness. We used CES-D to recruit roughly equal numbers in three score ranges (0–9, 10–19, ≥20); invitations were sent irrespective of score provided the quota for that range had not yet been filled. Eight participants were excluded due to having over 80% missing pupillometry data, resulting in a final sample of 42 participants with analyzable eye-tracking data.

Depression Group	Screened (CESD-based)	Recruited	Discarded	Final Sample
High (CESD > 20)	115	16	2	14
Medium (10 < CESD ≤ 20)	64	19	3	16
Low (CESD ≤ 10)	26	15	3	12
Total	**205**	**50**	**8**	**42**

Table 3.1: Sample stratified by depression severity.
ChatGPT said: Participants were prescreened with the CES D to ensure coverage across symptom severity. Of the 205 individuals who completed prescreening, 50 enrolled. Participation rates, defined as the share of prescreened individuals within each CES D range who accepted an invitation and completed the study, differed by group. In particular, participation was highest in the low CES D range (0 to 9; 42%), lower in the medium range (10 to 19; 26%), and lowest in the high range (20 or above; 13%). This pattern was not intentional; although fewer low CES D individuals were present in the prescreening pool, they were more likely to respond to invitations. Eight participants were excluded due to more than 80% missing pupillometry data, yielding a final sample of 42.

3.2.2 Coin Catching Task

The coin-catching task was largely similar to the one described in Chapter 2. In this task, a coin was released from the center of a large circle. A yellow bucket, positioned at the edge of the circle with a central angle of 30 degrees, was used by participants to catch the coin (see Figure 3.1a). The task consisted of four blocks, each containing 50 trials: high volatility low noise, high volatility high noise, low volatility low noise, and low volatility high noise.

The bucket was controlled by a rotary controller, which allowed participants to adjust the bucket's position, and pressing the controller confirmed their chosen

location. There was no time limit per trial. During the instruction phase, participants were shown how to use the rotary controller and given an overview of the experiment's layout. They were informed that their task was to catch as many coins as possible and that the coin's landing position would change with each trial. Unlike in Piray and Daw (2024), participants were not informed of the reasons behind the changes in the coin's location. After the instructions, participants completed 10 practice trials. Before beginning the main task, we asked participants to answer two multiple choice questions to check their understanding of the task: "What is the purpose of the task?" and "How do I decide where to put the bucket?"

The coin's location for each trial was generated using a von Mises random walk ($\theta_i \sim \text{von Mises}(\mu_i, \kappa_{\text{noise}})$), with μ_i changing on each trial. The update of this parameter was determined by: $\mu_i = \mu_{i-1} + v_t$, where v_t, the step taken by the mean at each trial, was drawn from a second von Mises distribution $v_t \sim \text{von Mises}(0, \kappa_{\text{volatility}})$. In low volatility blocks, the $\kappa_{\text{volatility}}$ was set to 84, while in high-volatility blocks, it was set to 5. The κ_{noise} was set to 9 for high noise blocks and 42 for low noise blocks.

As a result, the coin's locations was influenced by both the mean (μ_i) and the concentration (κ_{noise}) of a von Mises distribution, analogous to 1/standard deviation of a standardised normal distribution. This procedure introduced both expected and unexpected uncertainty into the environment. Each time the mean changed, participants needed to form a new belief about the mean of the distribution and distinguish these changes from noise (Figure 3.1b).

At the start of each trial, the bucket was positioned at a random location to maintain participant engagement. Once participants chose a location, they submitted their prediction by pressing the controller, which triggered the coin's release on the screen. After the participant's response, the coin was released from the center of the circle (where the bag is located; see Figure 3.1a) toward the outcome location on the blue track. The coin took 672 ms to reach the blue track, where it remained visible for 1 second. Following this, a jitter interval (ranging

from 1 to 5 seconds) was added while the outcome was still displayed. Thus, the outcome was shown on the screen for a minimum total of 2672 ms.

If the coin landed within the bucket, a green message reading "caught the coin" appeared at the top of the screen, and the counter increased by 1. If the coin missed the bucket, a red message reading "missed the coin" appeared, and the counter remained unchanged.

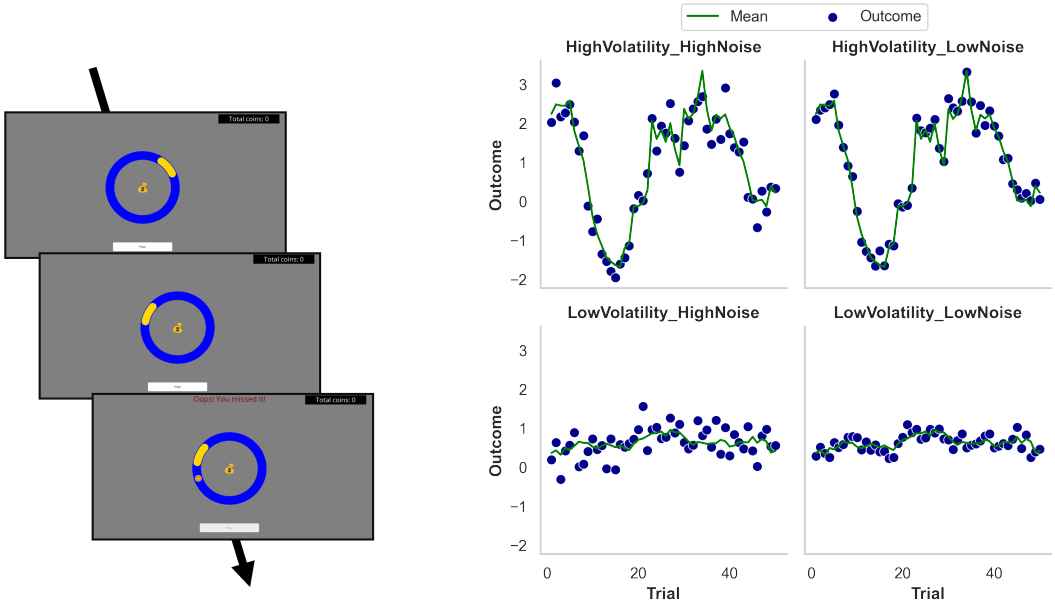


Figure 3.1: Task Structure: Coin catching task.
(a) Example trial. Participants were instructed to place the yellow bucket at the location where they believed the coin would most likely land. After positioning the bucket, a coin was released from the centre of the circle (taking 20 ms) and landed on the blue track. If the coin did not land in the yellow bucket (as shown in this example), a red message reading "Oops! You missed it" appeared at the top of the screen, and the total coin counter (located in the top right corner of the screen) remained unchanged. The outcome was displayed for a variable duration between 2 seconds and 7 seconds. (b) Experiment schedule. The landing positions of the coin for each block were generated using a von Mises random walk. In the high-volatility blocks, the mean of the distribution (green lines) changed every trial, alternating between clockwise and counterclockwise directions. In contrast, in the low-volatility blocks, the mean remained stable with minimal changes from trial to trial. The landing positions (blue dots) in high-noise blocks were more dispersed and farther from the mean, whereas, in low-noise blocks, the positions were closer to the mean.

3.2.3 Pupil Dilation Preprocessing

Pupil diameter data were recorded from both eyes using the Eyelink 1000 system (SR Research) at a sampling rate of 500 Hz. To clean the data, blinks were automatically removed using the Eyelink system's built-in filter, along with any saccades that occurred immediately before a blink. Missing data points were addressed through linear interpolation, with trials containing more than 50% interpolated data excluded from further analysis. A low-pass Butterworth filter with a 3.75 Hz cut-off was applied to the cleaned data. The filtered pupil diameters were z-transformed for normalization across the entire session and then averaged between the two eyes. For each trial, a 2-second pre-outcome window and a 5-second post-outcome window were extracted for analysis (Browning et al., 2015). To mitigate the effects of initial adaptation to luminance changes, the first five trials of each session were excluded (Browning et al., 2015; Nassar et al., 2012).

Although the primary analysis focused on tonic baseline pupil diameter, additional analyses were performed on baseline-corrected pupil dilation. Baseline correction involved subtracting the mean pupil size during the baseline period from each time point following the outcome presentation.

3.2.4 Pupil Dilation Analyses

Regression analyses were conducted to examine the effects of volatility (0, 1), noise (0, 1), their interaction, and fatigue on pupil dilation. Fatigue was included as a control variable. Given that there was a self-paced break between each block, fatigue was operationalized as the trial index within a block (ranging from 5 to 49). This definition captures the progressive decline in attention and increased cognitive fatigue as participants advanced through the block. Although an alternative definition of fatigue was considered, spanning trials 5 to 200, it was excluded due to significant collinearity with the interaction term, which biased the regression results. The analysis spanned 7 seconds, including a 2-second pre-outcome period and a 5-second post-outcome period. With an eye tracker sampling rate of 500

Hz, one data point was recorded every 2 ms, allowing for regression analyses to be conducted across all 3500 time intervals of 2 ms each.

Summary measures of beta weights for volatility, noise, the interaction between volatility and noise, and fatigue were calculated for each of the 2 ms timepoint. One-sample t-tests were used to determine whether these beta values significantly differed from zero. These beta weights were also used in correlational analyses to assess their relationship with behavioral adaptations and between-subject measures of psychiatric traits. Additionally, a median split was applied to analyze the effect of volatility within the low-anxiety group.

3.2.5 Computational Modelling

The primary measure of interest was the change in learning rates between high-volatility and low-volatility blocks, as well as between low-noise and high-noise blocks. To obtain these learning rates, we first fitted a simple learning model based on the Rescorla-Wagner framework for each block. We then tested two alternative models and performed parameter recovery for all three models, followed by model comparison.

Rescorla-Wagner Learning Model

The model follows the same implementation and design as the one introduced in Chapter 2.

Trend Learning model

The model described here involves a learning process based on the Rescorla-Wagner rule, with modifications to account for dynamic trends in the data. The objective of the model is to estimate how individuals predict the landing position of a coin, updating their beliefs over time using two main components: baseline learning and trend learning. Trend learning is the process of recognizing and adapting to systematic changes over time. It involves detecting consistent patterns within a volatile environment, such as a gradual increase or decrease. This type of learning

enables us to anticipate future outcomes based on the expected trajectory of the data. For instance, an example of a trend is the average prices of consumer products, which typically follow an upward trajectory over time. Not all trends are linear; for example, changes in maximum temperature throughout a year often exhibit a cyclical, non-linear pattern. This process is captured through four free parameters: LR_{baseline}, LR_{Δ}, a mixture parameter a and a precision parameter *prec*.

As outlined in Model 1, the Rescorla-Wagner learning rule was adapted to accommodate circular data. Specifically, participants' beliefs about the coin's landing position were updated on a trial-by-trial basis based on the prediction error. The belief update rule is:

$$\text{Belief}_t^{\theta} = \text{Belief}_{t-1}^{\theta} + LR_{\text{baseline}} \times PE_t, \tag{3.2.5.1}$$

where,

$$PE_t = \min\left(\left|\text{Outcome}_t - \text{Prediction}_t\right|,\ 2\pi - \left|\text{Outcome}_t - \text{Prediction}_t\right|\right) \tag{3.2.5.2}$$

In addition to baseline learning, the model assumes that individuals track changes in the coin's location from one trial to the next. This trend-tracking process is represented by Δ, with an associated learning rate denoted as LR_{Δ}. The update rule for trend learning is:

$$\text{Belief}_t^{\Delta} = \text{Belief}_{t-1}^{\Delta} + LR_{\Delta} \times (\Delta_{t-1} - \text{Belief}_{t-1}^{\Delta}) \tag{3.2.5.3}$$

where

$$\Delta_t = \text{Outcome}_t - \text{Outcome}_{t-1} \tag{3.2.5.4}$$

This component accounts for the dynamic nature of the environment. Specifically, trend learning involves identifying consistent tendencies that indicate movement in a specific direction. In this experiment, however, the volatility follows a non-structured pattern, and the coin does not exhibit a bias toward any particular direction.

The final belief about the coin's landing position combines the baseline belief and the trend belief. This combination is weighted by the mixture parameter α which determines the contribution of the delta effect:

$$\text{Belief}_t = \text{Belief}_t^{\theta} + a \times \text{Belief}_t^{\Delta} \tag{3.2.5.5}$$

After beliefs are updated, they are transformed into a probability distribution for predicting the coin's location. This is modelled using the von Mises probability density function, which is suitable for circular data (e.g., angles or periodic events) and includes a precision parameter to control the distribution's concentration around the mean. The probability of selecting a given location is:

$$\Pr(\text{Prediction}_t \mid \text{Belief}_t, \text{prec}) = \frac{\exp\left(\text{precision} \times \cos(\text{choice} - \text{belief})\right)}{2\pi I_0(\text{precision})} \quad (3.2.5.6)$$

Parameter Estimation

Learning rate parameters (LR_{baseline} and LR_{Δ}) and the mixture parameter (a) were allowed to vary between 0.001 and 0.999. Decision precision was sampled over a range of 0.01 to 100. Each parameter range was divided into 30 equally spaced discrete values. For Model 2, which has 4 free parameters, we tested all possible combinations, totalling 810,000 possibilities. The model was simulated on a trial-by-trial basis using each unique combination of parameter values. The likelihood across all trials was obtained by multiplying the trial-wise likelihoods of participants' choices given each parameter combination. This process was repeated for all combinations. The best estimate for each parameter was determined by calculating the expected value of its marginalized probability distribution, summing over all other parameters. The expected value was extracted by summing the discrete parameter values weighted by their marginalized probabilities.

The figure below illustrates the model's ability to replicate participant choices by comparing participant choice data with those generated by the model (see Figure 3.2).

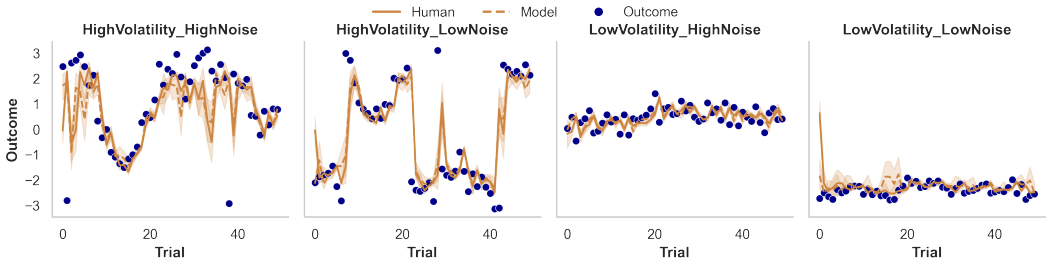
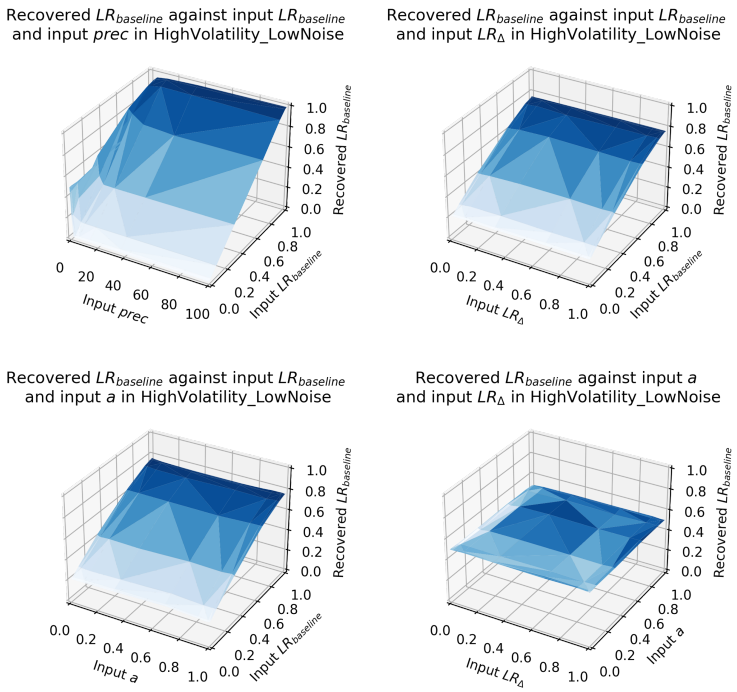


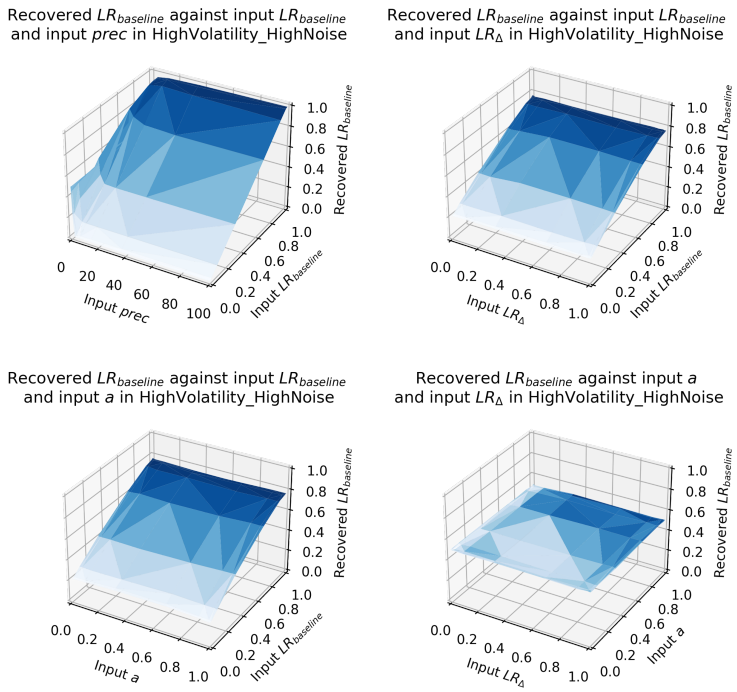
Figure 3.2: Comparison of Model-Simulated and Participant Behavioral Data. The plot shown here overlays the choice behaviour of participants (brown solid line) with the behaviour generated using model-fitted parameters (brown dash line) and the output (blue dots). The model output was generated by running the Trend learning model for each participant using their individually fitted parameter values on the task trial sequences. As shown, the model accurately reflects participant behaviours during the task.

Parameter Recovery

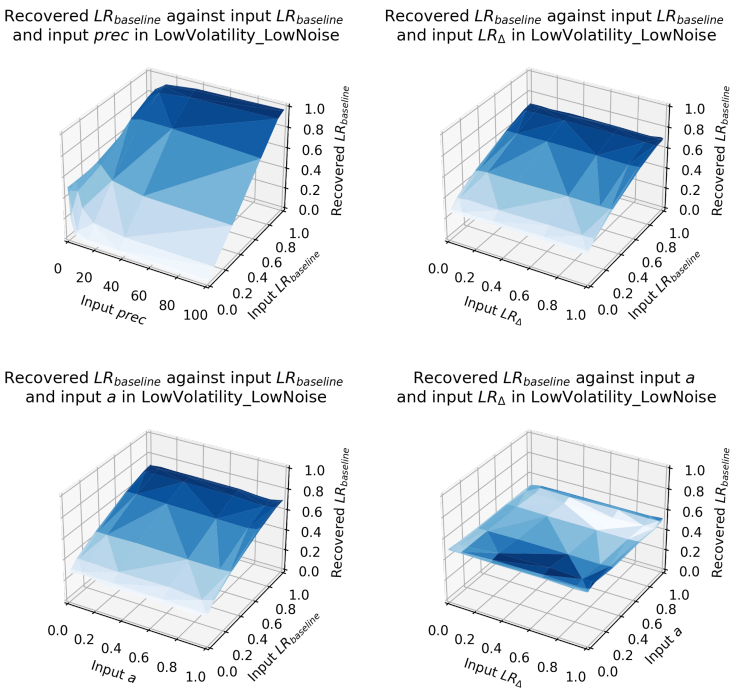
To assess whether our parameter estimation procedure reliably reproduces parameter values, parameter recovery was conducted for the trend learning model. First, predictions were simulated using 19,999 combinations of parameter values. For each of the four free parameters, 10 equally spaced values covering the entire parameter range used in the estimation procedure were applied to simulate choices. The trend learning model was then fitted to these simulated choices, and the resulting parameter estimates were compared to the input parameters used for simulation. Overall, LR_{baseline}, LR_{Δ}, and mixture parameter a were reasonably well recovered, closely matching their input values (see Figures 3.3, 3.4, and 3.5). However, the recovery of all three parameters was affected by low values of the precision parameter, *prec*. Specifically, under conditions of low input *prec*, predictions were more influenced by decision precision, resulting in a weaker correlation between the recovered and input learning rates and mixture parameter. However, LR_{baseline}, LR_{Δ}, and a influenced each other during recovery. Specifically, across all input values of a, LR_{baseline} could not be reliably recovered at smaller (< 0.2) or larger (> 0.8) values.



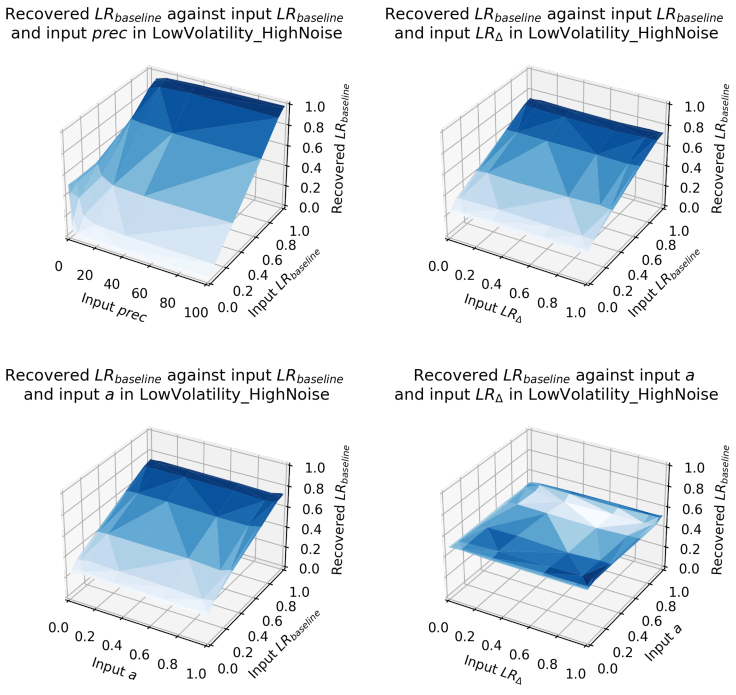
(a)Recoverability of LR_{baseline} in High Volatility Low Noise
The recovered LR_{baseline} (z-axis) is plotted for four input pairs arranged (top left) x = input *prec*, y = input LR_{baseline}; (top right) x = input LR_{Δ}, y = input LR_{baseline}; (bottom left) x = input α, y = input LR_{baseline}; (bottom right) x = input LR_{Δ}, y = input α. Recovered LR_{baseline} tracks the true LR_{baseline} approximately linearly across most values of prec, α, and LR_{Δ}, but recovery degrades when precision is low (small prec). Recovery is also unreliable when the true LR_{baseline} lies near the extremes (< 0.2 or > 0.8). Color shading from light to dark blue encodes low to high recovered LR_{baseline}.



(b)Recoverability of LR_{baseline} in High Volatility High Noise



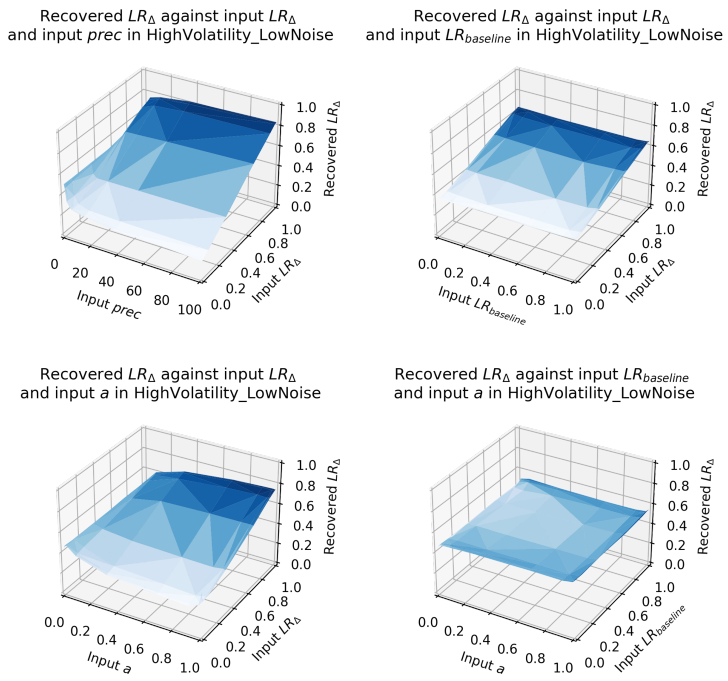
(c) Recoverability of LR_{baseline} in Low Volatility Low Noise



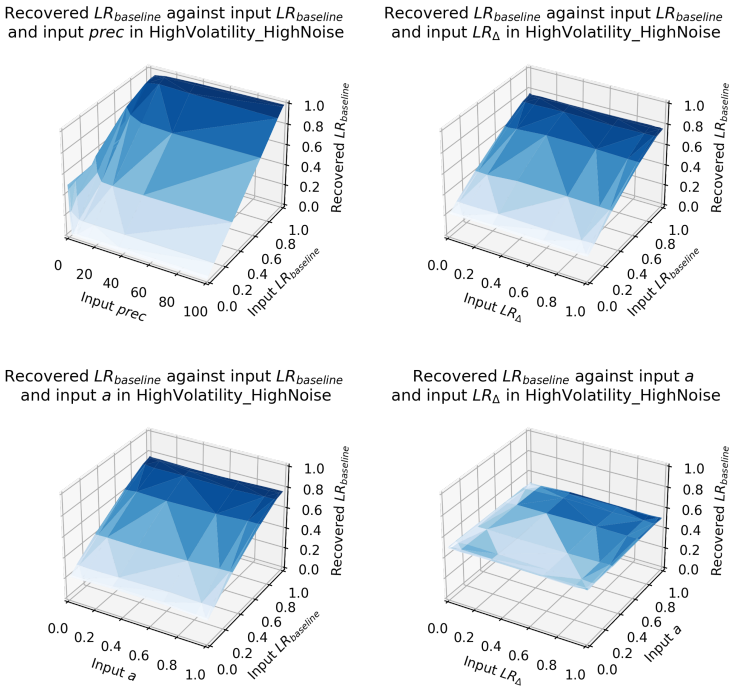
(d) Recoverability of LR_{baseline} in Low Volatility High Noise

Figure 3.3: Parameter recovery for $\mathbf{LR_{baseline}}$.
Parameter recovery for baseline learning rate ($LR_{baseline}$) was conducted for each block separately. Across all four blocks, $LR_{baseline}$ was reasonably well recovered, showing a clear linear relationship between input $LR_{baseline}$ and recovered $LR_{baseline}$. However, under conditions of low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered $LR_{baseline}$ and input $LR_{baseline}$. The input α and input LR_{Δ} affected the range of recoverable $LR_{baseline}$ values. Specifically, the recovered $LR_{baseline}$ was capped at 0.8, though a clear linear relationship between the input and recovered $LR_{baseline}$ persisted across all input α and input LR_{Δ} values. Additionally, there was no observed relationship between recovered LR and input precision or input α or input LR_{Δ}. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

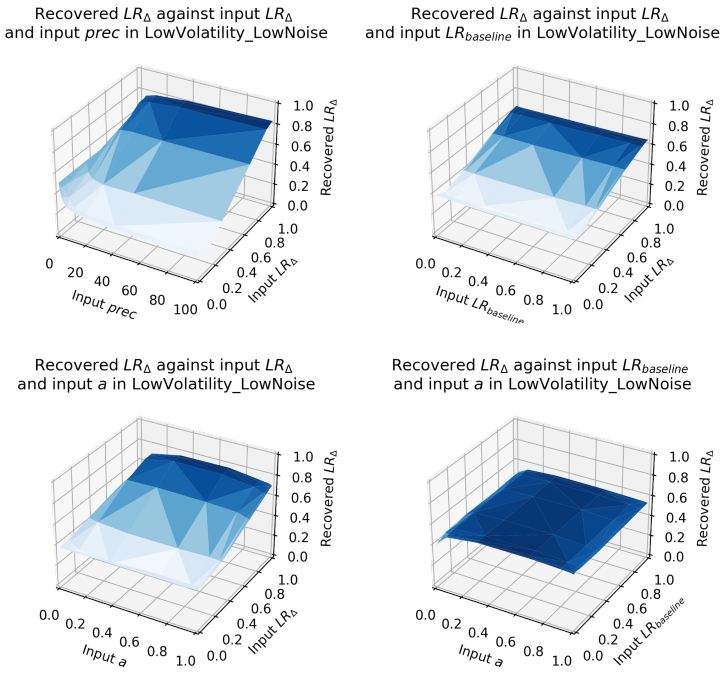
Similarly, when recovering LR_{Δ}, the recovery was affected by the input values of LR_{baseline}. Across all input values of LR_{baseline}, the recovery of LR_{Δ} was capped, with a reduced accuracy for values below 0.2 and above 0.8. Though input LR_{baseline} and input LR_{Δ} are not correlated with recovered a, what we observed here shows there is some level of interdependency between these three parameters (see Figure 3.4).



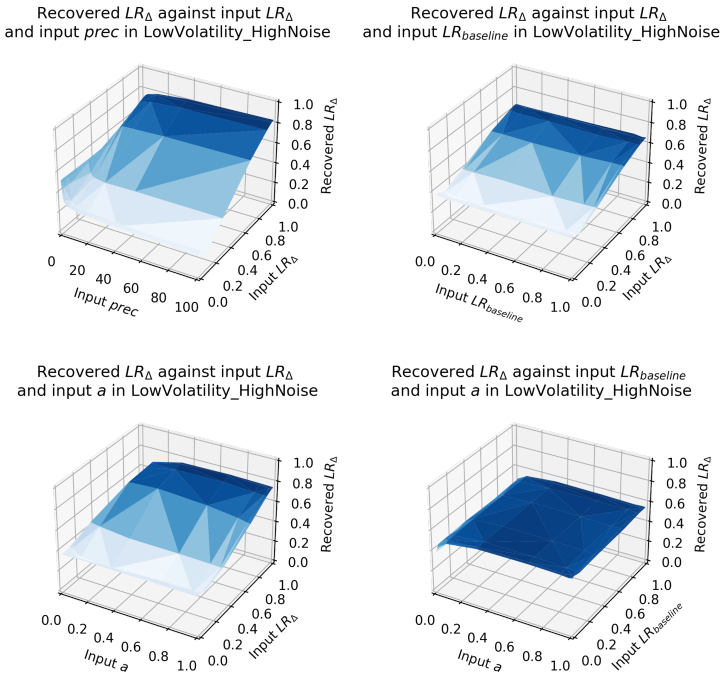
(a)Recoverability of LR_Δ in High Volatility Low Noise.
The recovered LR_Δ (z-axis) is shown for four input pairs arranged top-left, top-right, bottom-left, bottom-right, respectively: x = input *prec*, y = input LR_Δ; x = input LR_{baseline}, y = input LR_Δ; x = input α, y = input LR_Δ; x = input α, y = input LR_{baseline}. Recovery of LR_Δ is systematically capped across the range of LR_{baseline}, with reduced accuracy for true $LR_\Delta < 0.2$ and > 0.8. Although input LR_{baseline} and input LR_Δ show no direct association with recovered α, the joint surfaces suggest some interdependency among LR_{baseline}, LR_Δ, and α.



(b)Recoverability of LR_{baseline} in High Volatility High Noise



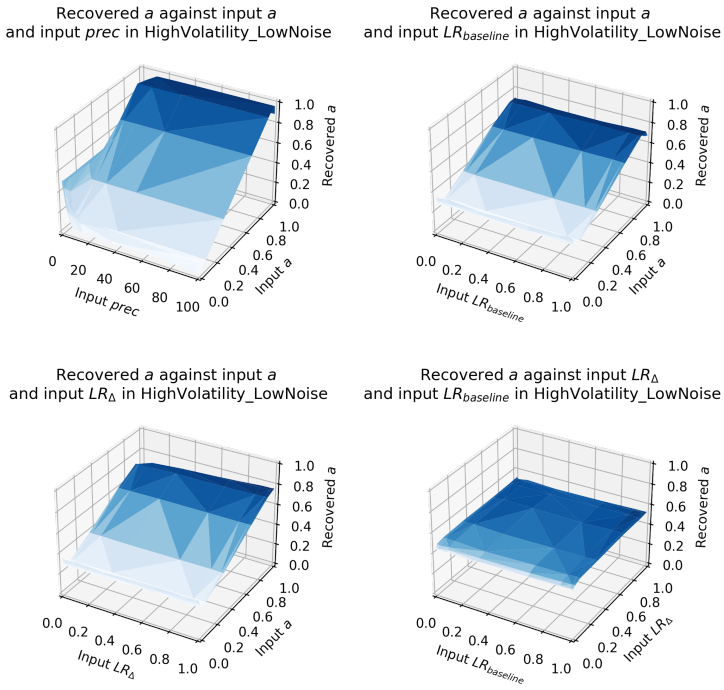
(c) Recoverability of LR_Δ in Low Volatility Low Noise



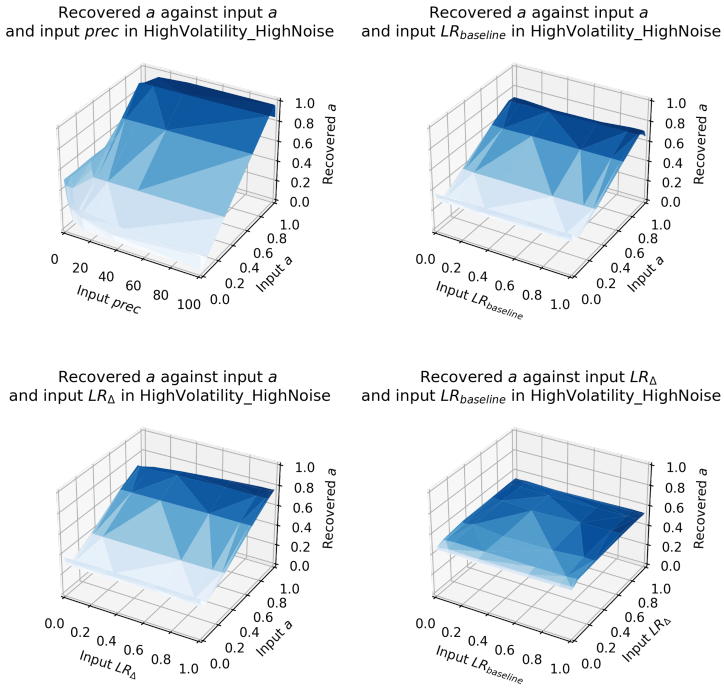
(d) Recoverability of LR_Δ in Low Volatility High Noise

Figure 3.4: Parameter recovery for LR_{Δ}.
Parameter recovery for trend learning rate (LR_{Δ}) was conducted for each block separately. Across all four blocks, LR_{Δ} was reasonably well recovered, showing a clear linear relationship between input LR_{Δ} and recovered LR_{Δ}. However, under conditions of low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered LR_{baseline} and input LR_{baseline}. The input α and input LR_{baseline} affected the range of recoverable LR_{Δ} values. Specifically, the recovered LR_{Δ} was capped at 0.6, though a clear linear relationship between the input and recovered LR_{Δ} persisted across all input α and input LR_{baseline} values. Additionally, there was no observed relationship between recovered LR and input precision or input α or input LR_{baseline}. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

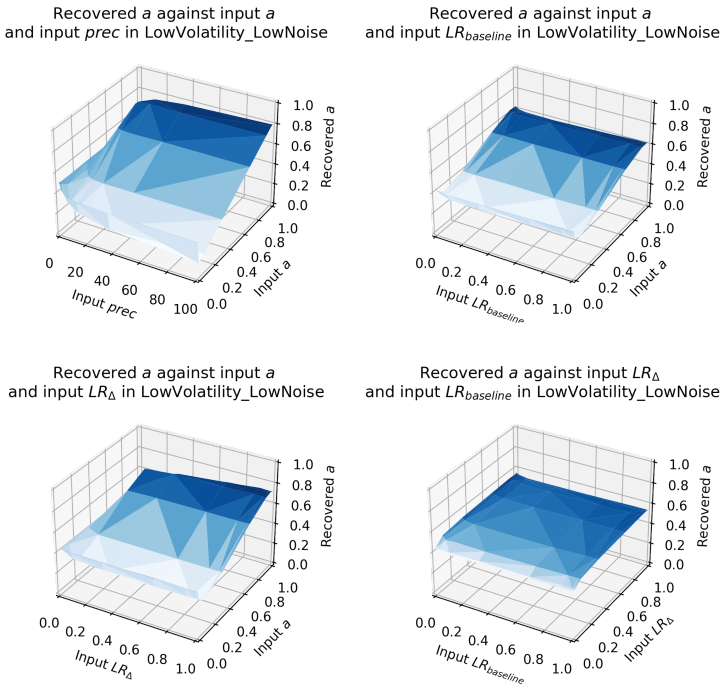
We assessed identifiability of the change-sensitivity parameter, α, via parameter-recovery in the High-Volatility/Low-Noise block (Fig. **??**; see caption for axis conventions). Recovered α tracked the true value approximately linearly across wide ranges of *prec*, LR_{baseline}, and LR_{Δ}, with accuracy degrading when *prec* was low and when the true α approached the bounds (< 0.2 or > 0.8). While LR_{baseline} and LR_{Δ} showed no systematic association with recovered α when α was held constant, the joint surfaces suggest mild interdependencies among these parameters in extreme regions.



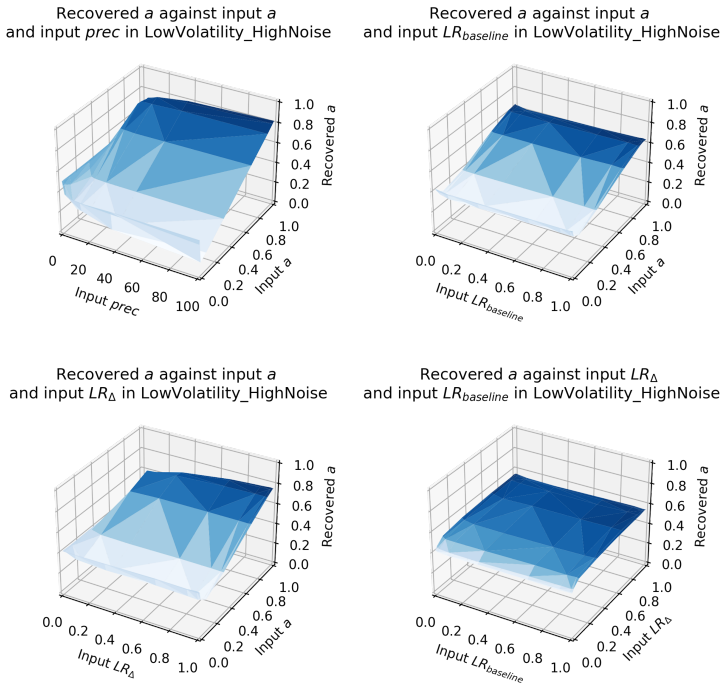
(a)Recoverability of α in High Volatility Low Noise.
The recovered α (z-axis) is plotted for four input pairs arranged top-left, top-right, bottom-left, and bottom-right, respectively: x = input *prec*, y = input α; x = input LR_{baseline}, y = input α; x = input LR_{Δ}, y = input α; x = input LR_{Δ}, y = input LR_{baseline}. Recovered α increases approximately linearly with the true α across most values of *prec*, LR_{baseline}, and LR_{Δ}, but recovery degrades when precision is low (small *prec*) and when the true α lies near the extremes (< 0.2 or > 0.8). The joint surfaces also suggest some interdependency among LR_{baseline}, LR_{Δ}, and α. Color shading from light to dark blue encodes low to high recovered α.



(b)Recoverability of α in High Volatility High Noise



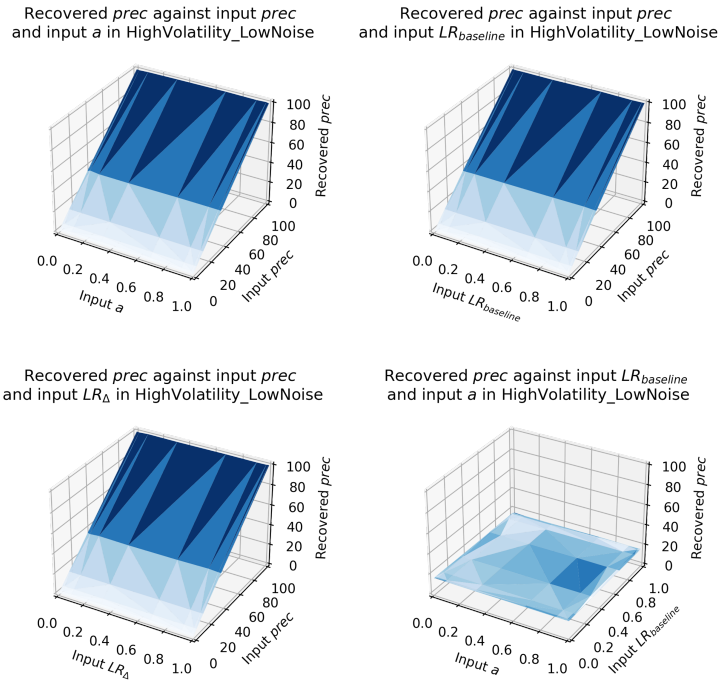
(c) Recoverability of α in Low Volatility Low Noise



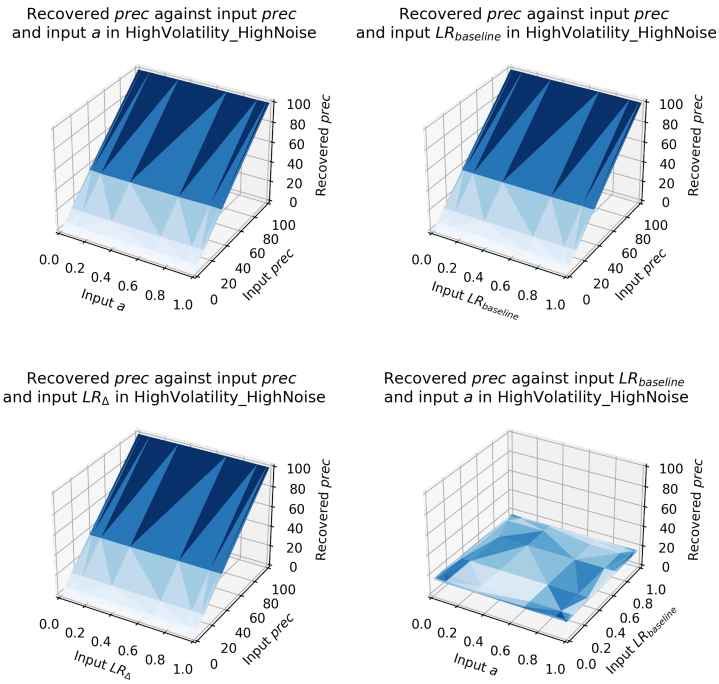
(d) Recoverability of α in Low Volatility High Noise

Figure 3.5: Parameter recovery for α.
Parameter recovery for mixture variable (α) was conducted for each block separately. Across all four blocks, α was reasonably well recovered, showing a clear linear relationship between input a and recovered a. Under low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered a and input a. Similar to recovery of both learning rates, input LR_{Δ} and input $LR_{baseline}$ affected the range of recoverable a values. Additionally, there was no observed relationship between recovered a and input precision or input LR_{Δ} or input $LR_{baseline}$. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

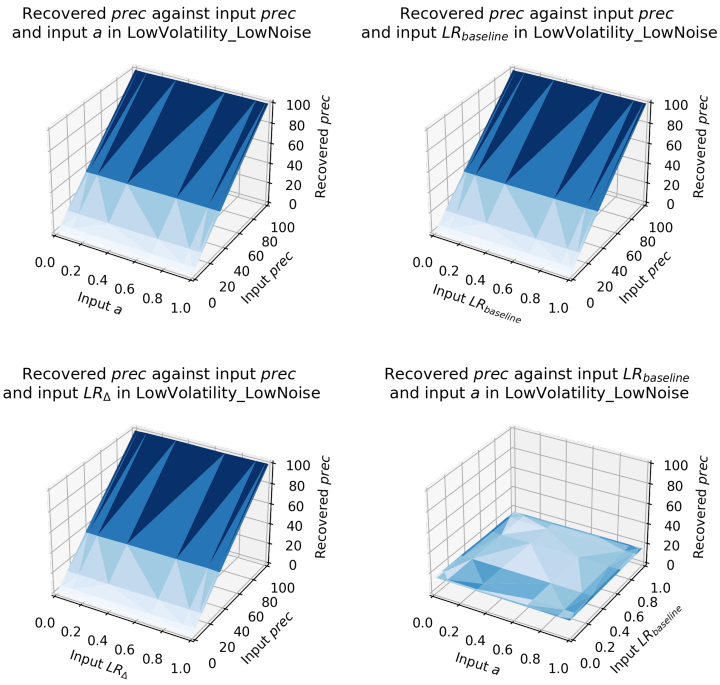
The precision parameter (*prec*) exhibited strong recoverability, which was independent of the input values of the other three free parameters (Figure 3.6).



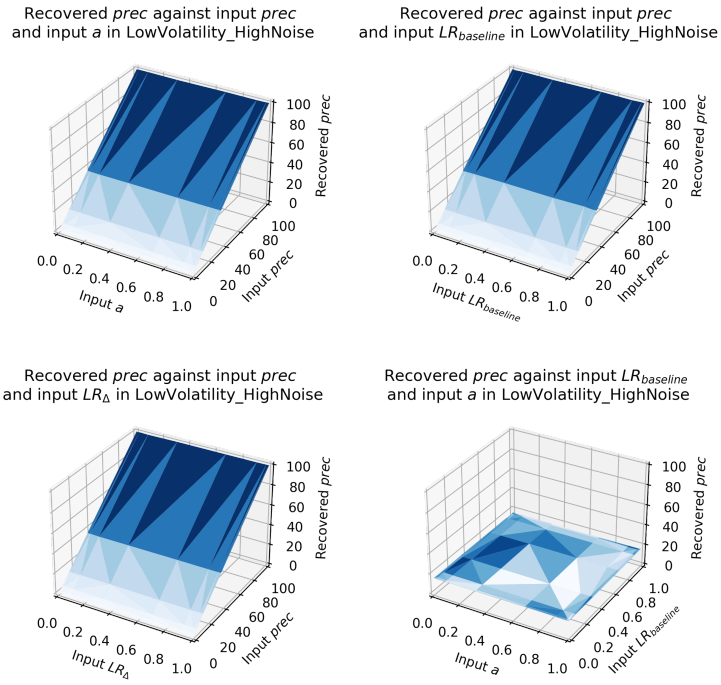
(a)Recoverability of *prec* in High Volatility Low Noise.
The recovered *prec* (z-axis) is shown for four input pairs arranged top-left, top-right, bottom-left, bottom-right, respectively: x = input α, y = input *prec*; x = input LR_{baseline}, y = input *prec*; x = input LR_{Δ}, y = input *prec*; x = input α, y = input LR_{baseline}. Recovery of *prec* was reasonably well across all combinations of input parameters. We observed a clear linear relationship was observed between input *prec* and recovered *prec*. No relationship was found between input learning rates (LR_{Δ}, $\text{LR}_{\text{baseline}}$), input α and recovered *prec*.



(b)Recoverability of *prec* in High Volatility High Noise



(c) Recoverability of *prec* in Low Volatility Low Noise



(d) Recoverability of *prec* in Low Volatility High Noise

Figure 3.6: Parameter recovery for *prec*.
Parameter recovery for decision precision (*prec*) was conducted for each block separately. Across all four blocks, a clear linear relationship was observed between input *prec* and recovered *prec*. No relationship was found between input learning rates (LR_Δ, $\text{LR}_{\text{baseline}}$), input α and recovered *prec*. The different shades of blue in these plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

Simplified Trend Learning Model

To improve parameter recovery and simplify the model, we removed the mixture parameter a. Consistent with trend learning model, the belief update rule is:

$$\text{Belief}_t^\theta = \text{Belief}_{t-1}^\theta + LR_{\text{baseline}} \times PE_t \tag{3.2.5.7}$$

The update rule for trend learning is:

$$\text{Belief}_t^\Delta = \text{Belief}_{t-1}^\Delta + LR_\Delta \times (\Delta_{t-1} - \text{Belief}_{t-1}^\Delta) \tag{3.2.5.8}$$

The final belief about the coin's landing position is a sum of the baseline belief and the trend belief.

$$\text{Belief}_t = \text{Belief}_t^\theta + \text{Belief}_t^\Delta \tag{3.2.5.9}$$

The probability of choosing a location given the belief of where the mean is:

$$\Pr(\text{Prediction}_t \mid \text{Belief}_t, \text{prec}) = f(\text{Prediction}_t \mid \text{Belief}_t, \text{prec}) \tag{3.2.5.10}$$

Parameter Estimation

Parameter estimation followed the same procedure as in the trend learning model. Particularly, learning rate parameters (LR_{baseline} and LR_Δ) were allowed to vary between 0.001 and 0.999. Decision precision was sampled over a range of 0.01 to 100. Each parameter's range was divided into 30 equally spaced discrete values. For the simplified trend learning model, consisting of 3 free parameters, we tested all possible combinations.

The figure below illustrates the model's ability to replicate participant choices by comparing participant data with those generated by the model (see Figure 3.7).

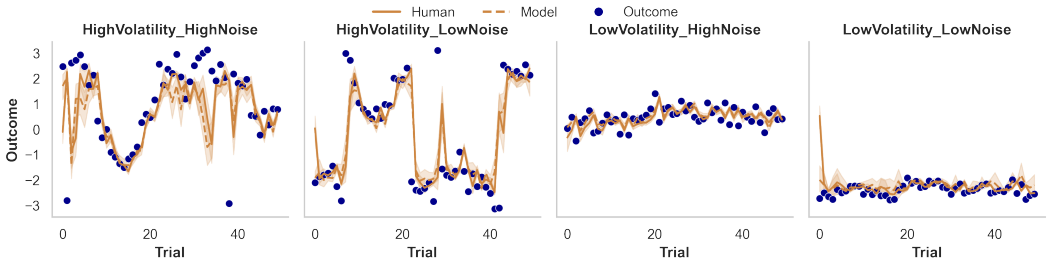
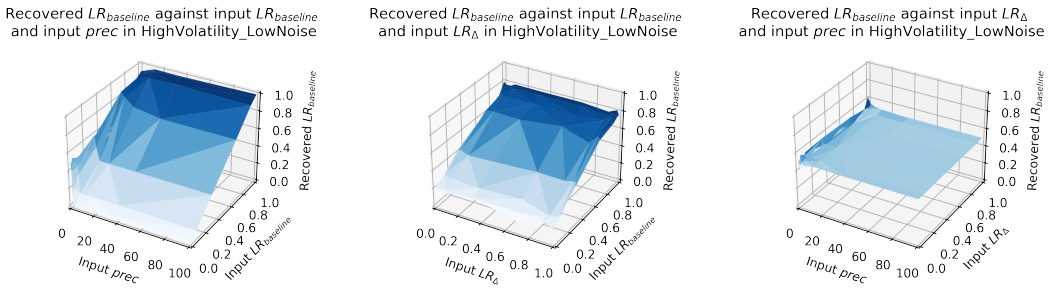


Figure 3.7: Comparison of Model-Simulated and Participant Behavioral Data. The plot shown here overlays the choice behaviour of participants (brown solid line) with the behaviour generated using model-fitted parameters (brown dash line) and the output (blue dots). The model output was generated by running the Simplified trend learning model for each participant using their individually fitted parameter values on the task trial sequences. As shown, the model accurately reflects participant behaviours during the task.

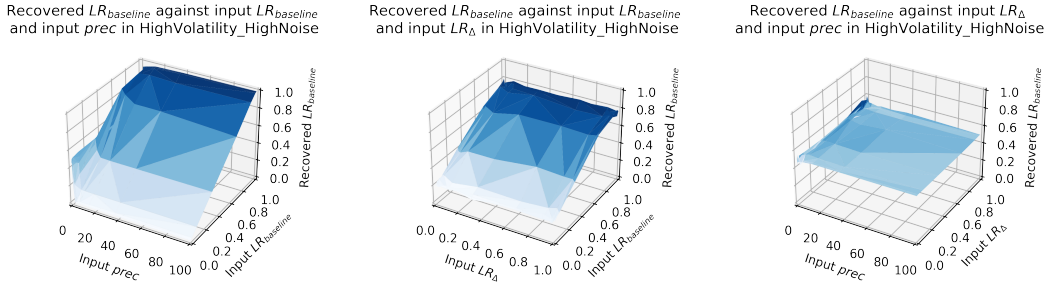
Parameter Recovery

To test whether our procedure of parameter estimation can reliably reproduce parameter values, parameter recovery was conducted for the simplified trend learning model. First, predictions were simulated from the simplified trend learning model using 4000 combinations of parameter values. For each of the four free parameters, 10 equally spaced values covering the entire parameter range considered in the estimation procedure were used for simulating choices. The simplified trend learning model was then fitted to simulated choices and the resulting parameter estimates were compared to the input parameters for simulating the choices.

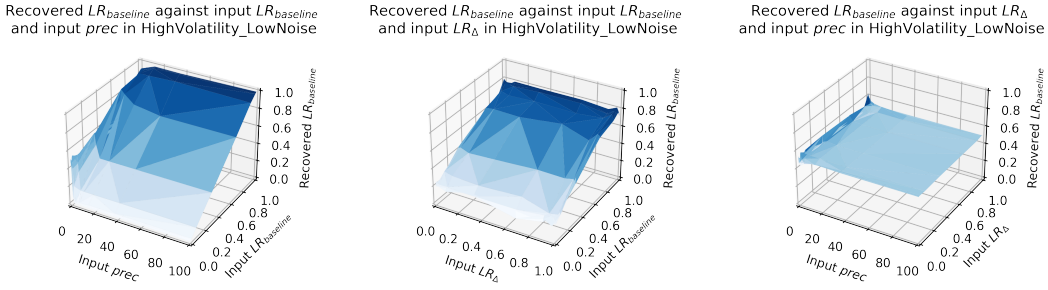
Overall, LR_{baseline} and LR_{Δ} recovered reasonably well, close to its input values (Figure 3.8 and 3.9). The precision parameter (*prec*) exhibited strong recoverability, which was independent of the input values of the other three free parameters (Figure 3.10).



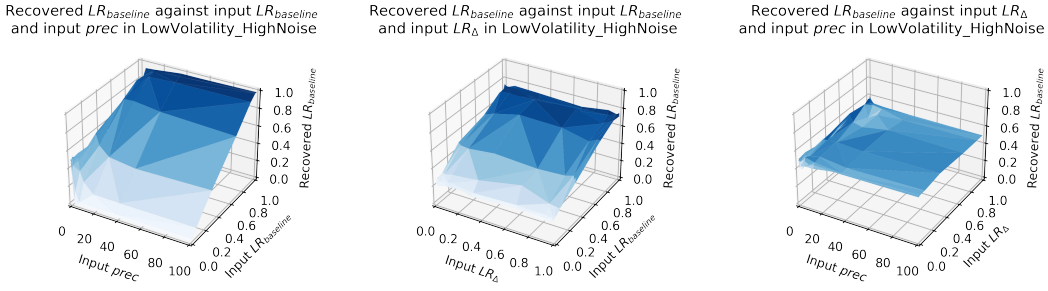
(a)Recoverability of LR_{baseline} in High Volatility Low Noise.
The recovered LR_{baseline} (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input *prec*, y = input LR_{baseline}; x = input LR_{Δ}, y = input LR_{baseline}; x = input *prec*, y = input LR_{Δ}. Recovery of LR_{baseline} was reasonably well recovered, showing a clear linear relationship between input and recovered values. However, under conditions of low input decision precision (*prec*), predictions were more influenced by the precision parameter, leading to a weaker correlation between input and recovered $\text{LR}_{\text{baseline}}$.



(b)Recoverability of LR_{baseline} in High Volatility High Noise



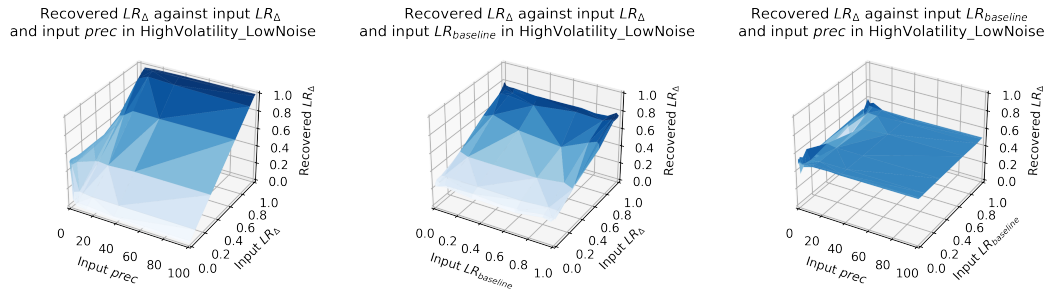
(c)Recoverability of LR_{baseline} in Low Volatility Low Noise



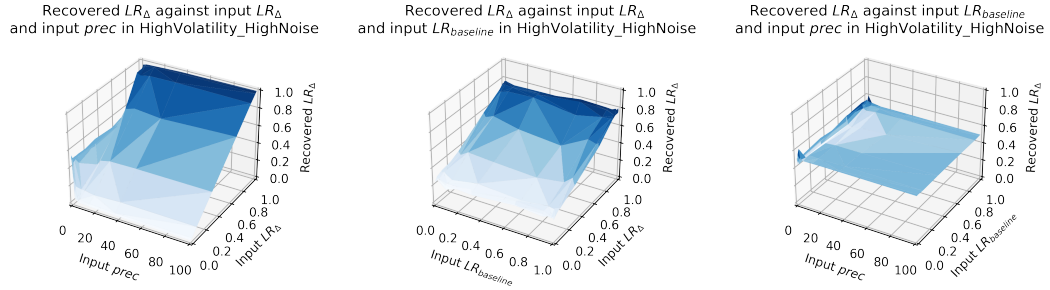
(d)Recoverability of LR_{baseline} in Low Volatility High Noise

Figure 3.8: Parameter recovery for $\mathbf{LR_{baseline}}$.
Parameter recovery for the baseline learning rate ($\mathrm{LR_{baseline}}$) was conducted separately for each block. Across all four blocks, $\mathrm{LR_{baseline}}$ was reasonably well recovered, showing a clear linear relationship between input and recovered values. However, under conditions of low input decision precision (*prec*), predictions were more influenced by the precision parameter, leading to a weaker correlation between input and recovered $\mathrm{LR_{baseline}}$. The input $\mathrm{LR_{\Delta}}$ affected the range of recoverable $\mathrm{LR_{baseline}}$ values: recovery was capped at 0.8 and failed to capture values below 0.2. Nonetheless, a consistent linear relationship persisted across all input $\mathrm{LR_{\Delta}}$ values. No systematic relationship was observed between recovered $\mathrm{LR_{baseline}}$ and either input precision or input $\mathrm{LR_{\Delta}}$. Different shades of blue represent variations in the recovered parameter values, with lighter shades indicating lower values and darker shades indicating higher values.

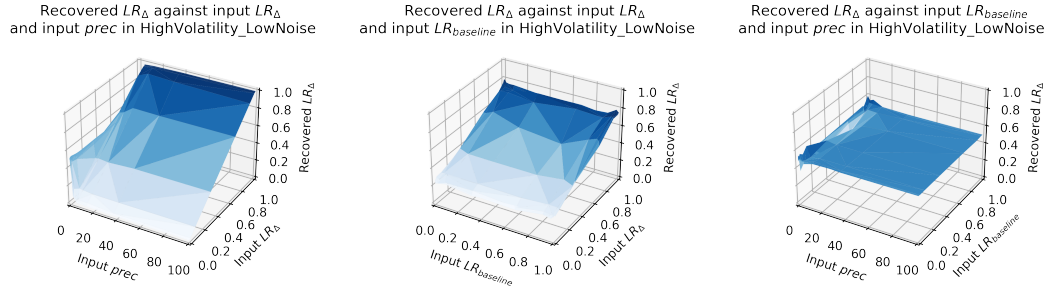
We evaluated parameter recovery of the simplified trend learning rate, LR_{Δ}, in each block (Figure 3.9). Across all four blocks, recovered LR_{Δ} tracked the true value approximately linearly, indicating generally good recoverability. Two qualifications emerged: (i) when decision precision was low (*prec* small), the correlation between input and recovered LR_{Δ} weakened, consistent with reduced recoverability; and (ii) across input LR_{baseline}, the dynamic range of recovered LR_{Δ} was compressed (capped at ≈ 0.7), even though the linear relation with the input parameter persisted. Beyond these boundary and precision effects, we found no evidence of a systematic dependence of recovered LR_{Δ} on input *prec*, or input LR_{baseline}.



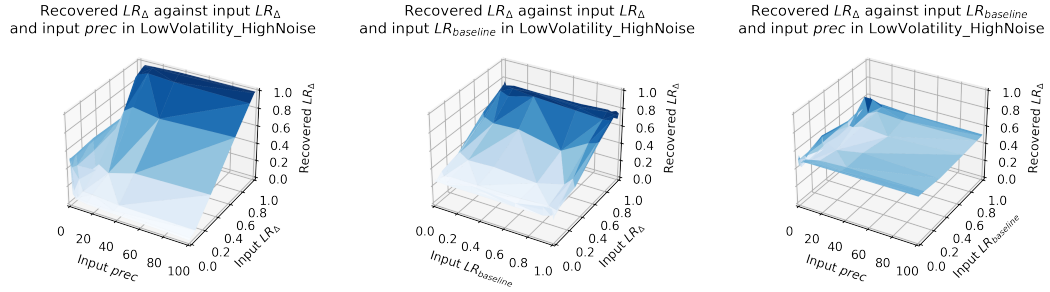
(a)Recoverability of LR_Δ in High Volatility Low Noise.
The recovered LR_Δ (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input *prec*, y = input LR_Δ; x = input LR_{baseline}, y = input LR_Δ; x = input *prec*, y = input LR_{baseline}. Recovery of LR_Δ was reasonably well recovered, showing a clear linear relationship between input and recovered values. However, under conditions of low input decision precision (*prec*), predictions were more influenced by the precision parameter, leading to a weaker correlation between input and recovered LR_Δ.



(b)Recoverability of LR_Δ in High Volatility High Noise



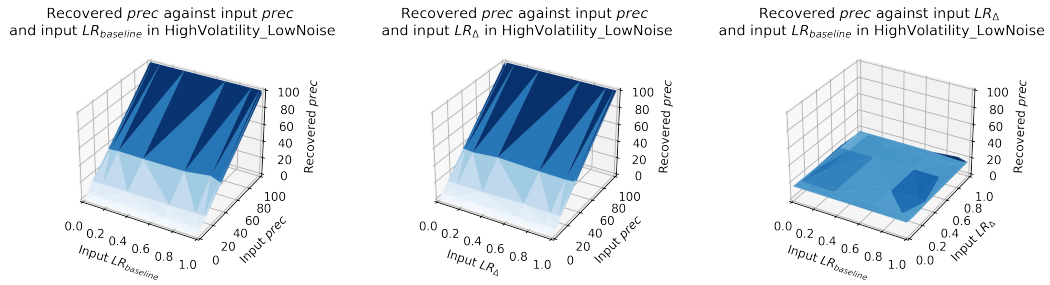
(c)Recoverability of LR_Δ in Low Volatility Low Noise



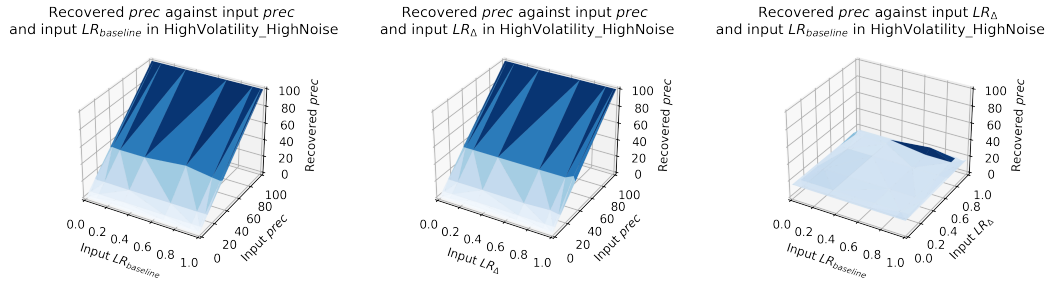
(d)Recoverability of LR_Δ in Low Volatility High Noise

Figure 3.9: Parameter recovery for LR_{Δ}.
Parameter recovery for the trend learning rate (LR_{Δ}) was conducted separately for each block. Across all four blocks, LR_{Δ} was reasonably well recovered, showing a clear linear relationship between input and recovered values. However, under conditions of low input decision precision (*prec*), predictions were more influenced by the precision parameter, leading to a weaker correlation between input and recovered LR_{baseline}. The input α and LR_{baseline} affected the range of recoverable LR_{Δ} values: recovered values were capped at 0.7, though the linear relationship between input and recovered LR_{Δ} persisted across all LR_{baseline} levels. No systematic relationship was observed between recovered LR_{Δ} and input precision, or LR_{baseline}. The different shades of blue in the plots represent variations in the recovered parameter values, with lighter shades indicating lower values and darker shades indicating higher values.

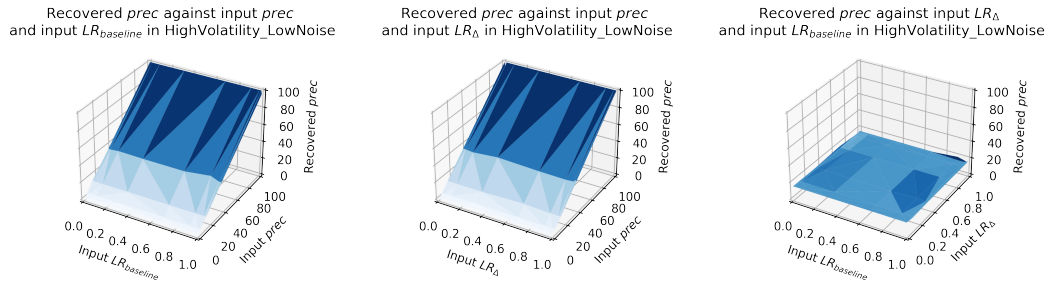
We assessed parameter recovery for decision precision (*prec*) in each block (Figure 3.10). Recovered *prec* tracked the true input values approximately linearly across all blocks, indicating good recoverability. Crucially, recovered *prec* showed no systematic dependence on the learning-rate parameters LR_{Δ} or LR_{baseline}; when *prec* was held constant, the recovery surfaces were essentially flat along the learning-rate axes. Together, these results suggest that *prec* can be estimated reliably with minimal confounding from LR_{Δ} and LR_{baseline}.



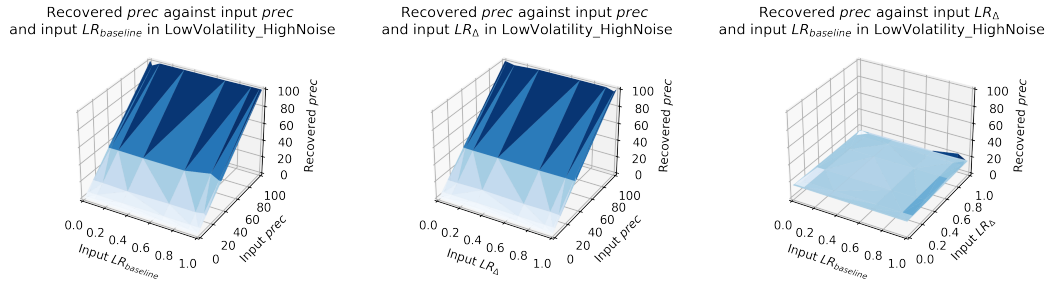
(a)Recoverability of *prec* in High Volatility Low Noise.
The recovered *prec* (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input LR_{baseline}, y = input *prec*; x = input LR_{Δ}, y = input *prec*; x = input LR_{baseline}, y = input LR_{Δ}. Recovery of *prec* was reasonably well recovered, showing a clear linear relationship between input and recovered values. Recovered *prec* was not related to input learning rates (LR_{Δ} and $\text{LR}_{\text{baseline}}$).



(b)Recoverability of *prec* in High Volatility High Noise



(c)Recoverability of *prec* in Low Volatility Low Noise



(d)Recoverability of *prec* in Low Volatility High Noise

Figure 3.10: Parameter recovery for *prec*.
Parameter recovery for decision precision (*prec*) was conducted separately for each block. Across all four blocks, a clear linear relationship was observed between input and recovered *prec* values. No systematic relationship was found between input learning rates (LR_{Δ} and $LR_{baseline}$) and recovered *prec*. The different shades of blue in the plots represent variations in the recovered parameter values (z-axis), with lighter blues indicating lower values and darker shades indicating higher values.

Recoverability Improvement

Recoverability of $LR_{baseline}$

To evaluate whether removing the mixture parameter improved the recoverability of $LR_{baseline}$ in the simplified trend-learning model, we divided LR_{Δ} into four equal bins and calculated the Pearson correlation between the input $LR_{baseline}$ and the recovered $LR_{baseline}$. There is no consistent or significant improvement observed in the recoverability of $LR_{baseline}$ in simplified trend learning model (see Table 3.2).

Block	Input LR_{Δ} range	Corr (input $LR_{baseline}$, recovered $LR_{baseline}$)		Δ corr
		Simplified trend learning model	Trend learning model	
High Volatility High Noise	[0, 0.25]	0.752	0.746	0.006
	[0.25, 0.5]	0.781	0.742	0.039
	[0.5, 0.75]	0.752	0.740	0.012
	[0.75, 1]	0.741	0.744	-0.003
High Volatility Low Noise	[0, 0.25]	0.746	0.746	0.000
	[0.25, 0.5]	0.733	0.740	-0.007
	[0.5, 0.75]	0.743	0.741	0.002
	[0.75, 1]	0.744	0.743	0.001
Low Volatility High Noise	[0, 0.25]	0.683	0.696	-0.013
	[0.25, 0.5]	0.693	0.682	0.011
	[0.5, 0.75]	0.643	0.683	-0.040
	[0.75, 1]	0.690	0.688	0.002
Low Volatility Low Noise	[0, 0.25]	0.659	0.654	0.005
	[0.25, 0.5]	0.662	0.645	0.017
	[0.5, 0.75]	0.637	0.644	-0.007
	[0.75, 1]	0.651	0.647	0.004

Table 3.2: Comparison of recoverability for $LR_{baseline}$ under LR_{Δ}.
We compared recoverability of $LR_{baseline}$ between simplified trend learning model and trend learning model under various input LR_{Δ} ranges. Changes in correlation between input $LR_{baseline}$ and recovered $LR_{baseline}$ are close to 0. We thus did not observe any significant or consistent improvement in the simplified trend learning model for $LR_{baseline}$.

We also divided *prec* into three equal bins and calculated the Pearson correlation between the input LR_{Δ} the recovered LR_{Δ}. There is consistent and significant improvement observed in the recoverability of LR_{Δ} in simplified trend learning model (see Table 3.3).

Block	**Input *prec* range**	**Corr (input LR_{baseline}, recovered LR_{baseline})**		**Δ corr**
		Simplified trend learning model	**Trend learning model**	
High Volatility High Noise	[0, 33.67]	0.674	0.665	0.009
	[33.67, 67.33]	0.996	0.999	-0.003
	[67.33, 101]	1.000	1.000	0.000
High Volatility Low Noise	[0, 33.67]	0.664	0.665	-0.001
	[33.67, 67.33]	0.999	0.999	0.000
	[67.33, 101]	1.000	1.000	0.000
Low Volatility High Noise	[0, 33.67]	0.577	0.588	-0.011
	[33.67, 67.33]	0.995	0.998	-0.003
	[67.33, 101]	0.998	1.000	-0.002
Low Volatility Low Noise	[0, 33.67]	0.540	0.530	0.010
	[33.67, 67.33]	0.972	0.988	-0.016
	[67.33, 101]	0.990	0.997	-0.007

Table 3.3: Comparison of recoverability for LR_{baseline} under *prec*.
We compared recoverability of LR_{baseline} between the simplified trend learning model and the trend learning model across various input *prec* ranges. Correlation differences were minimal.

To evaluate whether removing the mixture parameter improved the recoverability of LR_{Δ} in the simplified trend-learning model, we divided LR_{baseline} into four equal bins and calculated the Pearson correlation between the input LR_{Δ} and the recovered LR_{Δ}. There is significant improvement observed in the recoverability of LR_{Δ} in simplified trend learning model (see Table 3.4).

Block	Input LR_{baseline} range	Corr (input LR_{Δ}, recovered LR_{Δ})		Δ corr
		Simplified trend learning model	Trend learning model	
High Volatility High Noise	[0, 0.25]	0.713	0.565	0.148
	[0.25, 0.5]	0.698	0.559	0.139
	[0.5, 0.75]	0.696	0.551	0.145
	[0.75, 1]	0.728	0.561	0.167
High Volatility Low Noise	[0, 0.25]	0.694	0.535	0.159
	[0.25, 0.5]	0.684	0.525	0.159
	[0.5, 0.75]	0.711	0.527	0.184
	[0.75, 1]	0.712	0.530	0.182
Low Volatility High Noise	[0, 0.25]	0.693	0.544	0.149
	[0.25, 0.5]	0.689	0.548	0.141
	[0.5, 0.75]	0.668	0.557	0.111
	[0.75, 1]	0.688	0.554	0.134
Low Volatility Low Noise	[0, 0.25]	0.634	0.524	0.110
	[0.25, 0.5]	0.622	0.537	0.085
	[0.5, 0.75]	0.654	0.539	0.115
	[0.75, 1]	0.624	0.541	0.083

Table 3.4: Comparison of recoverability for LR_{Δ} under LR_{baseline}.
We compared recoverability of LR_{Δ} between simplified trend learning model and trend learning model under various input LR_{baseline} ranges. Changes in correlation between input LR_{Δ} and recovered LR_{Δ} are around 0.15. Correlation between input LR_{Δ} and recovered LR_{Δ} in simplified trend learning model is closer to 1.

We divided *prec* into three equal bins and calculated the Pearson correlation between the input LR_{Δ} and the recovered LR_{Δ}. There is consistent and significant improvement observed in the recoverability of LR_{Δ} in simplified trend learning model (see Table 3.5).

Block	Input *prec* range	Corr (input LR_Δ, recovered LR_Δ)		Δ corr
		Simplified trend learning model	Trend learning model	
High Volatility High Noise	[0, 33.67]	0.625	0.468	0.157
	[33.67, 67.33]	0.997	0.822	0.175
	[67.33, 101]	1.000	0.852	0.148
High Volatility Low Noise	[0, 33.67]	0.604	0.433	0.171
	[33.67, 67.33]	0.999	0.793	0.206
	[67.33, 101]	1.000	0.794	0.206
Low Volatility High Noise	[0, 33.67]	0.585	0.463	0.122
	[33.67, 67.33]	0.992	0.799	0.193
	[67.33, 101]	0.998	0.816	0.182
Low Volatility Low Noise	[0, 33.67]	0.503	0.439	0.064
	[33.67, 67.33]	0.980	0.797	0.183
	[67.33, 101]	0.987	0.822	0.165

Table 3.5: Comparison of recoverability for LR_Δ under *prec*.
We compared recoverability of LR_Δ between simplified trend learning model and trend learning model under various input *prec* ranges. Changes in correlation between input LR_Δ and recovered LR_Δ are around 0.17. Correlation between input LR_Δ and recovered LR_Δ in simplified trend learning model is closer to 1.

3.2.6 Model Comparison

We employed both the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) to assess which model best fit the participants' choices (see Figure 3.11). Our analysis indicated that a simplified Rescorla-Wagner model provided the optimal fit across all blocks, with the exception of the high-volatility, low-noise condition. However, the differences in AIC and BIC values between models were negligible, suggesting that traditional model comparison methods might be insufficient for this project. Consequently, we report the results of both the alpha-beta filter and the simplified trend-learning model.

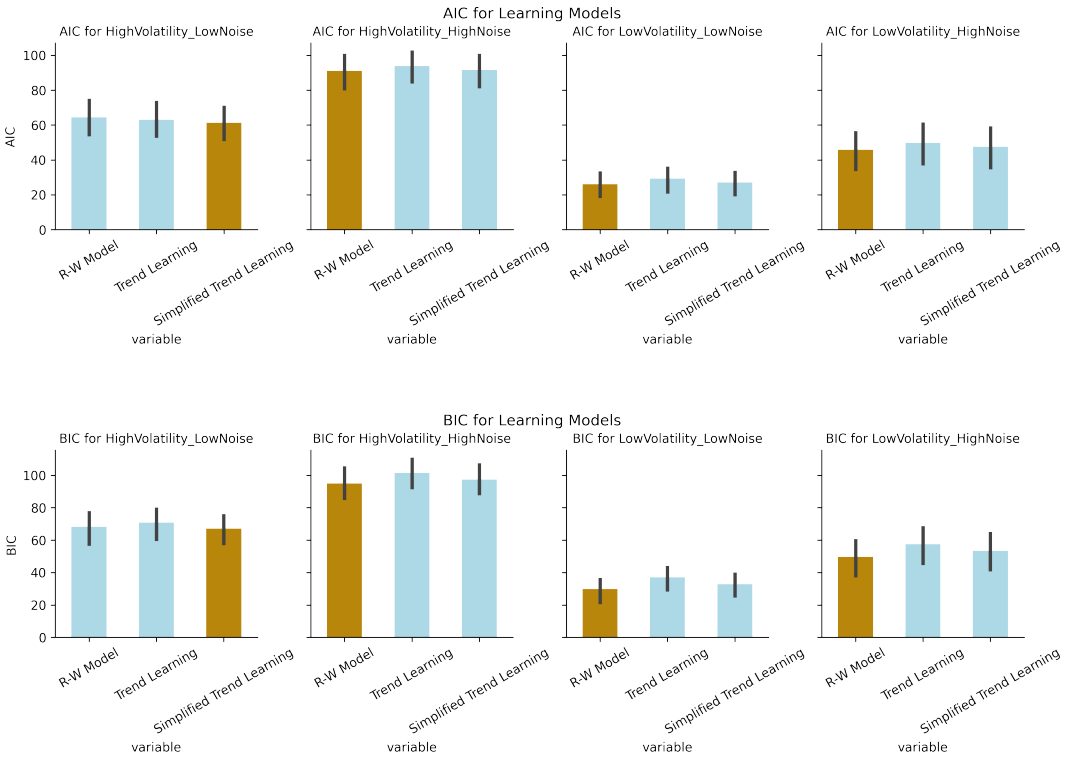


Figure 3.11: BIC and AIC of each model.
While both BIC and AIC yielded consistent results across models, neither measure demonstrated a consistent preference throughout all experimental blocks.

3.3 Results

In the Results section, we present the learning rates from both the Rescorla-Wagner model and the simplified trend-learning model. This decision was based on the fact that both models demonstrated reasonable recovery and neither the AIC nor BIC consistently favoured one model over the other across all experimental blocks.

3.3.1 Increased learning rate when volatility is high and noise is low

The learning rate (LR) indicates the extent to which participants' current estimations are influenced by recent observations compared to those further in the past. A higher learning rate implies that participants place greater weight on recent outcomes when updating their beliefs, while a lower learning rate suggests that recent outcomes

have less influence. We fitted a Rescorla-Wagner model to participants' choice data and estimated their learning rate for each block. Beyond capturing the belief updating process, we also modeled participants' decision precision, similar to the inverse decision temperature variable used in binary choice tasks. Decision precision reflects how accurately participants placed their bucket based on their estimations. A smaller precision parameter indicates more random and less precise bucket placement, suggesting lower confidence in their current estimations. For statistical analyses, model parameters were transformed to approximate a normal distribution: an inverse logit transformation was applied to learning rates, and decision precision was log-transformed.

Trial-wise Learning Rate

We first examined the effect of volatility and noise on adaptation of trial-wise learning rate. Our analysis showed that participants' learning rates were significantly higher in high-volatility blocks ($F(1, 41) = 41.109$, $p < 0.001$, Figure 3.12). However, there was no noise effect, as participants did not reduce learning rate when noise was high compared to when noise was low ($F(1, 41) = 0.008$, $p = 0.930$). In Chapter 2, we showed that average trial-wise learning rate by block have a number of shortcomings, including sensitivity to extreme LR values. Furthermore, while computing trial-wise learning rate, there is also the issue with change of belief.

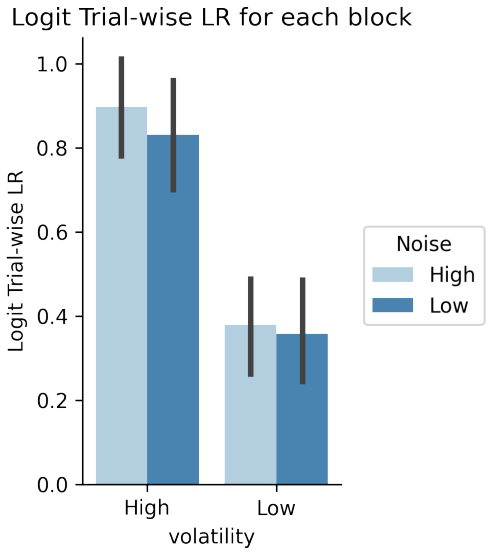


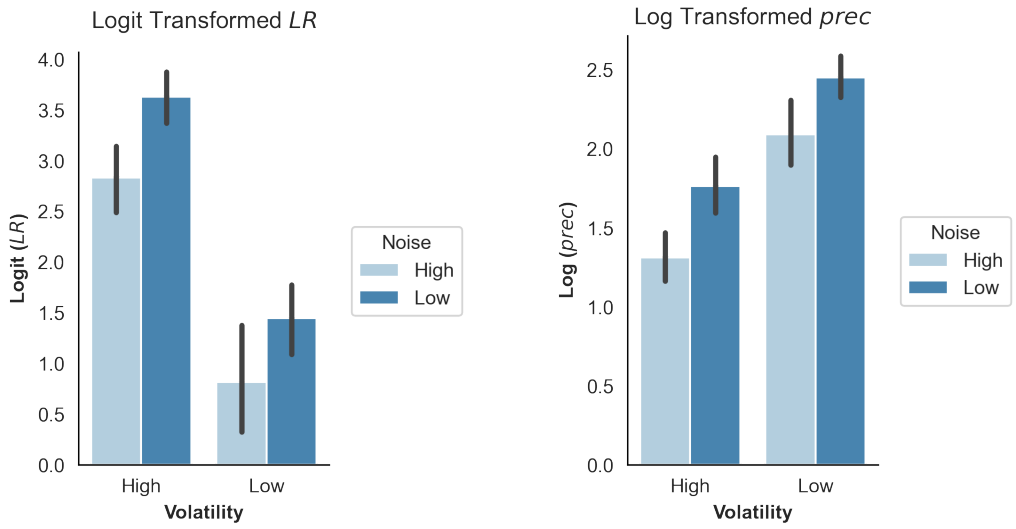
Figure 3.12: Estimates of participants' trial-wise learning rates across volatility and noise conditions.
Trial-wise learning rates were significantly higher in high-volatility blocks compared to low-volatility blocks, indicating participants adapted more strongly in volatile environments. However, there was no significant effect of noise on learning rate. Error bars represent ±1 SEM.

Rescorla-Wagner learning model

Our analysis showed that participants' learning rates were significantly higher in high-volatility blocks ($F(1, 41) = 175.665$, $p < 0.001$). Additionally, participants exhibited lower learning rates in high-noise blocks ($F(1, 41) = 18.239$, $p < 0.001$). These findings replicated the results from Chapter 2. Across participants, we observed a volatility effect, with decision precision being higher in low-volatility blocks ($F(1, 41) = 121.342$, $p < 0.001$). Participants also demonstrated greater prediction precision in low-noise blocks compared to high-noise blocks ($F(1, 41) = 35.837$, $p < 0.001$). No interaction effect was found between volatility and noise on decision precision ($p = 0.539$, see Figure 3.13).

To examine whether the order in which participants viewed the blocks influenced their subsequent adaptation, we re-ran the ANOVA including block order as a between subjects factor. Participants were classified as having seen two high volatility blocks first, two low volatility blocks first, or a mixed sequence. The analysis

revealed no significant main effect of block order on performance, $F(1, 40) = 1.80$, $p = .188$. This suggests that the sequence of blocks presented in the first half of the experiment, whether high volatility, low volatility, or mixed, did not affect participants' subsequent adaptation to task demands.



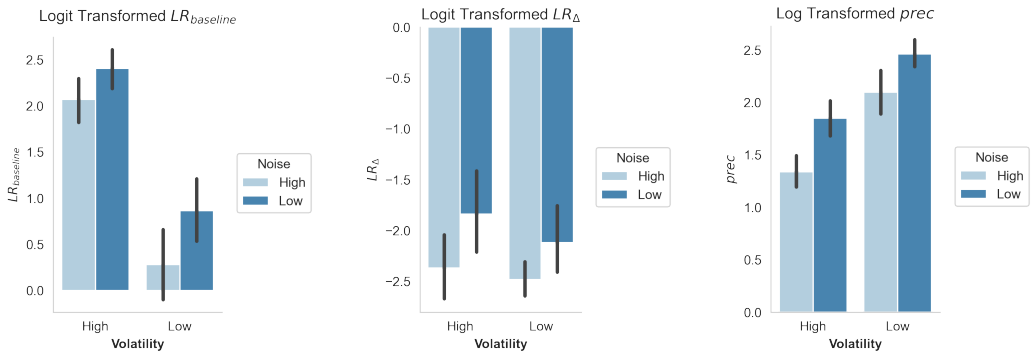
(a) **(b)**

Figure 3.13: Estimates of Participants' Learning Rates from Rescorla-Wagner Model.
We fitted a Rescorla-wagner learning model to participants' choices for each block. We applied logit and log transformation to learning rates and decision precision, respectively. The black vertical lines represent ±1 SEM of participant learning rates and decision precision. Participants showed higher learning rates in high volatile versus low volatile blocks ($F(1, 41) = 175.665, p < 0.001$) and showed lower learning rates in low noise versus high noise blocks ($F(1, 41) = 18.239, p < 0.001$).

Simplified Trend Learning Model

Consistent with the results from the Rescorla-Wagner model, participants exhibited a lower baseline learning rate in low-volatility blocks compared to high-volatility block ($F(1, 41) = 118.283$, $p < 0.001$) and high-noise blocks compared to low-noise blocks ($F(1, 41) = 12.343$, $p = 0.001$, Figure 3.14). Across participants, we observed an effect of volatility, with decision precision being higher in low-volatility blocks ($F(1, 41) = 121.146$, $p < 0.001$). Participants also demonstrated greater prediction precision in low-noise blocks ($F(1, 41) = 35.821$, $p < 0.001$). There was

no interaction effect between volatility and noise on decision precision ($p = 0.539$). Volatility did not significantly increase LR_Δ ($F(1, 41) = 2.36$, $p = 0.132$), although LR_Δ was reduced in low-noise blocks ($F(1, 41) = 8.821$, $p = 0.005$).



(a) **(b)** **(c)**

Figure 3.14: Estimates of participants' learning rates from simplified trend learning model.
We fitted a simplified trend learning model to participants' choices for each block. Learning rates and decision precision were logit- and log-transformed, respectively. Black vertical lines represent ±1 SEM. Participants showed higher baseline learning rates ($\text{LR}_{\text{baseline}}$) in high-volatility versus low-volatility blocks ($F(1, 41) = 118.283$, $p < 0.001$), and lower baseline learning rates in low-noise versus high-noise blocks ($F(1, 41) = 12.343$, $p = 0.001$).

Overall, these results indicate that both volatility and noise influence learning rates as hypothesized, independent of the specific used.

3.3.2 Adaptation to volatility is not modulated by Trait Anxiety and Depressive Symptoms

The results indicate that adaptation to volatility is not significantly influenced by trait anxiety or depressive symptoms. The analysis, which measured the difference in participants' learning rates between high- and low-volatility blocks, showed that neither trait anxiety nor depressive symptoms significantly modulated this adaptation ($r_s(40) = -0.167$, $p = 0.291$). Although a near-significant negative correlation was observed between trait anxiety and behavioral adaptation in the low-noise condition ($r_s(40) = -0.274$, $p = 0.079$), this did not reach statistical significance. Furthermore, using a simplified trend-learning model, no significant relationships were found between adaptation to volatility and CES-D scores or

trait anxiety ($p > 0.247$). These findings suggest that sensitivity to volatility and behavioral adjustments are not modulated by these psychological traits.

3.3.3 Adaptation to noise is not modulated by Trait Anxiety and Depressive Symptoms

Contrary to our hypothesis, no significant correlations were found between adaptation to noise and trait anxiety (high-volatility condition: $r_s(40) = -0.005$, $p = 0.976$; low-volatility condition: $r_s(40) = 0.113$, $p = 0.474$), nor between adaptation to noise and depressive symptoms (high-volatility condition: $r_s(40) = -0.096$, $p = 0.546$; low-volatility condition: $r_s(40) = 0.166$, $p = 0.292$). Using the simplified trend-learning model, there was also no correlation between the adaptation of learning rates to noise and either trait anxiety or CESD scores ($p > 0.538$).

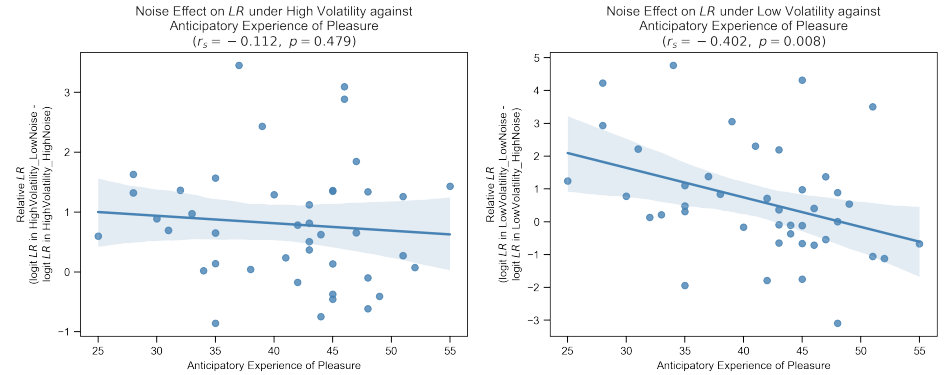
3.3.4 Exploratory evidence for a link between anticipatory pleasure and adaptation to noise

As part of an exploratory analysis, we examined whether anhedonic ability, measured by the anticipatory and consummatory subscales of the Temporal Experience of Pleasure Scale (TEPS), was related to the adaptation of the learning rate to noise and volatility. In the uncorrected results, we found a negative correlation between anticipatory TEPS and adaptation to noise under low-volatility conditions ($r_s(40) = -0.402, p = 0.008$; Fig. 3.15a), such that participants reporting greater anticipatory pleasure exhibited less adaptation to noise.

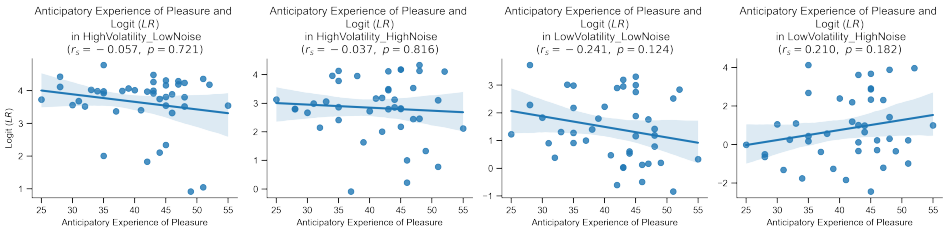
To better understand the source of this association, we examined block-wise learning rates. The correlation was specific to *changes* in learning rate between low-noise and high-noise blocks under low volatility; correlations with learning rate *within* each block were non-significant (high-volatility/low-noise: $r_s(40) = -0.046, p = 0.774$; high-volatility/high-noise: $r(42) = -0.048, p = 0.762$; low-volatility/low-noise: $r_s(40) = -0.045, p = 0.776$; low-volatility/high-noise: $r_s(40) = 0.254, p = 0.104$). These block-wise analyses were conducted as a localisation step and were not treated as independent hypothesis tests.

To ensure that this relationship was not sensitive to the choice of computational model, we repeated the correlation using learning-rate estimates from a simplified trend learning model. The negative association persisted ($r_s(40) = -0.357, p = 0.020$; Fig. 3.15c), suggesting that the finding was not driven by model specification.

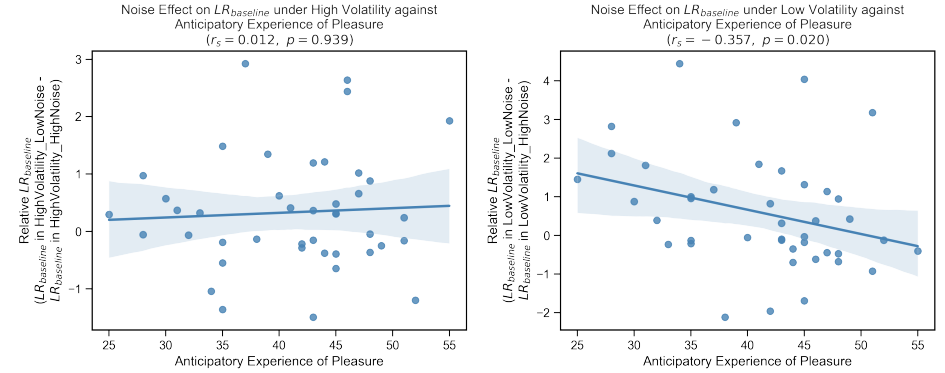
Finally, because both TEPS subscales were tested against four adaptation metrics (eight correlations in total), we applied a Holm–Bonferroni correction to control the familywise error rate at $\alpha = 0.05$. The anticipatory TEPS–noise adaptation correlation did not survive correction ($p_{\text{adj}} = 0.064$), indicating that this pattern should be interpreted cautiously as an exploratory finding, albeit one consistent across two computational models.



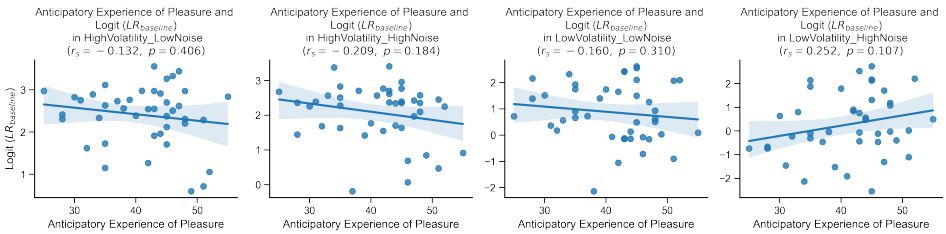
(a)



(b)



(c)



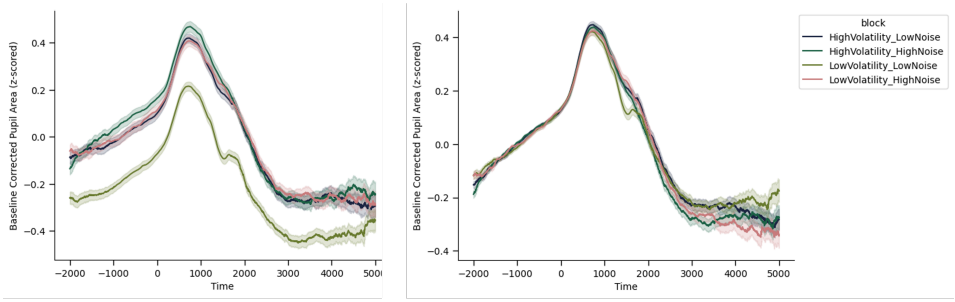
(d)

Figure 3.15: Adaptation to noise is modulated by anticipatory pleasure. Each row shows scatter plots relating model-derived learning parameters to anticipatory pleasure scores. The relative log learning rate for the high noise versus the low noise blocks was negatively correlated with anticipatory experience of pleasure in certain conditions. High anticipatory participants showed reduced adaptation to noise. Error bars indicate standard deviation of estimated parameters.

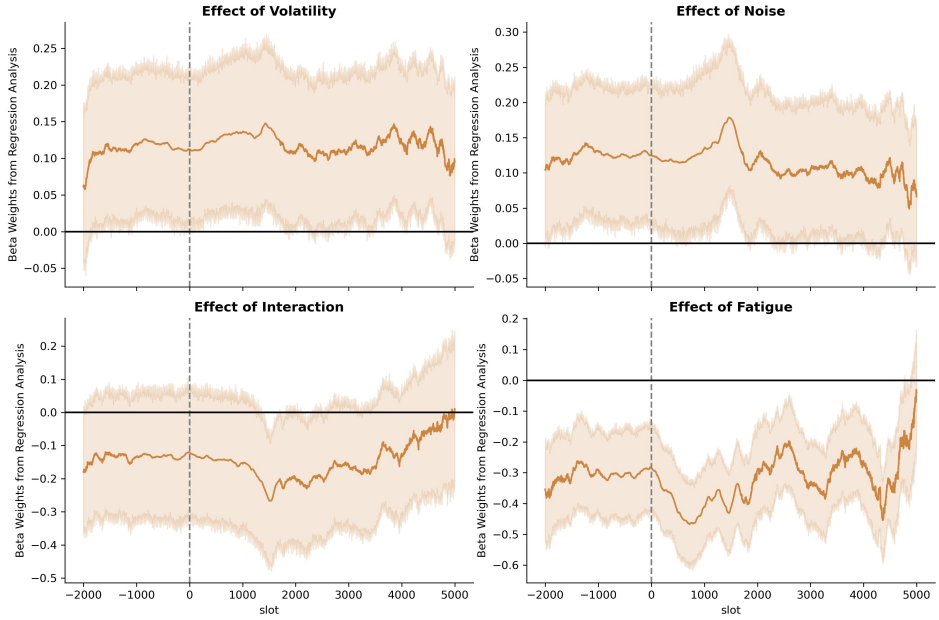
3.3.5 Changes in pupil size does not track volatility and noise

We initially plotted pupil diameter for each block across the 7-second window extracted for each trial. The results indicated that the effects of volatility and noise were present at the tonic level, while the baseline-corrected pupil dilation time series showed no differences across blocks. Therefore, we conducted linear regressions using baseline-uncorrected pupil diameter data. We found no main effect of volatility on pupil size ($t(41) = 1.123$, $p = 0.268$) and no main effect of noise ($t(41) = 1.167$, $p = 0.250$, Figure 3.16). Additionally, no significant interaction between volatility and noise was observed ($t(41) = -0.744$, $p = 0.461$). However, fatigue was found to significantly reduce pupil size across trials ($t(41) = -2.509$, $p = 0.016$).

Regression weights for volatility and noise did not show a correlation with behavioral adaptation to volatility and noise. Specifically, the relative learning rate (LR) between high-volatility and low-volatility blocks did not correlate with the mean effect of trial volatility on pupil dilation ($p = 0.372$). Similarly, the relative LR between low-noise and high-noise blocks did not correlate with the mean effect of trial noise on pupil size ($p = 0.225$). The same results were observed in the simplified trend-learning model, where behavioral adaptations were not associated with pupillary responses to volatility and noise ($p > 0.194$).



(a)



(b)

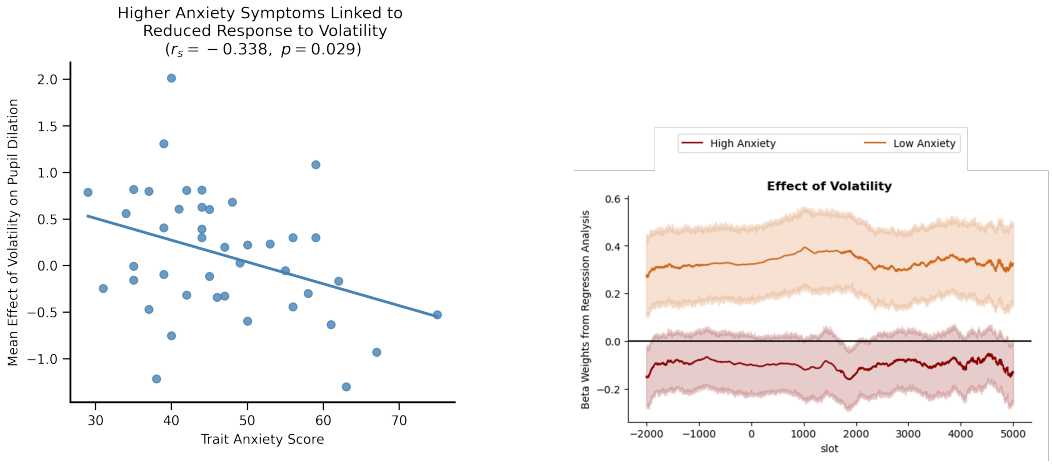
Figure 3.16: Pupil dilation does not track volatility nor noise.
(a) Z-scored pupil area and baseline-corrected pupil data from 2 seconds before to 5 seconds after outcome presentation. Lines illustrate average pupil size over time across task blocks. (b) Time course of the regression weights from a linear model predicting pupil dilation from volatility, noise, their interaction, and fatigue. Shaded areas represent ±1 standard deviation of the regression coefficients.

3.3.6 Higher trait anxiety, but not anticipatory experience of pleasure, is associated with reduced pupillary response to trial volatility

Higher levels of trait anxiety were associated with reduced regression weights for volatility ($r_s(40) = -0.338$, $p = 0.029$, see Figure 3.17a). Conversely, trait anxiety did not modulate the regression weights for noise, fatigue, or their interaction.

Participants were then divided into high-anxiety and low-anxiety groups based on the median trait anxiety score. In the low-anxiety group, volatility significantly increased pupil dilation, with beta weights across the 6-second window being significantly different from zero ($t(21) = 2.095$, $p = 0.049$, two-sided t-test, see Figure 3.17b). Based on the pupillometry result, we subsequently did a median split on relative learning rates (logit learning rates in high volatility blocks – logit learning rates in low volatility blocks) from the R-W model. Although the low-anxiety group had a slightly higher mean relative learning rate (M = 4.307, SD = 2.299) than the high-anxiety group (M = 4.086, SD = 1.823), a two-samples t-test indicated that the difference was not statistically significant, $t(40) = 0.345, p = .73$. Likewise, in the simplified trend learning model, the relative learning rates LR_{baseline} and LR_{Δ} did not differ (all $ps > .705$).

Anticipatory experience of pleasure is not associated with pupillary response to volatility ($p = 0.762$) or noise ($p = 0.515$).



(a) **(b)**

Figure 3.17: Tonic pupil dilation varies as a function of trait anxiety. (a) The degree to which participants' tonic pupil dilation tracked volatility was negatively related to trait anxiety, $r_s(40) = -0.338$, $p = 0.029$. (b) Using a median split on participants' trait anxiety scores, pupil dilation of low-anxious participants ($n = 22$) tracked trial volatility. Shaded error bands represent ±1 standard deviation of the regression coefficients.

3.3.7 Working memory capacity protects learning from fatigue

We also examined working memory capacity using the Operation Span Task (OSPAN). Higher working memory capacity was associated with a reduced impact of fatigue on pupil dilation ($r_s(40) = 0.401$, $p = 0.008$; see Figure 3.18). However, working memory capacity did not show significant associations with any measures of behavioral adaptation to volatility ($p > 0.086$) or noise across the tested models ($p > 0.072$).

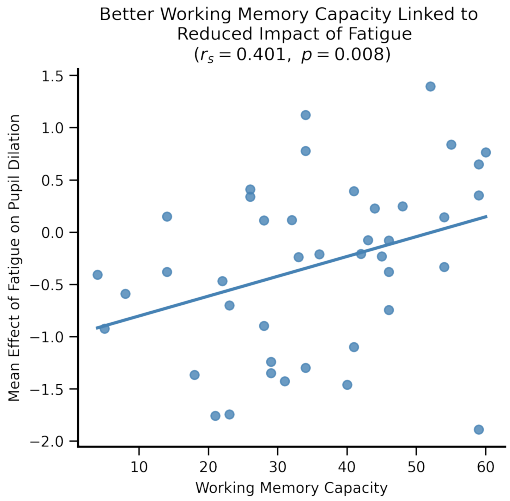


Figure 3.18: High working memory capacity is associated with smaller effects of fatigue.
Working memory capacity was measured using the Operation Span Task (OSPAN). Participants with higher working memory capacity experienced a reduced effect of fatigue on pupil dilation ($r_s(40) = 0.401$, $p = 0.008$).

3.4 Discussion

In this experiment, we showed that humans adapted their learning rates normatively by increasing learning rates in high-volatility environments and reducing them in high-noise conditions. However, contrary to predictions, pupil dilation did not directly track volatility or noise across all participants. Tonic pupil dilation only tracked volatility in individuals with low trait anxiety, suggesting that anxiety levels may influence physiological responses to environmental uncertainty. Importantly,

neither trait anxiety nor depressive symptoms significantly modulated behavioural adaptation to volatility and noise. We also observed a positive correlation between working memory capacity and effect of fatigue on pupil diameter.

Previous research has shown working memory is heavily involved in reinforcement learning (Collins & Frank, 2012). We thus measured working memory capacity to examine if it affects the ability to learn under uncertainty. At the behavioural level, we did not observe a significant effect of working memory capacity and adaptation to volatility and noise. However, better working memory capacity was found to be associated with a reduced effect of fatigue on pupil area. This aligns with Otto et al. (2013)'s findings, where higher working memory capacity protected against stress in goal-directed, model-based learning. The association between working memory capacity and reduced fatigue effects may reflect an individual's ability to allocate cognitive resources efficiently, mitigating the physiological toll of prolonged effort.

Contrary to our initial hypothesis, we observed a negative correlation between anticipatory TEPS and adaptation to noise under low-volatility conditions, consistent across two computational models. However, this association did not survive correction for multiple comparisons and should therefore be interpreted with caution.

A subsequent analysis of the relationship between learning rate and anticipatory TEPS by block revealed that individuals with higher hedonic ability showed lower learning rates in low volatility low noise condition. However, in low volatility high noise conditions, higher hedonic ability was associated with higher learning rates, suggesting an increased response to prediction errors caused by random fluctuations. This pattern raises the possibility that individuals with higher hedonic ability may misinterpret noise as an indicator of volatility and thus treat noise as if it was informative, leading to higher learning rates when true volatility is minimal but noise is high.

This observed relationship between anticipatory hedonic ability and adaptation to noise, although exploratory, aligns with Alloy and Abramson (1979)'s description of how depressed and non-depressed individuals perceive action-outcome contingency. Specifically, non-depressed individuals tend to overestimate action-outcome

contingency, particularly when the outcomes are desirable. In contrast, depressed individuals appear to perceive environmental contingencies more accurately. In one experiment, pressing a button led to an outcome (a green light) 75% of the time, and not pressing also resulted in the outcome appearing 75% of the time. Under this noncontingent condition, depressed participants correctly identified that their actions (pressing or not pressing) did not influence the outcome, accurately reflecting reality. Non-depressed participants, however, often overestimated their control in the same condition, perceiving some level of control despite the independence of outcomes from their actions.

It is worth noting that Alloy and Abramson (1979) observed these group differences using Beck's Depression Index (BDI-I, 1961). In our study, higher anticipatory pleasure, often associated with lower depression severity, was linked to a pattern of behaviour consistent with overestimating the informativeness of noise. This discrepancy could be related to the measures. The BDI-II and CES-D have distinct conceptual foundations (Choi et al., 2014). The BDI-II, rooted in cognitive-behavioral therapy for depression (Beck et al., 1979), places greater emphasis on cognitive symptoms (Demyttenaere & De Fruyt, 2003). In contrast, the CES-D was developed for epidemiological studies and prioritizes the affective dimensions of depression (Radloff, 1977). Since learning under uncertainty is a cognitive process, the BDI-II maybe more sensitive to this cognitive ability, as opposed to CES-D.

In this study, we did not observe an effect of volatility or noise on tonic or phasic pupil dilation. The tonic-level effect of volatility we observed aligns with prior research indicating that average pupil dilation signals volatility, whether through reversals (Pulcu & Browning, 2025) or change-point paradigms (Nassar et al., 2012). However, it contrasts with results showing phasic increases in pupil dilation in response to volatility (Browning et al., 2015; Lawson et al., 2017). This discrepancy could stem from differences in experimental design. Specifically, the schedule used in these experiments varies, and depending on how easily detectable environmental changes are, the LC-NE system may respond differently with tonic and phasic responses. In particular, sharp or highly noticeable changes (e.g., change-point,

odd-ball) might elicit phasic responses, as participants would need to quickly shift attention. This consistent with work by O'Reilly et al. (2013), which highlights a dissociation between surprise and model updating, as reflected in pupil diameter responses. O'Reilly et al. (2013) found that surprise was associated with an increase in pupil diameter, reflecting a phasic response to unexpected events that require immediate attention reallocation or reprogramming. In contrast, model updating was associated with a decrease in pupil diameter, appearing later than the response to surprise. This decrease suggests a different underlying mechanism tied to the adjustment of internal models rather than immediate attentional shifts.

We also observed that higher trait anxiety is associated with lower tonic level response to volatility. A median split based on trait anxiety show that tonic pupil dilation tracks volatility in low trait anxiety group. These findings regarding trait anxiety are also observed in the phasic response to volatility, as reported in Browning et al. (2015). Together, these results suggest that phasic and tonic pupil responses reflect different aspects of how anxiety influences perception and response to volatility. Phasic reductions, as observed in Browning et al. (2015) may indicate an immediate, momentary impairment due to anxiety, leading to reduced processing of sudden changes. In contrast, the reduction in tonic response observed in this study may imply that anxiety affects the ongoing, background level adaptability, potentially resulting in a chronic state of reduced cognitive flexibility and readiness.

3.5 Conclusion

In this chapter, we explored how individuals adapt to environmental volatility and noise, examining the role of pupil dilation and its modulation by trait anxiety, depressive symptoms, and anticipatory pleasure. Our findings are consistent with our hypotheses that participants increase their learning rates when volatility is high and noise is low.

Contrary to our hypotheses, neither trait anxiety nor depressive symptoms significantly influenced participants' adaptation to volatility or noise. However, our exploratory analyses indicated that higher anticipatory hedonic ability was

associated with reduced adaptation to both environmental factors. This pattern suggests that the extent to which individuals anticipate future enjoyment may be linked to their perceived controllability of the environment. Specifically, those who perceive greater control may expect to experience more pleasure in the future.

Pupil dilation did not directly track environmental volatility or noise, nor did it correlate with behavioral adaptations. Nonetheless, trait anxiety was associated with a reduced pupillary response to volatility, particularly in individuals with lower anxiety levels, where the dilation was less pronounced.

4

Adaptive Learning in Response to Momentum Under Noise

4.1 Introduction

In Chapters 2 and 3, we established that individuals adapt to environmental volatility in a normative manner, even in the presence of noise. Volatility, in our task, was operationalized as the variability of the means over time. However, the generative distributions underlying the data we observe may exhibit higher-order regularities in their mean. Specifically, the mean of these distributions might follow systematic and predictable patterns. For instance, cyclical patterns, such as annual average temperatures, result in periodic fluctuations in the mean of the generative distribution. Similarly, linear trends, like those observed in inflation rates, generate data that align with a steadily increasing mean.

We refer to such linear trends in volatility as momentum, which reflects the mean consistently shifting in a predictable direction over time. Recognizing these higher-order structures enables individuals to move beyond simply tracking a shifting mean, allowing them to incorporate the underlying linear structure of volatility in the data. This study specifically focused on the momentum and examines the extent to which individuals can identify and leverage such structural information to enhance their learning and predictions in environments characterized by structured volatility.

Previous literature on this topic is limited. To recapitualte Chapter 1, Ritz et al. (2018) demonstrated that when individuals are presented with linear trends, they integrate information over time to track gradual changes in the trend using a proportional-integrative-derivative controller (PID). We tested this model in supplementary analyses. Eldar et al. (2016) proposed a model similar to the alpha-beta filter (reviewed in Chapter 1) but did not provide empirical data to validate it.

This study is the first to directly manipulate momentum in an experimental task under noisy conditions and systematically test a range of computational models to capture how people learn in such environments. We hypothesized that humans increase their learning rates in response to momentum in the environment, with pupil dilation increasing during momentum blocks compared to no momentum blocks. Additionally, we aimed to replicate previous findings on noise, which suggest that humans reduce their learning rates when noise levels are high. We further explored how symptoms of mood disorders might influence adaptation to momentum in linearly structured volatile environments with noise. Findings from Chapter 3 indicate that higher trait anxiety is associated with less tonic pupil diameter in response to volatility. Previous research has also shown that trait anxiety is linked to a reduced ability to adapt to changes in the environment (Browning et al., 2015). Building on these findings, we hypothesized that symptoms of mood disorders would similarly be associated with diminished adaptation of learning rates and pupillary responses to momentum.

4.2 Methods

The experimental procedure was largely the same as in our previous study described in Chapter 3. Participants first completed an online survey that included the CES-D, demographic information, and questions about their mental health. Those meeting the selection criteria were invited to the Warneford Hospital for a 1-hour testing session. After providing informed consent, participants received instructions for the coin-catching task. The task blocks were presented in random order, and the task itself lasted approximately 25 minutes. Following the coin-catching task, participants

completed the Automated Operation Span Task (AOSPAN) to assess working memory capacity and a series of questionnaires, including the trait anxiety subscale of the STAI, WSAS, and TEPS. The entire experimental session lasted about 1 hour.

4.2.1 Sample

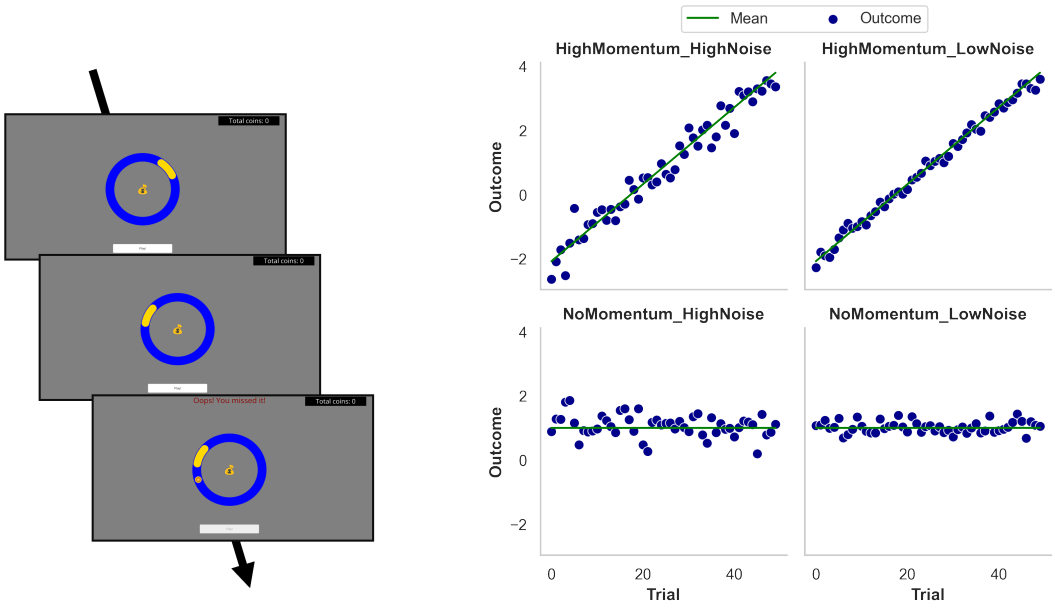
Fifty (N female = 36) English speaking, healthy participants, aged between 19 and 73 (mean = 32.74, S.D = 13.06), were recruited from Oxford, UK area. As in the previous project, recruitment was conducted via Facebook Ads targeted by geolocation to users whose registered location fell within a 15 km radius of the Oxford Botanic Garden. The advertisement appeared in users' news feeds and directed interested individuals to the study page. Pre-screening on Center for Epidemiologic Studies Depression Scale (CESD) was conducted in order to achieve an approximately the same number of participants with scores in the ranges 0-10, 10-20, >20; invitations were sent irrespective of score provided the quota for that range had not yet been filled. Exclusion criteria included current receipt of psychoactive medication or psychological therapy and neurological illness. Analyzable eye tracking data was obtained on all but six of the remaining participants. Six participants were excluded due to having over 80% missing pupillometry data, resulting in a final sample of 44 participants with analyzable eye-tracking data (Table 4.1).

Depression Group	Screened (CESD-based)	Recruited	Discarded	Final Sample
High (CESD > 20)	128	17	0	17
Medium (10 < CESD ≤ 20)	68	18	3	15
Low (CESD ≤ 10)	35	15	3	12
Total	**231**	**50**	**6**	**44**

Table 4.1: Sample stratified by depression severity.
Participants were prescreened to ensure coverage across symptom severity. In total, 231 individuals completed the prescreen; 50 were recruited. Participation rates (the share of prescreened individuals within each CES-D range who accepted an invitation and completed the study) differed by group. In particular, participation was highest in the low CES-D group (0 to 9; 41%), lower in the medium group (10 to 19; 26%), and lowest in the high group (20; 13%). This pattern was not intentional; although fewer low CES-D individuals were in the prescreening pool, they were more likely to accept and complete. Six participants were excluded due to more than 80% missing pupillometry data, yielding a final sample of 44.

4.2.2 Coin Catching Task

The experiment design was the same as in our previous study described in Chapter 3, except for the schedule (Figure 4.1). The four blocks are high momentum low noise, high momentum high noise, no momentum low noise and no momentum high noise. Overall, the task comprises four blocks of 50 trials. We used a von Mises biased random walk to generate the schedule. As before, the coin landing position was drawn from $\theta_i \sim \text{von Mises}(\mu_i, k_{\text{noise}})$ with μ_i changed every trial in high momentum blocks and μ_i keeping constant every trial in no momentum blocks. The k_{noise} was set to 9 for high noise blocks and 42 for low noise blocks. The mean of the distribution in high momentum blocks was determined by $\mu_i = \mu_{i-1} + v_t$, where v_t was step that the mean take at each trial and was determined by a second von Mises distribution $v_t \sim \text{von Mises}(a, 0)$. In high momentum blocks, a was set to 0.12. In no momentum blocks, a was set to 0. Setting the a to 0.12 made sure that the average of all steps v_t would be 0.12, and thus, an overall clockwise trend in the schedule. This gave a clear linear structure in the unexpected uncertainty, which allowed us to test the extent of which people could detect and learn such secondary structure and used it to improve their predictions. k_{noise} introduced expected uncertainty into the environment. Introducing both unexpected and expected uncertainty in the environment require people to differentiate between meaningful changes caused by momentum and those random signals generated by noise.



(a) (b)

Figure 4.1: Task Structure: Coin catching task.
(a) Example trial. Participants were instructed to place the yellow bucket at a location where they believed the coin would most likely land. After positioning the bucket, a coin was released from the center of the circle (taking 20 ms) and landed on the blue track. If the coin did not land in the yellow bucket (as shown in this example), a red message reading "Oops! You missed it" appeared at the top of the screen, and the total coin counter (located in the top right corner of the screen) remained unchanged. The outcome was displayed for a variable duration between 2 seconds and 7 seconds. (b) Experiment schedule. Coin's landing position for each block, generated by a biased von Mises random walk. In the two high momentum blocks, the means of the distribution (green lines) changed every trial in both clockwise directions. Meanwhile, the means of two no momentum blocks remained constant. The landing positions (blue dots) in high noise clocks were further away from the means, while low noise blocks the positions were closer to the means.

4.2.3 Model-free Analysis

We analysed trial-by-trial learning rates by calculating them as the ratio between the change in belief (the estimated bucket position) and the prediction error.

$$\mathrm{LR}_t = \frac{\text{Change of belief}_t}{\text{Prediction error}_t} \tag{4.2.3.1}$$

Since the momentum direction was defined as clockwise, we analysed learning rates separately for prediction errors aligned with this direction ($\mathrm{LR}_{\mathrm{same}}$) versus those

in the anticlockwise direction (LR_{opp}). If participants were taking the momentum into consideration, they would be expected to be more influenced by the clockwise PE than by anticlockwise PE of the same magnitude. In other words, learning rates for prediction errors in the same direction as the momentum (i.e., clockwise) would be higher compared to those in the opposite direction as the momentum (i.e., anticlockwise). This provides a relatively simple, model free metric of whether people are influenced by momentum.

To assess adaptation to momentum using model-free learning rates, we estimated learning rates for each block by computing the difference between the learning rate for PEs in the same direction as momentum and the learning rate for PEs in the opposite direction within a block

$$\text{Diff} = \text{LR}_{\text{same}} - \text{LR}_{\text{opp}} \tag{4.2.3.2}$$

.

To quantify the extent of individual adaptation to momentum, we calculated the difference between these block-level differences in the High Momentum Low Noise and No Momentum Low Noise conditions,

$$\text{Adaptation}_{\text{low noise}} = \text{Diff}_{\text{high momentum low noise}} - \text{Diff}_{\text{no momentum low noise}} \tag{4.2.3.3}$$

Adaptation to momentum under high noise conditions were obtained using the same logic,

$$\text{Adaptation}_{\text{high noise}} = \text{Diff}_{\text{high momentum high noise}} - \text{Diff}_{\text{no momentum high noise}} \tag{4.2.3.4}$$

Overall adaptation to momentum is the summation of how much one adapts under high noise and low noise,

$$\text{Adaptation} = \text{Adaptation}_{\text{low noise}} + \text{Adaptation}_{\text{high noise}} \tag{4.2.3.5}$$

4.2.4 Computational Modelling

We investigated various learning models, including the Rescorla-Wagner model, the simplified trend-learning model, and the alpha-beta filter, to examine how humans learn and incorporate linear volatility into their predictions. The alpha-beta filter, in particular, has been proposed as an alternative framework for describing trend learning. Additionally, we explored the proportional-integrative-derivative (PID) controller as another potential model (see Supplementary Material). This section outlines the models themselves, as well as our methodologies for parameter estimation, parameter recovery, and model recovery. Finally, we present evidence of positive model selection based on data features identified through model-free analysis.

Rescorla-Wagner Learning Model

The model follows the same implementation and design as the one introduced in Chapter 3.

Trend Learning Model

The model follows the same implementation and design as the one introduced in Chapter 3.

Simplified Trend Learning Model

The model follows the same implementation and design as the one introduced in Chapter 3. The figure below illustrates the model's ability to replicate participant choices by comparing participant data with those generated by the model (see Figure 4.2).

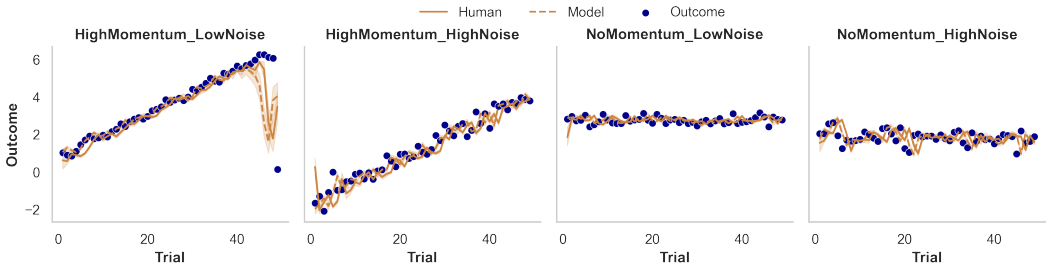


Figure 4.2: Comparison of Model-Simulated and Participant Behavioral Data. The plot shown here overlays the choice behaviour of participants (brown solid line) with the behaviour generated using model-fitted parameters (brown dash line) and the output (blue dots). The model output was generated by running the simplified trend learning model for each participant using their individually fitted parameter values on the task trial sequences. As shown, the model accurately reflects participant behaviours during the task.

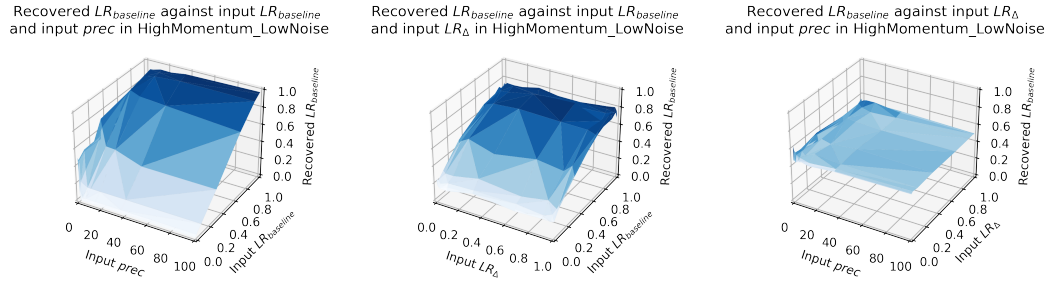
Parameter Recovery

In the previous chapter, we demonstrated that the simplified trend learning model could be successfully recovered using the schedule outlined in Chapter 3. To evaluate whether our parameter estimation procedure reliably reproduces parameter values under the schedule employed in Chapter 4, we first defined the parameter ranges for the simplified trend learning model.

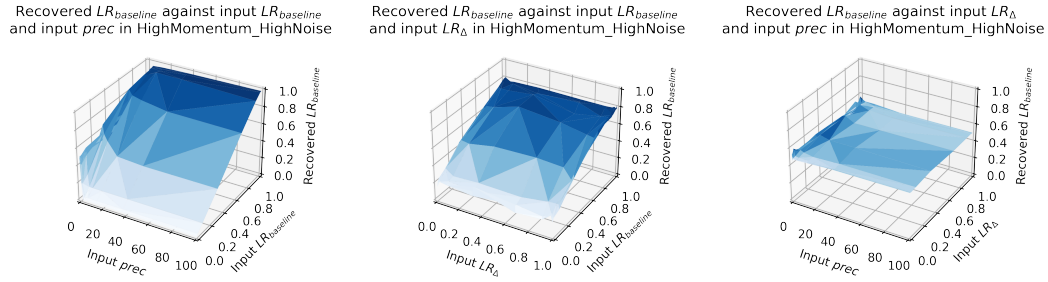
Each of the three free parameters was divided into 10 equally spaced values, spanning the entire parameter range used in the estimation procedure. Using these parameter combinations, predictions were simulated for a total of 10,000 combinations per block. Since the simulations were conducted across four blocks, this resulted in a total of 40,000 simulations. The simplified trend learning model was then fitted to these simulated choices, and the resulting parameter estimates were compared to the original input parameters.

Overall, $LR_{baseline}$ and LR_{Δ} were reasonably well recovered, closely matching their input values (see Figure 4.3 and 4.4). However, the recovery of both parameters was affected by low values of the precision parameter, *prec*. Specifically, under conditions of low input *prec*, predictions were more influenced by decision precision, resulting in a weaker correlation between the recovered and input learning

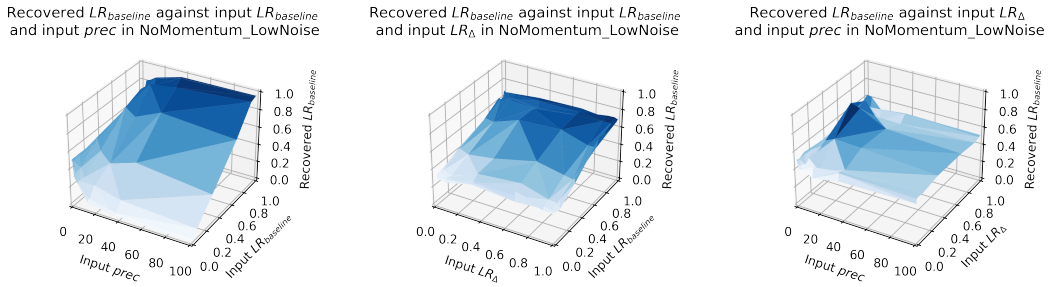
rates. The precision parameter (*prec*) exhibited strong recoverability, which was independent of the input values of the other two free parameters (Figure 4.5).



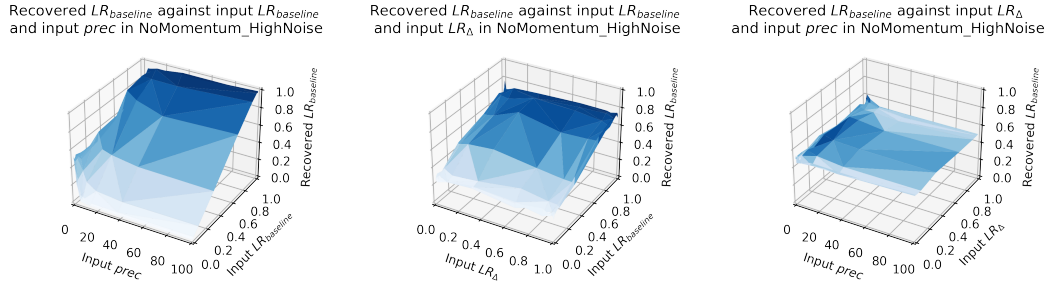
(a)Recoverability of LR_{baseline} in High Momentum Low Noise.
The recovered LR_{baseline} (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input *prec*, y = input LR_{baseline}; x = input LR_{Δ}, y = input LR_{baseline}; x = input *prec*, y = input LR_{Δ}. Recovery of LR_{baseline} was reasonably well recovered, showing a clear linear relationship between input and recovered values. However, under conditions of low input decision precision (*prec*), predictions were more influenced by the precision parameter, leading to a weaker correlation between input and recovered $\text{LR}_{\text{baseline}}$.



(b)Recoverability of LR_{baseline} in High Momentum High Noise



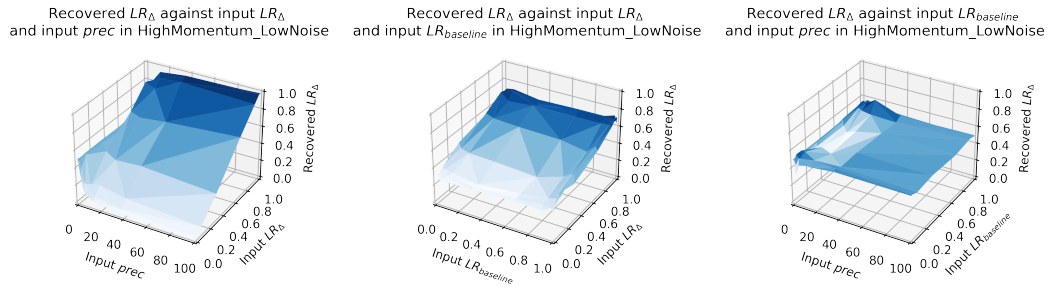
(c)Recoverability of LR_{baseline} in No Momentum Low Noise



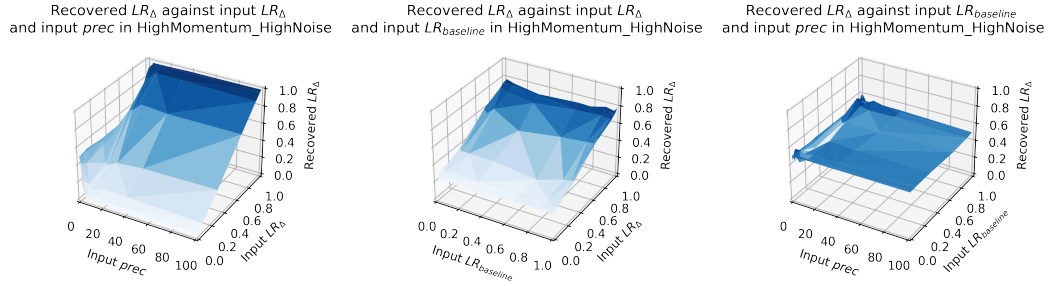
(d)Recoverability of LR_{baseline} in No Momentum High Noise

Figure 4.3: Parameter recovery for baseline learning rate ($LR_{baseline}$). Parameter recovery was conducted for each block separately. Across all four blocks, $LR_{baseline}$ was reasonably well recovered, showing a clear linear relationship between input $LR_{baseline}$ and recovered $LR_{baseline}$. However, under conditions of low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered $LR_{baseline}$ and input $LR_{baseline}$. Input LR_{Δ} affected the range of recoverable $LR_{baseline}$ values. Specifically, the recovered $LR_{baseline}$ was capped at 0.8, though a clear linear relationship between the input and recovered $LR_{baseline}$ persisted across all input LR_{Δ} values. Additionally, there was no observed relationship between recovered $LR_{baseline}$ and input precision or input LR_{Δ}. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

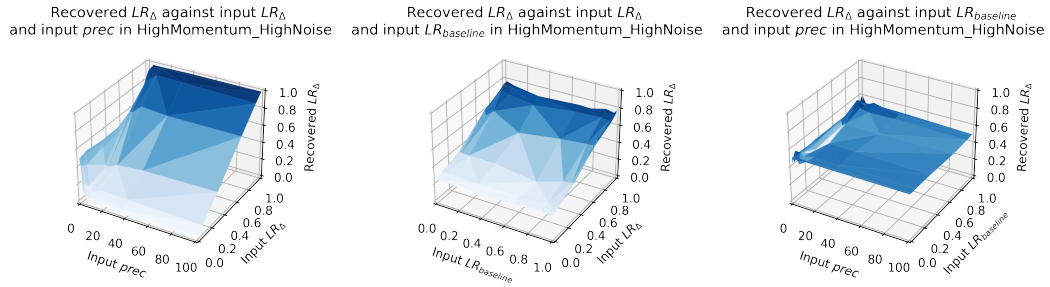
We examined recoverability of the trend learning rate, LR_{Δ}, using parameter–recovery analyses in each block (Figure 4.4). Across all four blocks, recovered LR_{Δ} mapped approximately linearly onto its generating value, indicating good identifiability. Similar to what we have found in Chapter 3 parameter recovery, two caveats emerged: (i) when decision precision was low (*prec* small), the association between input and recovered LR_{Δ} weakened; and (ii) across LR_{baseline}, the dynamic range of recovered LR_{Δ} was compressed (saturating near ~ 0.7), even though the linear trend with the input parameter remained. Aside from these precision and boundary effects, we found no systematic dependence of recovered LR_{Δ} on *prec*, or LR_{baseline}.



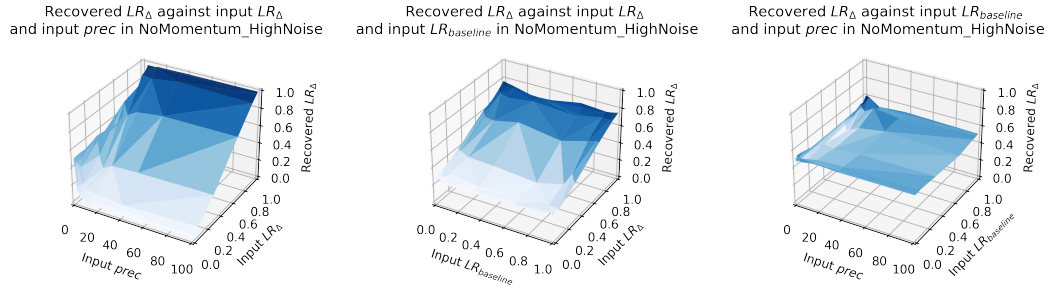
(a)Recoverability of LR_Δ in High Momentum Low Noise.
The recovered LR_Δ (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input *prec*, y = input LR_Δ; x = input LR_{baseline}, y = input LR_Δ; x = input *prec*, y = input LR_{baseline}. Recovery of LR_Δ was reasonably well recovered, showing a clear linear relationship between input and recovered values. However, under conditions of low input decision precision (*prec*), predictions were more influenced by the precision parameter, leading to a weaker correlation between input and recovered LR_Δ.



(b)Recoverability of LR_Δ in High Momentum High Noise



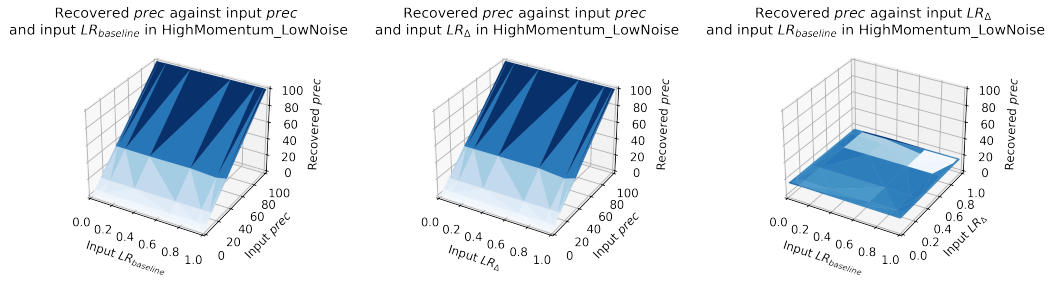
(c)Recoverability of LR_Δ in No Momentum Low Noise



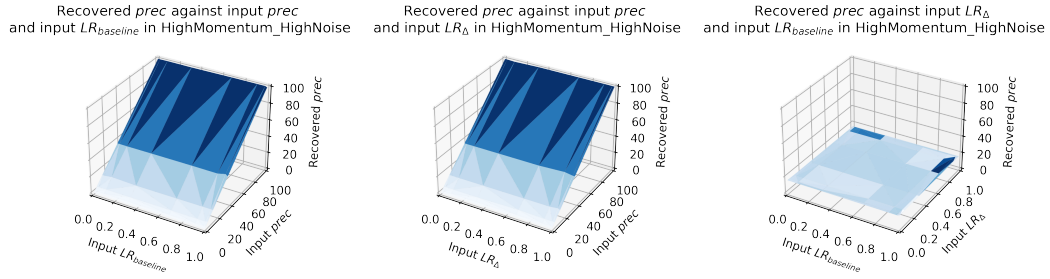
(d)Recoverability of LR_Δ in No Momentum High Noise

Figure 4.4: Parameter recovery for trend learning rate (LR_{Δ}).
Parameter recovery was conducted for each block separately. Across all four blocks, LR_{Δ} was reasonably well recovered, showing a clear linear relationship between input LR_{Δ} and recovered LR_{Δ}. However, under conditions of low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered LR_{Δ} and input LR_{Δ}. Input $LR_{baseline}$ affected the range of recoverable LR_{Δ} values. Specifically, the recovered LR_{Δ} was capped at 0.7, though a clear linear relationship between the input and recovered LR_{Δ} persisted across all input $LR_{baseline}$ values. Additionally, there was no observed relationship between recovered LR_{Δ} and input precision or input $LR_{baseline}$. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

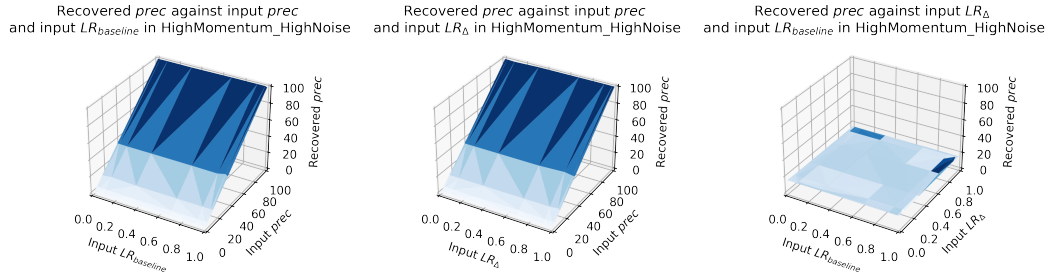
We examined recoverability of decision precision (*prec*) across blocks (Figure 4.5). Estimated *prec* scaled approximately linearly with its generating value in every block, indicating robust identifiability. Importantly, *prec* estimates were effectively orthogonal to the learning-rate parameters LR_{Δ} and $LR_{baseline}$: when *prec* was held fixed, the recovery surfaces were flat along the learning-rate axes. Taken together, these findings indicate that *prec* can be recovered reliably with minimal confounding from LR_{Δ} or $LR_{baseline}$.



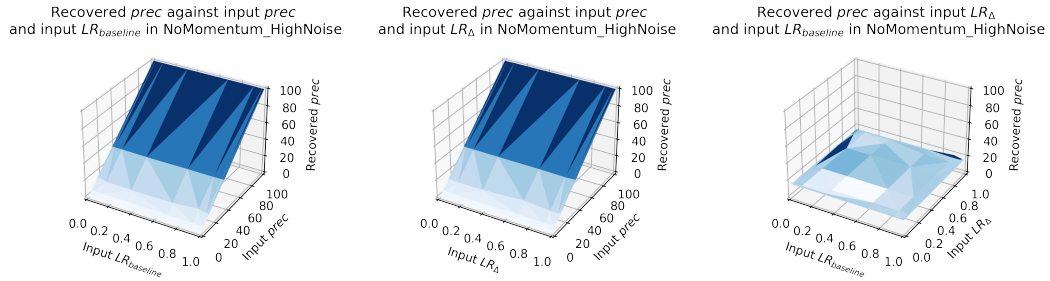
(a)Recoverability of *prec* in High Momentum Low Noise.
The recovered *prec* (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input LR_{baseline}, y = input *prec*; x = input LR_{Δ}, y = input *prec*; x = input LR_{baseline}, y = input LR_{Δ}. Recovery of *prec* was reasonably well recovered, showing a clear linear relationship between input and recovered values. Recovered *prec* was not related to input learning rates (LR_{Δ} and $\text{LR}_{\text{baseline}}$).



(b)Recoverability of *prec* in High Momentum High Noise



(c)Recoverability of *prec* in No Momentum Low Noise



(d)Recoverability of *prec* in No Momentum High Noise

Figure 4.5: Parameter recovery for decision precision (*prec*).
Parameter recovery was conducted for each block separately. Across all four blocks, a clear linear relationship was observed between input *prec* and recovered *prec*. No relationship was found between input learning rates (LR_{Δ} and $LR_{baseline}$) and recovered *prec*. The different shades of blue in these plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

Alpha-Beta Filter

The alpha-beta filter is traditionally used to estimate and predict the future state of a system based on noisy observations (Benedict & Bordner, 2003). It is an alternative model that people have used to describe trend learning. The filter operates with two main components: position estimation (controlled by the alpha parameter) and velocity estimation (controlled by the beta parameter). These two components allow the model to track both the current position and the trend (i.e., rate of change) in the data over time.

The process of which it updates its position is similar to the Rescorla-Wagner learning model, where the position estimate at each time step is updated by incorporating new observations. In particular, the alpha parameter, α, determines the rate at which the position estimate responds to new information.

$$\text{Belief}_{\theta_t} = \text{Belief}_{t-1} + \alpha \cdot PE \tag{4.2.4.1}$$

Where

$$\text{Belief}_{t-1} = \text{Belief}_{\theta_{t-1}} + \text{Belief}_{v_{t-1}} \tag{4.2.4.2}$$

$$PE = \text{Outcome}_{t-1} - \text{Belief}_{t-1} \tag{4.2.4.3}$$

In addition to position estimation, the alpha-beta filter also tracks the rate of change of the position, or the velocity, through trend learning. This process is governed by the beta parameter, β, which controls how quickly the model adapts to changes in the trend.

In addition to position estimation, the alpha-beta filter also tracks the rate of change of the position, or the velocity, through trend learning. This process is governed by the beta parameter, β, which controls how quickly the model adapts to changes in the trend.

$$\text{Belief}_{v_t} = \text{Belief}_{v_t} + \beta \cdot \left(\frac{\text{Outcome}_{t-1} - \text{Belief}_{t-1}}{\Delta t} \right) \tag{4.2.4.4}$$

With $\Delta t = 1$ in our experiment:

$$\text{Belief}_{v_t} = \text{Belief}_{v_t} + \beta \cdot PE \tag{4.2.4.5}$$

Belief about where the coin is going to land at trial t is a summation of the position estimate and velocity estimate:

$$\text{Belief}_t = \text{Belief}_{\theta_t} + \text{Belief}_{v_t} \tag{4.2.4.6}$$

After beliefs are updated, they are transformed into a probability distribution for predicting the coin's location. This is done using the von Mises probability density function, which models circular data (such as angles or periodic events) and includes a precision parameter to control the concentration of the distribution around the mean. The probability of choosing a given location is:

$$\begin{aligned} \Pr(\text{Prediction}_t \mid \text{Belief}_t, \text{prec}) &= f(\text{Prediction}_t \mid \text{Belief}_t, \text{prec}) \\ &= \frac{\exp(\text{precision} \cdot \cos(\text{choice} - \text{belief}))}{2\pi I_0(\text{precision})} \end{aligned} \tag{4.2.4.7}$$

Parameter Estimation

Learning rate parameters (α and β) were allowed to vary between 0.001 and 0.999, while decision precision (prec) was sampled over a range of 0.01 to 100. Each parameter's range was divided into 30 equally spaced discrete values. For the Alpha-beta model, which has 3 free parameters, we tested all possible combinations, resulting in 27,000 possibilities per block. Since the simulations were conducted across four blocks, this yielded a total of 81,000 simulations. The model was

simulated trial by trial based on each unique combination of parameter values within each block. The posterior probability was then calculated by multiplying the trial-wise likelihood of participants' choices, given the corresponding parameter combination. This process was repeated for each possible combination. The best estimate for each parameter was determined by computing the expected value of its marginalized probability distribution, summing over all other parameters. The expected value was extracted by summing the discrete parameter values weighted by their marginalized probabilities.

The figure below illustrates the model's ability to replicate participant choices by participant data with those generated by the model (see Figure 4.6).

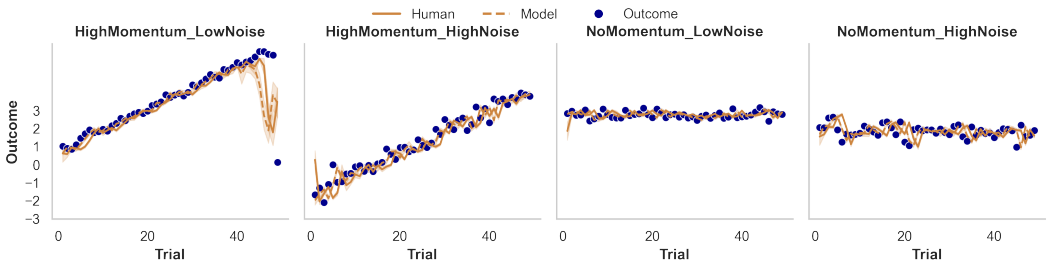
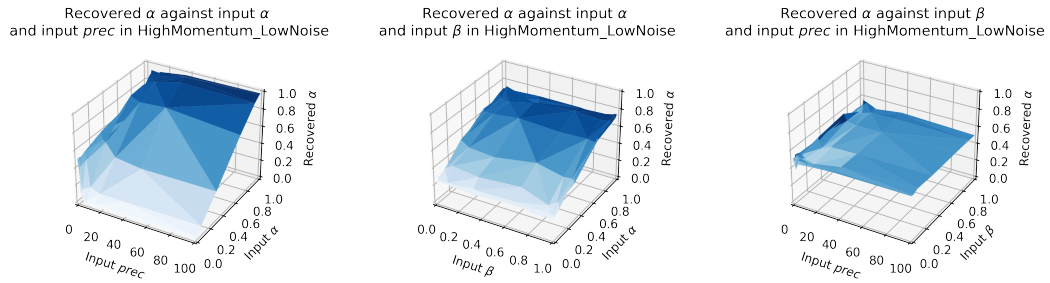


Figure 4.6: Comparison of Model-Simulated and Participant Behavioral Data. The plot shown here overlays the choice behaviour of participants (brown solid line) with the behaviour generated using model-fitted parameters (brown dash line) and the output (blue dots). The model output was generated by running the alpha-beta filter for each participant using their individually fitted parameter values on the task trial sequences. As shown, the model accurately reflects participant behaviours during the task.

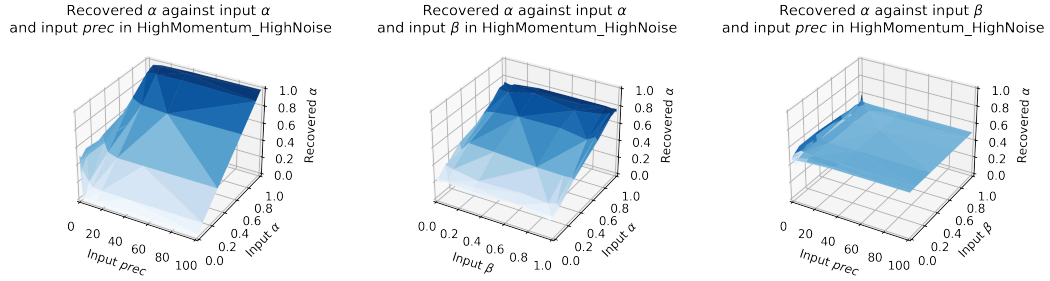
Parameter Recovery

Similar to the simplified trend learning model, the alpha-beta filter also had three free parameters. Each of these three parameters was divided into 10 equally spaced values, spanning the entire parameter range used in the estimation procedure. Using these parameter combinations, predictions were simulated for a total of 10,000 combinations per block. Since the simulations were conducted across four blocks, this resulted in a total of 40,000 simulations. The alpha-beta filter was then fitted to these simulated choices, and the resulting parameter estimates were compared to the original input parameters.

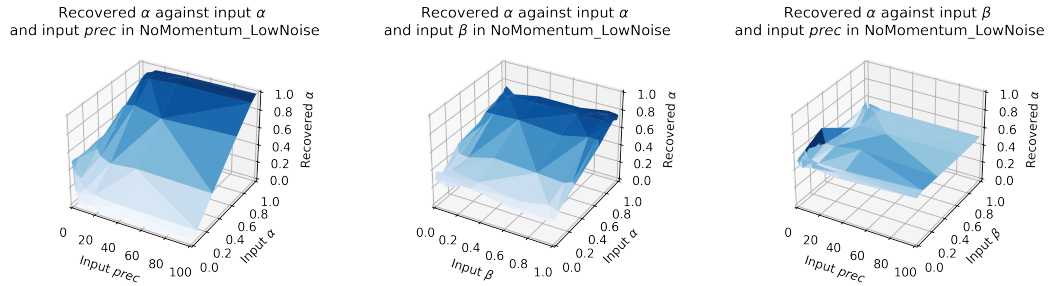
Overall, α and β were reasonably well recovered, closely matching their input values (see Figures 4.7 and 4.8). However, the recovery of all three parameters was affected by low values of the precision parameter, *prec*. Specifically, under conditions of low input *prec*, predictions were more influenced by decision precision, resulting in a weaker correlation between the recovered and input learning rates. The precision parameter (*prec*) exhibited strong recoverability, which was independent of the input values of the other two free parameters (Figure 4.9).



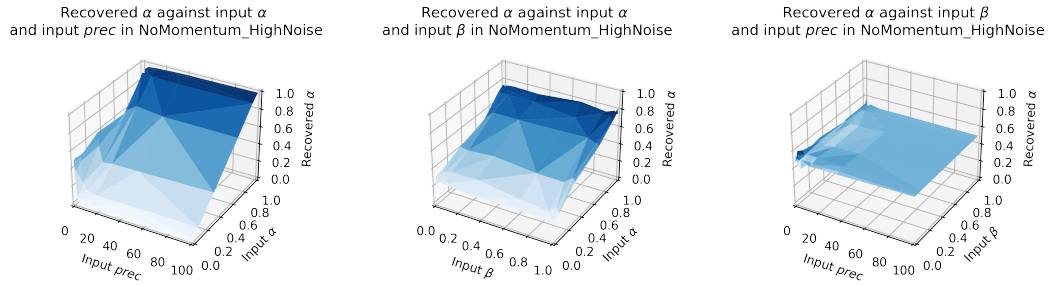
(a)Recoverability of α in High Momentum Low Noise.
The recovered α (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input *prec*, y = input α; x = input β, y = input α; x = input *prec*, y = input β. Recovery of α was reasonably well recovered, showing a clear linear relationship between input and recovered values. Recovery weakened when decision precision (*prec*) was low, and across input β the dynamic range of recovered α saturated near 0.8, though the linear trend persisted. Color shading from light to dark blue indicates low to high recovered α.



(b)Recoverability of α in High Momentum High Noise



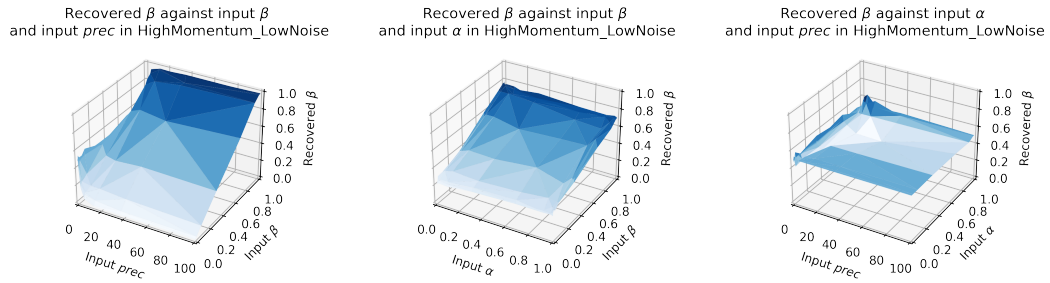
(c)Recoverability of α in No Momentum Low Noise



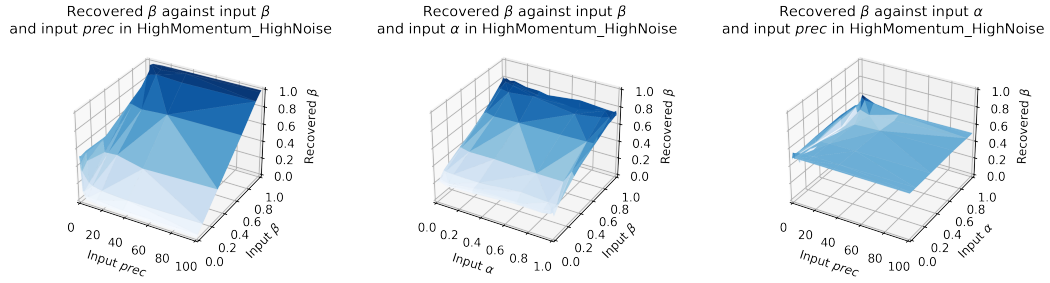
(d)Recoverability of α in No Momentum High Noise

Figure 4.7: Parameter recovery for α.
Parameter recovery for α was conducted for each block separately. Across all four blocks, α was reasonably well recovered, showing a clear linear relationship between input α and recovered α. However, under conditions of low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered α and input α. Input β affected the range of recoverable α values. Specifically, the recovered α was capped at 0.8, though a clear linear relationship between the input and recovered α persisted across all input β values. Additionally, there was no observed relationship between recovered α and input precision or input β. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

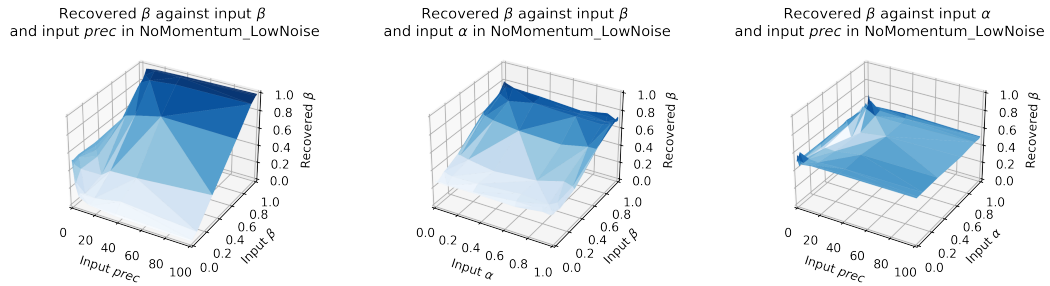
We evaluated recoverability of the trend parameter β in each block (Figure 4.8). Recovered β tracked the generating value approximately linearly across all blocks, indicating good identifiability. Two qualifications emerged: (i) when decision precision was low (*prec* small), the association between input and recovered β weakened; and (ii) across input α, the dynamic range of recovered β was compressed, saturating near ~ 0.7, although the linear trend with the input parameter persisted. Aside from these precision and boundary effects, we found no systematic dependence of recovered β on *prec* or α.



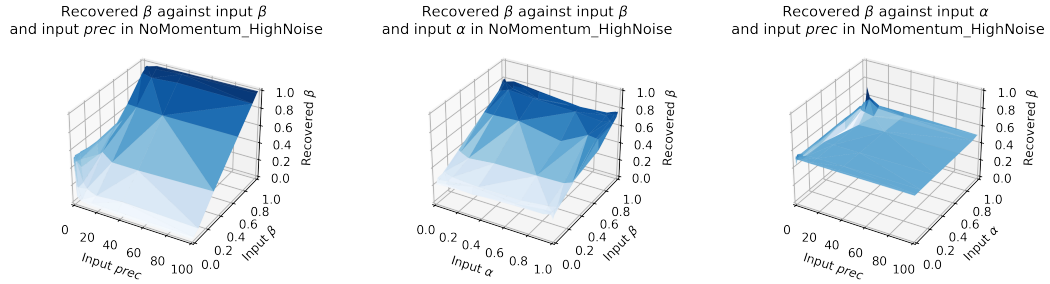
(a)Recoverability of β in High Momentum Low Noise.
The recovered β (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input *prec*, y = input β; x = input α, y = input β; x = input *prec*, y = input α. Recovery of β was reasonably well recovered, showing a clear linear relationship between input and recovered values. Similar to parameter recovery of α, weaker recovery at low decision precision (*prec*) and saturation near ~ 0.7 across input α.



(b)Recoverability of β in High Momentum High Noise



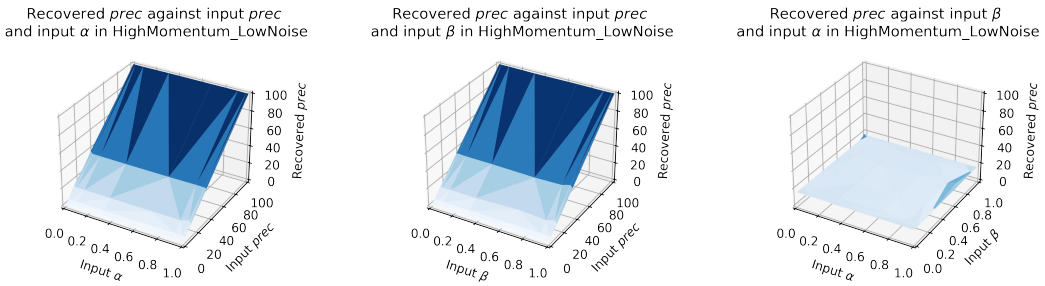
(c)Recoverability of β in No Momentum Low Noise



(d)Recoverability of β in No Momentum High Noise

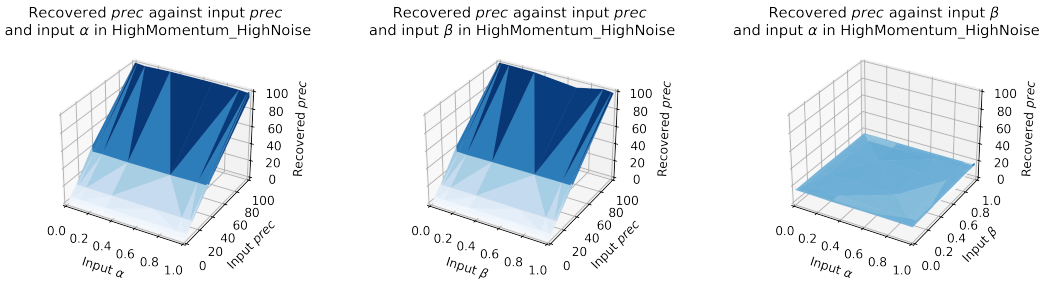
Figure 4.8: Parameter recovery for β.
Parameter recovery for β was conducted for each block separately. Across all four blocks, β was reasonably well recovered, showing a clear linear relationship between input β and recovered β. However, under conditions of low input precision (prec), predictions were more influenced by decision precision, leading to a weaker correlation between recovered β and input β. Input α affected the range of recoverable β values. Specifically, the recovered β was capped at 0.7, though a clear linear relationship between the input and recovered β persisted across all input α values. Additionally, there was no observed relationship between recovered β and input precision or input α. The different shades of blue in the plots represent variations in the z-axis values (i.e., recovered parameter values), with lighter blues indicating lower values and darker shades representing higher values.

We evaluated recovery of decision precision (*prec*) in each block (Figure 4.9). Recovered *prec* tracked its generating value approximately linearly across all blocks, indicating strong identifiability. Importantly, *prec* estimates showed no systematic dependence on the learning-rate parameters α or β; when *prec* was held fixed, the recovery surfaces were flat along the learning-rate axes. These results indicate that *prec* can be estimated reliably with minimal confounding from α and β.

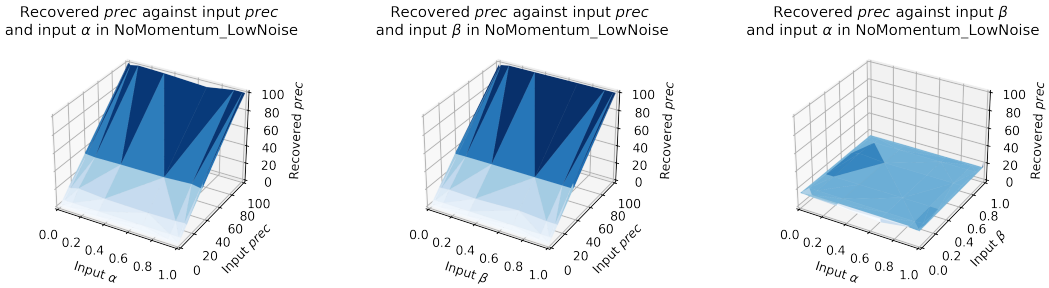


(a)Recoverability of *prec* in High Momentum Low Noise.
The recovered *prec* (z-axis) is shown for three input pairs arranged left, middle, and right, respectively: x = input α, y = input *prec*; x = input β, y = input *prec*; x = input α, y = input β. Recovery of $\prec$ was reasonably well recovered. In particular, a clear linear relationship was observed between input *prec* and recovered *prec*, while no relationship was found between input learning rates (α and β) and recovered *prec*

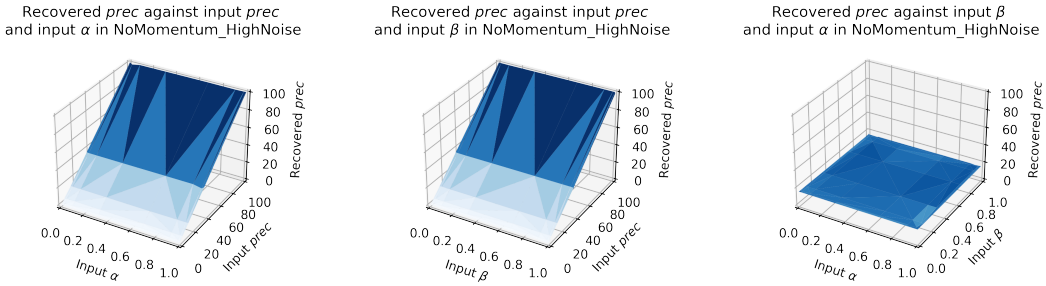
.



(b)Recoverability of *prec* in High Momentum High Noise



(c)Recoverability of *prec* in No Momentum Low Noise



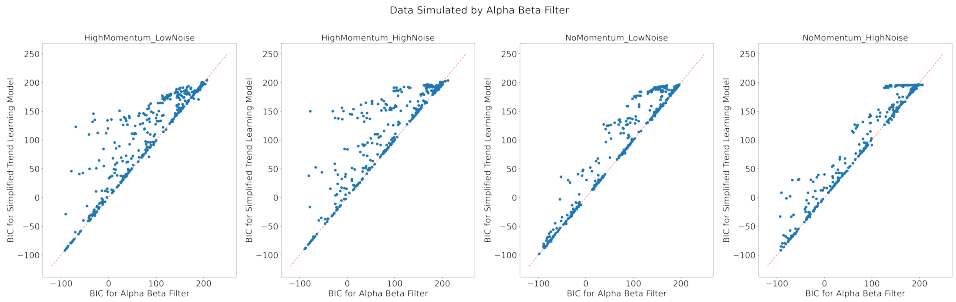
(d)Recoverability of *prec* in No Momentum High Noise

Figure 4.9: Parameter recovery for *prec*.
Parameter recovery for decision precision (*prec*) was performed separately for each block. Across all four blocks, a clear linear relationship was observed between input *prec* and recovered *prec*, while no relationship was found between input learning rates (α and β) and recovered *prec*. The different shades of blue in the plots represent variations in recovered values on the z-axis, with lighter shades indicating lower values and darker shades representing higher values.

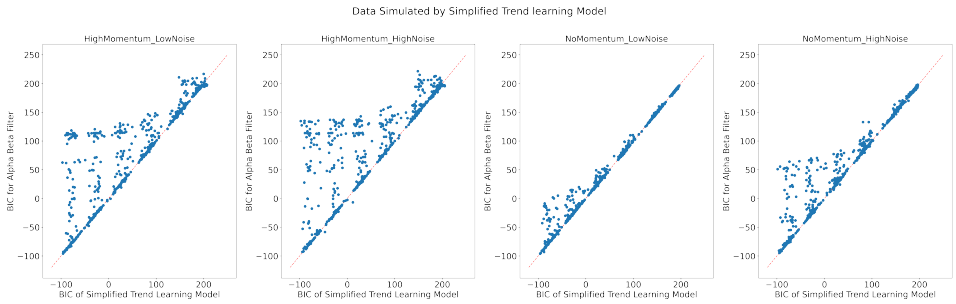
4.2.5 Model Recovery

The Alpha-Beta filter and the simplified trend learning model share a similar design. We summarized the key characteristic in Table 4.2. To test whether these two models are identifiable, we simulated synthetic data using both models. Each model includes three free parameters: two learning rates and one precision parameter. Data were generated in blocks, with 10 possible values assigned to each parameter, resulting in 1,000 parameter combinations per model for each block. We then estimated parameters for each dataset using both models and computed the Bayesian Information Criterion (BIC). Model recovery was assessed by checking whether the model used to estimate the parameters achieved a lower BIC when applied to the data it originally generated.

We compared the BIC values of alpha-beta filter and the simplified trend learning models to alpha beta filter generated data (Figure 4.10a) and simplified trend learning model generated data (Figure 4.10b). Scatterplots of BIC values demonstrate that each model is better suited to the synthetic data it generated. For alpha-beta filter-generated data, most points lie above the y = x line, indicating that the alpha-beta filter provides a better fit. Similarly, for simplified trend learning model-generated data, most points lie below the y = x line, favouring the trend learning model. While some data points are indistinguishable, these results confirm that each model is more likely to recover its own data.



(a)



(b)

Figure 4.10: Model Recovery: Alpha-Beta Filter vs. Simplified Trend-Learning Model.
(a) The scatterplot comparing the BIC values for the alpha-beta filter and the simplified trend learning model, applied to alpha-beta filter-generated data. While some data points lie directly on the y = x line, indicating that the BIC values for the two models are indistinguishable and thus cannot identify a clear preference, the majority of points lie above the y = x line. This pattern suggests that, for most cases, the alpha-beta filter provides a better fit to the data it generated, as indicated by the lower BIC values. These results support the conclusion that the alpha-beta filter is generally more likely to have generated the synthetic data compared to the simplified trend-learning model. (b) The scatterplot comparing the BIC values for the alpha-beta filter and the simplified trend learning model, applied to trend-learning model-generated data, shows a similar pattern. While some data points lie directly on the y = x line, indicating that the BIC values for the two models are indistinguishable and thus cannot identify a clear preference, the majority of points lie below the y = x line. This indicates that the simplified trend-learning model provides a better fit to the data it generated, as reflected by the lower BIC values. These results suggest that the trend-learning model is generally more likely to have generated the synthetic data compared to the alpha-beta filter.

		Simplified trend learning model *Outcome of the current trial* *Outcome of the trial before* *Prediction*	**Alpha-beta filter** *Outcome* *and prediction* *of the current trial*
Baseline learning / State Estimation	**Differences**	Baseline learning computes PE using the baseline belief alone, without incorporating the trend estimate.	State estimation already incorporates velocity from the last trial. Specifically, PEs for state estimation are computed based on the prior belief, which combines both position and velocity estimates.
	PEs	$PE_t = Outcome_{t-1} - Belief^{\theta}_{t-1}$	$PE_t = Outcome_{t-1} - Belief_{t-1}$ $= Outcome_{t-1} - (Belief^{\theta}_{t-1} + Belief^{v}_{t-1})$
Trend learning / Velocity Estimation	**Differences**	Learned through trial-to-trial changes in outcome position (Δ) with its own learning rate (LR_{Δ}). Trend is updated separately and combined with baseline learning only at the prediction step.	Estimated from the same PEs as state estimation using a separate learning rate for velocity (β).
	PEs	$PE_t = \Delta_{t-1} - Belief^{\Delta}_{t-1}$	$PE_t = Outcome_{t-1} - Belief_{t-1}$ $= Outcome_{t-1} - (Belief^{\theta}_{t-1} + Belief^{v}_{t-1})$

Table 4.2: Summary of key characteristics of simplified trend learning model and alpha-beta filter.
This table compares how baseline and trend/velocity estimates are computed in both models.

Positive Model Selection

Our primary interest in model selection was whether the models could replicate the participants' performance using the simple behavioural metric of asymmetric learning rates described above. A key behavioural signature of the task was a differential response to directed PEs. Specifically, we aimed to determine whether prediction errors (PE) aligned with the momentum direction resulted in higher learning rates compared to those generated from PEs in the opposite direction. This differential response to PEs highlights an important aspect of participants adaption of learning, which we present in greater detail in the Result section.

In our behavioural data, we observed these directional differences in trial-by-trial learning rates. To validate this finding, we conducted the same trial-wise learning rate analysis on model-generated choices using fitted parameters. Results showed that the alpha-beta filter, trend-learning, and simplified trend-learning models reproduced the essential behavioural features of the data. In contrast, the Rescorla-Wagner model did not, so we excluded it from further comparisons. (Figure 4.11).

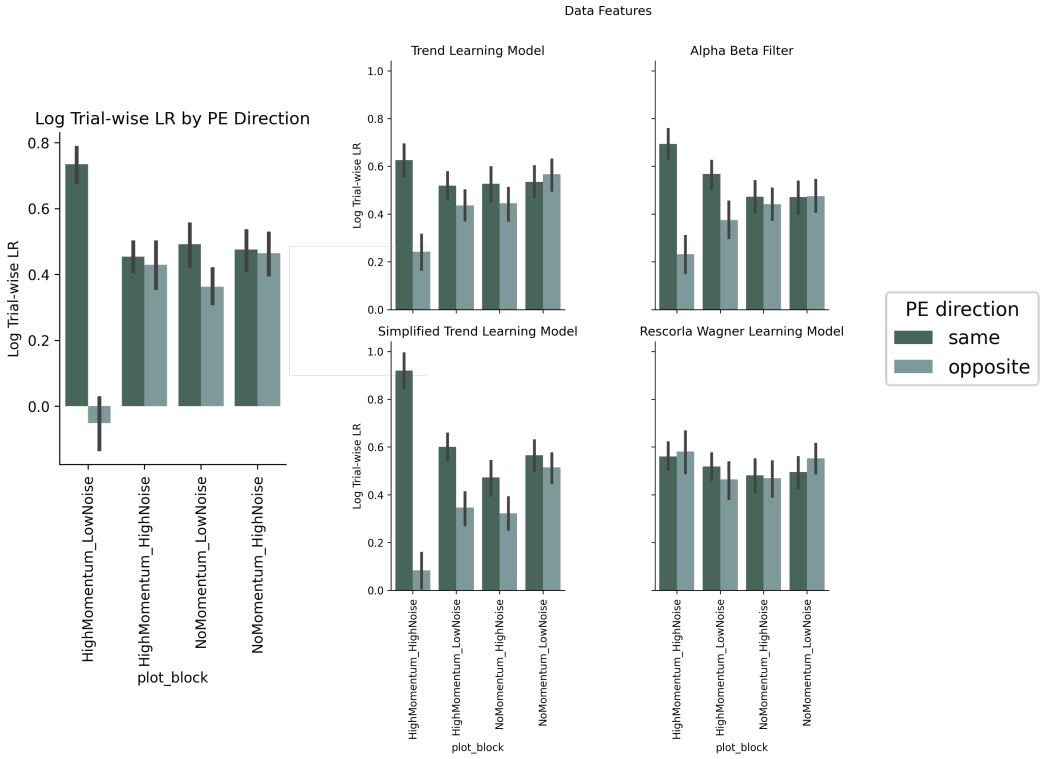


Figure 4.11: Comparison of trial-wise learning rates (LR) by prediction error (PE) direction for different models.
The behavioral data revealed an asymmetric response to PEs, with higher learning rates for PEs aligned with the momentum direction ("same") compared to those in the opposite direction ("opposite"). The Alpha-Beta Filter, Trend Learning Model, and Simplified Trend Learning Model successfully captured this asymmetric learning behavior observed in participants. Conversely, the Rescorla-Wagner model failed to replicate these key data features, leading to its exclusion from further comparisons. Bars represent mean log trial-wise LR, and error bars denote standard errors of the mean (SEM).

Model Comparison

The results suggest that the AIC and BIC do not provide a strong basis for preferring one model over the other. We used both the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) to evaluate which of the two models better fit participants' choices. Based on these metrics, the alpha-beta filter emerged as the winning model for all blocks except the no momentum low noise block. However, the absolute differences between AIC and BIC values were minimal, making these traditional model selection criteria insufficient for drawing robust conclusions (Figure 4.12). Consequently, we report the results from both the

alpha-beta filter and the simplified trend-learning model.

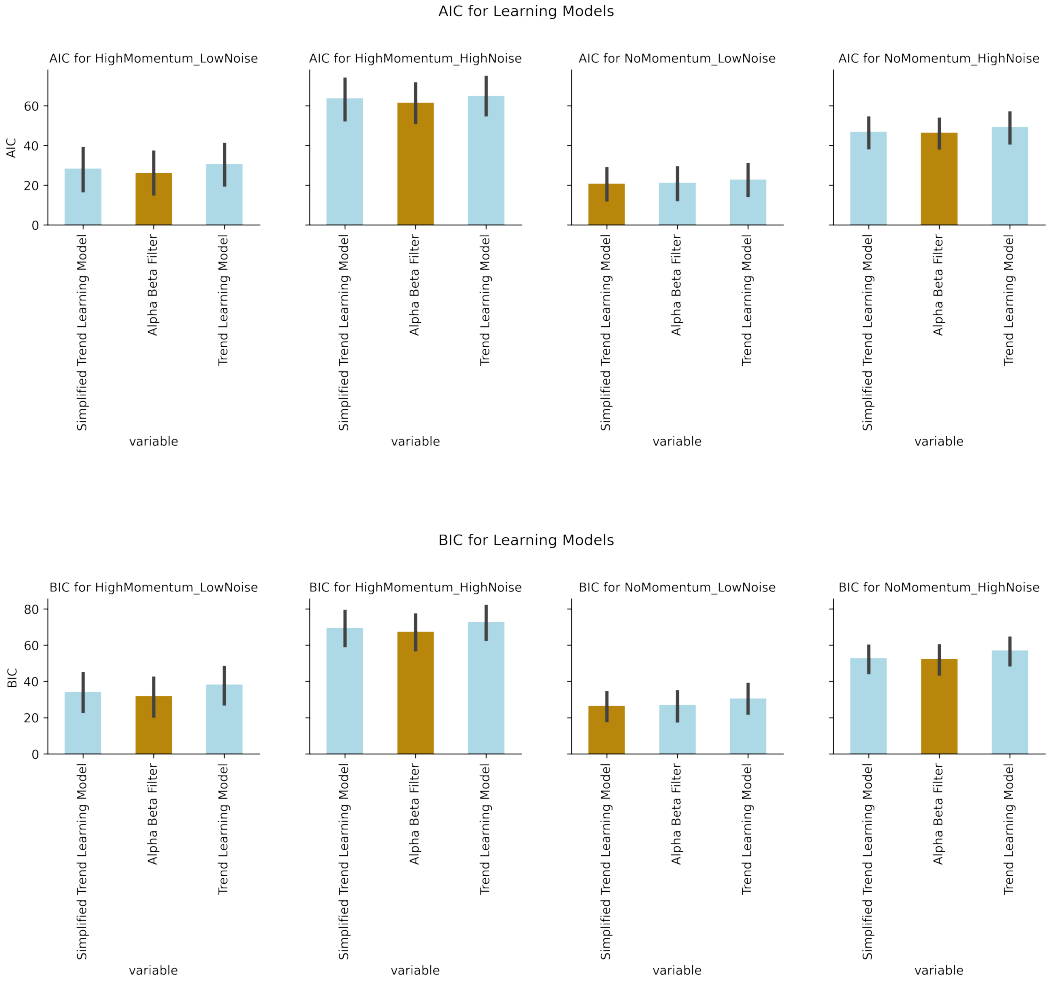


Figure 4.12: BIC and AIC of each model
While both BIC and AIC yielded consistent results across models, neither measure demonstrated a consistent preference throughout all experimental blocks.

4.2.6 Pupil Dilation Preprocessing

Pupil diameter data were recorded from both eyes using the Eyelink 1000 system (SR Research) at a sampling rate of 500 Hz. To clean the data, blinks were automatically removed using the Eyelink system's built-in filter, along with any saccades that occurred immediately before a blink. Missing data points were addressed through linear interpolation, with trials containing more than 50% interpolated data excluded from further analysis. A low-pass Butterworth filter with a 3.75 Hz cut-off was

applied to the cleaned data. The filtered pupil diameters were z-transformed for normalization across the entire session and then averaged between the two eyes. For each trial, a 2-second pre-outcome window and a 5-second post-outcome window were extracted for analysis (Browning et al., 2015). To mitigate the effects of initial adaptation to luminance changes, the first five trials of each session were excluded (Browning et al., 2015; Nassar et al., 2012).

We focused on both tonic baseline pupil dilation and phasic pupil dilation. Baseline correction was done by subtracting the mean pupil size during the baseline period from each time point following the outcome presentation.

4.2.7 Pupil Dilation Analyses

Regression analyses were done separately for tonic and phasic pupil dilation. Specifically, pupil dilation was regressed on model estimated trial-wise momentum, trial-wise noise, interaction between the two, while controlling for the effect of fatigue. Trial-wise momentum was derived from the simplified trend learning model with fitted trend learning rate (LR_Δ), which gives us individual-level estimates of momentum on each trial. Specifically, we computed $Belief_{\Delta t}$ for each subject and block by applying the trend learning update rule to the observed trial-by-trial changes in outcome, using the fitted LR_Δ as a fixed parameter. Trial-wise noise was represented by block labels (0 = low noise, 1 = high noise). Fatigue was operationalized as the trial index within a block (0–49). There were 7s in total with 2s in the pre-outcome period and 5s in the post-outcome period. Since the eyetracker was sampling at 500 hz, there was one data point per 2 ms. We thus conducted the regression analyses on all 7s time intervals.

In this thesis, we examined both phasic and tonic pupil responses, distinguished by whether or not baseline correction was applied. Phasic dilation refers to transient, outcome-related changes in pupil size measured relative to a pre-outcome baseline, consistent with prior work linking baseline-corrected pupil changes to event-evoked arousal (e.g., Nassar et al. (2012) and Preuschoff et al. (2011)). In contrast, tonic dilation refers to overall baseline-uncorrected pupil size, reflecting sustained activity

across a trial and often interpreted as an index of ongoing arousal or control state (e.g.,Aston-Jones and Cohen (2005) and Gilzenrat et al. (2010)).

For phasic pupil dilation, the resultant time series of beta-weights for estimated momentum, noise, and interaction were down-sampled to generate beta estimates reflecting the effects of momentum and noise on pupil dilation. These estimates were calculated for 10 sequential 500-ms time bins spanning the outcome period from its onset. The beta values were then analysed using separate analyses of variance (ANOVAs) for momentum, noise, and interaction. Time bin was treated as a within-subject factor, while block order (first two blocks as momentum, first two blocks as no momentum, and mixed) was included as a between-subject factor. Block order was included as a variable of interest because participants who first completed two consecutive momentum blocks could potentially overgeneralize the trends learned to subsequent no-momentum blocks. Additionally, post-hoc one-sample t-tests were conducted on individual time bins where a significant time-bin effect was found, comparing them to zero and applying Bonferroni correction to account for multiple comparisons. To examine the relationship between pupil dilation to momentum and psychiatric traits, we also averaged the beta-weights for momentum, noise, and their interaction across the entire outcome period (i.e., across the 10 time bins) for each participant. These averaged beta weights were then correlated with self-reported psychiatry trait scores using Spearman's correlation.

For tonic pupil dilation, we calculated cross-time bin summary measures of pupil-beta weights for momentum, noise, the interaction between momentum and noise, and fatigue across the entire 7-second period. We conducted a univariate ANOVA with mean pupil-beta weights of intercept, momentum and noise as dependent variable and block order as the fixed factor. These mean pupil-beta weights were also used in correlational analyses to examine their relationships with behavioural adaptations and between-subject measures of psychiatric traits.

4.3 Results

4.3.1 Increased learning rate when momentum is present under low noise

Model-free Analysis

People significantly increased their learning rates under low noise conditions for prediction errors going in the same direction as momentum compared to learning rates for prediction errors going in the opposite direction of momentum ($t(43) = 9.99, p < 0.001$). However, the same momentum effect was not present in the high momentum high noise block ($p = 0.92$, Figure 4.13a).

We also investigated effect of block order on learning rate adaptations. Specifically, participants who experienced high momentum blocks first might have learnt that "momentum exists", leading them to overgeneralize this pattern to subsequent no-momentum blocks. Conversely, starting with no momentum blocks might have created a "momentum does not exist" belief, influencing how they processed subsequent high momentum blocks. This carryover effect suggests that participants may have been primed to expect similar patterns across blocks. To explore this further, participants were grouped by block order. Through randomization, 10 participants played two high momentum blocks first and 10 played two no momentum blocks first. Though we saw the momentum effect across all block orders, the momentum effect was most pronounced in the two high momentum block first group. And such momentum effect carried over to no momentum blocks (Figure 4.13b).

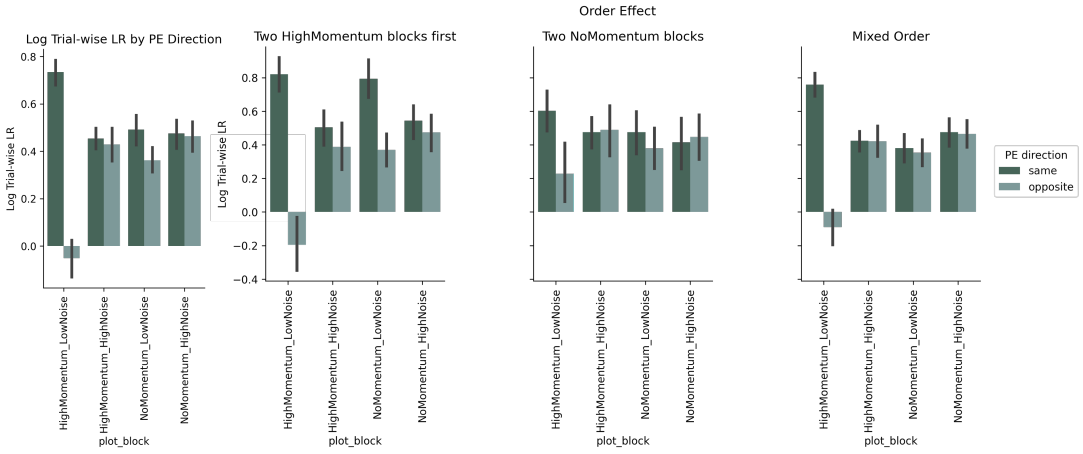


Figure 4.13: Order effect.
a) The figure shows that under low noise conditions, participants significantly increased learning rates for prediction errors aligned with momentum compared to those opposing momentum ($t(43) = 9.99$, $p < 0.001$), but this effect was absent in high noise, high momentum blocks ($p = 0.92$). b) Block order influenced learning rate adaptations: participants starting with high momentum blocks showed a stronger momentum effect, which carried over to subsequent no-momentum blocks, while those starting with no-momentum blocks exhibited a weaker momentum effect in later high momentum blocks.

Simplified Trend Learning Model

We computed trial-by-trial estimates of momentum using the fitted trend learning rate (LR_Δ) and then averaged these values across trials within each block. Estimated momentum was higher in the high momentum blocks compared to no momentum blocks ($F(1, 43) = 2605.837$, $p < 0.001$). There was also a main effect of noise on estimated momentum ($F(1, 43) = 311.862$, $p < 0.001$). There was a significant interaction between momentum and noise ($F(1, 43) = 794.504$, $p < 0.001$), suggesting that the influence of momentum cues on trial-wise estimates depended on noise level, and vice versa (Figure 4.14).

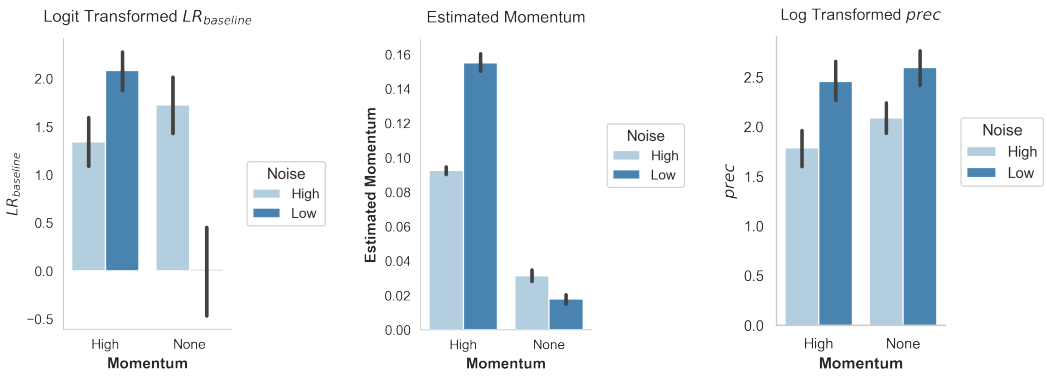
To unpack this interaction, we conducted a series of paired-sample t-tests. Within the high momentum condition, estimated momentum was significantly lower in the high noise blocks compared to the low noise blocks ($t(43) = -22.878$, $p < 0.001$). Within the no momentum condition, estimated momentum was significantly higher in the high noise blocks than in the low noise blocks ($t(43) = 86.566$, $p < 0.001$). At high noise levels, estimated momentum was higher in high momentum blocks

compared to no momentum blocks ($t(43) = 23.776$, $p < 0.001$), and this difference was even greater under low noise ($t(43) = 64.285$, $p < 0.001$).

To test whether the momentum effect varied depending on the level of noise, we calculated how much learning rates increased in response to momentum under low-noise and high-noise conditions. Specifically, we subtracted learning rates in no momentum blocks from those in high momentum blocks, separately for low noise ($\text{diff}_{\text{low noise}} = \text{HighMomentum}_{\text{LowNoise}} - \text{NoMomentum}_{\text{LowNoise}}$) and high noise ($\text{diff}_{\text{high noise}} = \text{HighMomentum}_{\text{HighNoise}} - \text{NoMomentum}_{\text{HighNoise}}$) conditions.

A paired-sample t-test revealed that the momentum effect was significantly larger in the low noise condition than in the high noise condition ($t(43) = 28.187$, $p < 0.001$). This suggests that participants were better able to detect and use momentum cues when the environment was less noisy.

We also tested whether the impact of noise differed between momentum conditions by computing $\text{diff}_{\text{high momentum}} = \text{HighMomentum}_{\text{LowNoise}} - \text{HighMomentum}_{\text{HighNoise}}$ for noise adaptation under high momentum and $\text{diff}_{\text{low momentum}} = \text{NoMomentum}_{\text{LowNoise}} - \text{NoMomentum}_{\text{HighNoise}}$ for noise adaptation under low momentum. These difference scores were significantly different ($t(43) = 28.187$, $p < 0.001$), indicating that participants' sensitivity to noise was greater when momentum cues were present. In other words, the presence of momentum signals amplified participants' response to noise.

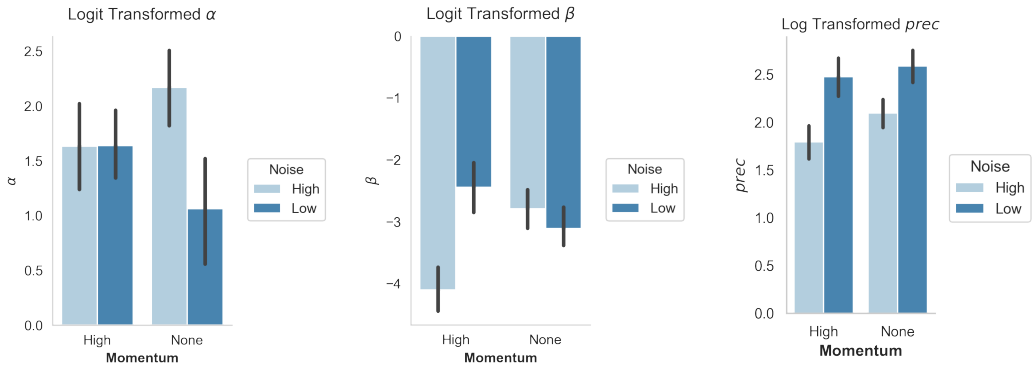


(a) **(b)** **(c)**

Figure 4.14: Estimates of Participants' Learning Rates from simplified trend learning model.
We fitted the simplified trend learning model to participants' choices for each block, applying logit and log transformations to learning rates and decision precision, respectively. The black vertical lines indicate ±SEM for participant learning rates and decision precision. (a) Participants exhibited higher baseline learning rates in momentum blocks compared to no-momentum blocks ($F(1, 43) = 25.130$, $p < 0.001$) and lower learning rates in low-noise blocks compared to high-noise blocks ($t(43) = 8.123$, $p < 0.001$). A post-hoc t-test showed that the momentum effect was primarily driven by an increase in learning rates during high momentum, low noise blocks compared to no-momentum, low-noise blocks. (b) There was a significant effect of momentum on estimated momentum ($F(1, 43) = 2605.837$, $p < 0.001$) and a significant noise effect was observed ($F(1, 43) = 311.862$, $p < 0.001$). (c) Both momentum and noise effects were evident for the precision parameter.

Alpha-Beta Filter

The alpha (α) parameter from the alpha-beta filter showed no effect of momentum ($F(1, 43) = 0.014, p = 0.908$) and a significant effect of noise ($F(1, 43) = 19.251, p < 0.001$, with participants increasing α in high noise conditions Figure 4.15a). The beta (β) parameter showed a near-significant effect of momentum ($F(1, 43) = 3.802, p = 0.058$; Figure 4.15b) and a significant effect of noise ($F(1, 43) = 10.051, p = 0.003$; Figure 4.15b). A post-hoc analysis on the near-significant momentum effect on β showed that this effect was driven by a higher learning rate in the high momentum, low noise block compared to the no momentum, low noise block ($t(43) = 3.044, p = 0.004$; Figure 4.15c).



(a) **(b)** **(c)**

Figure 4.15: Estimates of Participants' Learning Rates from alpha-beta filter. We fitted the alpha-beta filter to participants' choices for each block, applying logit and log transformations to learning rates and decision precision, respectively. The black vertical lines indicate ±SEM for participant learning rates and decision precision. (a) We did not observe any significant effect of momentum on α ($p = 0.908$) but there is a noise effect ($F(1, 43) = 19.251$, $p < 0.001$). (b) There is a near significant momentum effect on the beta (β) parameter, where participants showed higher learning rates in momentum blocks compared to no momentum blocks ($F(1, 43) = 3.802$, $p = 0.058$). A post-hoc t-test showed that the momentum effect was primarily driven by an increase in learning rates during high momentum, low noise blocks compared to no momentum, low noise blocks, $t(43) = 3.044$, $p = 0.004$. There is also a significant effect of noise on β, $F(1, 43) = 10.051$, $p = 0.003$. (c) Both momentum and noise effects were evident for the precision parameter.

Across all analytical approaches, including model free trial wise learning rates, the simplified trend learning model, and alpha beta filters, we consistently found strong evidence for adaptation to momentum, as expected. However, this effect was primarily driven by learning rates under low noise conditions. Most importantly, in the simplified trend learning model, model estimated momentum was higher in high momentum blocks compared to blocks without momentum.

4.3.2 Trait Anxiety and depressive symptoms do not modulate behavioral adaptation to momentum

To assess adaptation to momentum using model-free learning rates, we calculated the difference between learning rates for prediction errors in the same and opposite directions within each block and compared these differences between the High Momentum Low Noise and No Momentum Low Noise conditions $\text{Adaptation}_{\text{low noise}}$. A greater difference indicates higher sensitivity and behavioural adaptation to

momentum. $\text{Adaptation}_{\text{low noise}}$ was not related with CES-D ($r_s(42) = 0.037, p = 0.811$) or with trait anxiety ($r_s(42) = -0.0319, p = 0.836$). Similarly, when examining adaptation to momentum under high noise conditions $\text{Adaptation}_{\text{high noise}}$, no significant correlations were found with CES-D ($r_s(42) = -0.130, p = 0.400$) or trait anxiety ($r_s(42) = -0.232, p = 0.129$).

Using model based approaches, we assessed adaptation to momentum by examining estimated momentum in the simplified trend learning model and the beta parameter in the alpha beta filter. As in the model free analysis, greater differences reflected stronger behavioural adjustment in response to momentum. Contrary to our hypothesis, neither trait anxiety nor depressive symptoms significantly influenced adaptation to volatility or noise. Specifically, the difference in estimated momentum between high and low momentum blocks under high noise was not significantly correlated with depression ($r_s(42) = -0.15,\ p = 0.34$) or trait anxiety ($r_s(42) = -0.13,\ p = 0.39$). Similarly, under low noise, the difference was not associated with depression ($r_s(42) = -0.03,\ p = 0.85$) or trait anxiety ($r_s(42) = -0.23,\ p = 0.13$). In the alpha beta filter, beta adaptation under both low and high noise conditions also showed no significant association with trait anxiety ($p > 0.12$).

4.3.3 Post-outcome phasic level pupil dilation tracks momentum

We examined trial-wise pupil responses to outcome using baseline-corrected data. The pre-outcome period was sampled into 2000 bins of 1 ms each, and the post-outcome period into 5000 bins of 1 ms each, resulting in a total of 7000 time bins. For each bin, we ran a linear regression to estimate how trial-wise model-derived momentum, noise, their interaction, and fatigue predicted pupil dilation. This yielded four separate time series of beta weights, reflecting the effect of each predictor over time.

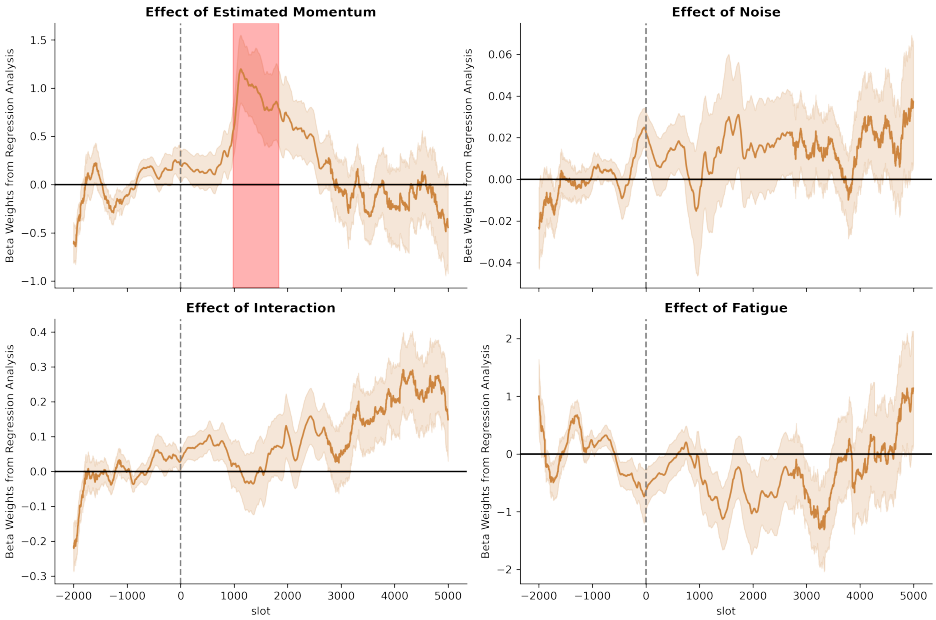
We downsampled the post-outcome beta time series into ten 500-ms bins and conducted repeated-measures ANOVAs. For momentum, there was a significant

effect of time bin on beta weights ($F(9, 33) = 3.361, p < 0.001$), but no effect of block order ($p = 0.369$).

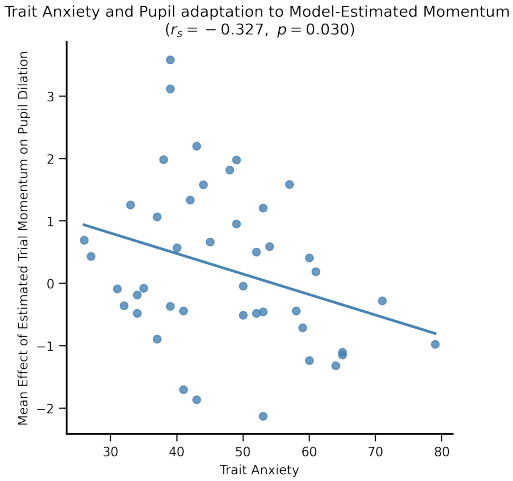
To more precisely identify when momentum significantly predicted pupil responses while controlling for multiple comparisons across time, we applied a cluster-based permutation test to the full-resolution momentum beta time series (Maris & Oostenveld, 2007). At each time slot, we performed a one-sample t-test across participants to test whether the mean beta coefficient differed from zero. Time points exceeding a critical threshold (uncorrected two-tailed $p < 0.05$) were grouped into clusters of adjacent significant time points. Cluster-level statistics were defined as the sum of the absolute t-values within each cluster. To determine significance, we constructed a null distribution of maximum cluster statistics from 1000 permutations, randomly flipping the sign of each participant's beta time series to simulate the null hypothesis. This analysis revealed a significant positive cluster from 0.98 s to 1.84 s post-outcome ($p = 0.040$, corrected; corresponding to slots 980–1838), indicating a sustained period during which estimated momentum consistently predicted increased pupil dilation. No significant clusters were found for the other predictors, including noise, fatigue, or the momentum × fatigue interaction (Figure 4.16).

To examine individual differences in pupil sensitivity to model estimates, we calculated average beta weights across the entire 5-second post-outcome period for both momentum and noise regressors. We then tested whether trial-wise learning rate adaptation to momentum—defined as the sum of $\text{Adaptation}_{\text{high noise}} + \text{Adaptation}_{\text{low noise}}$—was associated with average momentum-related pupil responses. No significant correlation was observed ($r_s(42) = -0.0717, p = 0.643$). Similarly, adaptation to noise, quantified as the difference in baseline learning rate between low- and high-noise blocks, was not significantly associated with average pupil responses to noise ($r_s(42) = -0.195, p = 0.204$). Beta learning rates estimated from the alpha-beta filter also showed no association with either average momentum or noise beta weights ($p > 0.223$).

Finally, we found that participants' average pupil response to momentum was significantly associated with trait anxiety, such that individuals with higher anxiety showed reduced pupil responses to momentum ($r_s(42) = -0.319, p = 0.035$).



(a)



(b)

Figure 4.16: Post-outcome phasic level pupil dilation tracks momentum (a) Time courses for the effect of trial-wise estimates of momentum and noise on pupil dilation before and following presentation of outcome. The graphs show the mean of across participants (n = 44) of the beta weights obtained by regressing pupil dilation against trial-wise momentum and noise. Post outcome pupil dilation was greater for trials where momentum is present, ($F(3.139, 41) = 3.475, p = 0.017$) but not for high noise versus low noise ($F(2.707, 41) = 0.658, p = 0.564$). Shaded area in pink marks the time interval during which the effect of momentum on pupil dilation post-outcome was significantly different from zero, as determined by a two-sided cluster-based permutation test (corrected for multiple comparisons across time). Error bars represent the standard deviation of the regression coefficients (beta weights) across participants. (b) Mean pupil response to momentum post outcome was correlated negatively with trait anxiety $r(43) = -0.319, p = 0.035$.

4.3.4 Tonic level pupil size tracks noise

We investigated trial-wise pupil responses to outcome using baseline uncorrected data. Using the same analyses conducted on baseline-corrected data, we performed regression analyses on each of the 7000 1-ms bins spanning the pre- and post-outcome periods. The independent variables in these regressions included model estimated trial-wise momentum, noise, their interaction, and fatigue. These analyses produced four time series of beta weights corresponding to model estimated momentum, noise, interaction, and fatigue, respectively.

To quantify the overall effects of these variables, we calculated the mean beta weights for all variables (model estimated momentum, noise, interaction, fatigue, and intercept) by averaging across the entire 7-second period. The mean intercept represents the baseline level of pupil diameter across all time points, capturing the overall tonic pupil size during the pre- and post-outcome intervals. The single averaged beta weights for momentum, noise, interaction, and fatigue quantify their respective effect on pupil diameter.

A one-way ANOVA was conducted to examine the effect of block order on mean beta weights of noise. The analysis revealed no significant effect of block order on beta of noise, $F(2, 44) = 2.866, p = 0.068$. However, the mean beta of noise was significantly different from zero, $F(1, 44) = 6.875, p = 0.012$, Figure 4.17, indicating that participants, on average, demonstrated a non-zero beta of noise across all block orders.

We repeated the same analysis on mean beta weights for momentum. There was no significant effect of block order on mean beta weight of momentum, $F(2, 44) = 1.737, p = 0.189$. However, the mean beta of momentum was not significantly different from zero, $F(1, 44) = 1.336, p = 0.255$.

To determine the extent of which behavioural adaptation to momentum and noise correlate with pupillary response to momentum and noise, we ran correlation analyses between the mean beta weight for momentum and noise and relative learning rates. We again checked if adaptation of momentum quantified by differential learning rates ($\text{Adaptation}_{\text{high noise}} + \text{Adaptation}_{\text{low noise}}$) is related to beta weights of

momentum. There was no correlation between behavioural adaptation and pupillary adaptation to momentum ($r_s(42) = -0.079, p = 0.609$). Similarly, relative trial-wise learning rate (learning rates in low noise blocks – learning rates in high noise blocks) was not related to beta weights of noise ($r_s(42) = 0.0786, p = 0.611$).

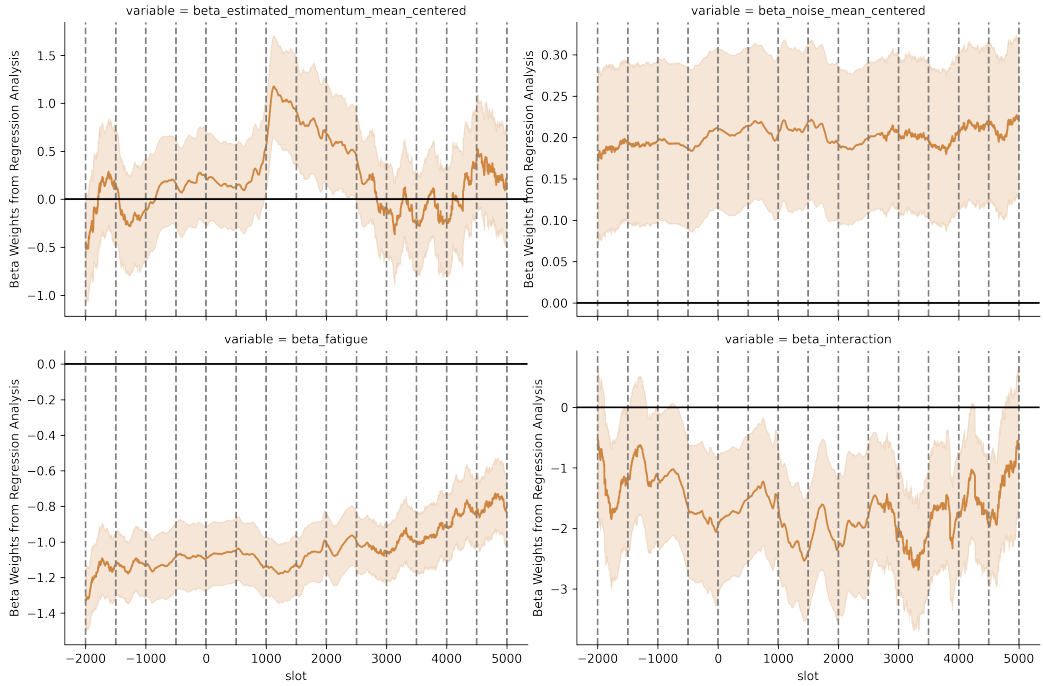


Figure 4.17: Tonic level pupil size tracks noise.
Time courses for the effect of trial-wise estimates of momentum and noise on baseline uncorrected pupil size before and following presentation of outcome. Post-outcome pupil dilation was greater for trials when noise is high versus low, ($F(1, 44) = 7.648, p = 0.008$) but not for the presence of momentum ($F(1, 44) = 1.336, p = 0.255$)). Error bars represent the standard deviations of the regression coefficients (beta weights) from the pupil analysis.

4.4 Discussion

The aim of this study was to examine how individuals adapt their learning strategies in environments where both momentum and noise are present. Using model-free analysis, we found that learning rates for prediction errors in the same direction as the momentum were significantly higher than those for prediction errors in the opposite direction to the momentum. But this asymmetry was observed only under low noise conditions. The simplified trend learning model provided consistent evidence,

where estimated momentum was higher in high momentum blocks. Additionally, the alpha-beta filter analyses showed that momentum has a near-significant effect on model-beta (β), which affects the velocity estimate, capturing how quickly the model adapts to changes in the trend of observed data. Post-hoc analysis showed that model-betas (β) in high momentum low noise block were significantly greater than model-betas (β) in in no momentum low noise block. We also replicated our previous findings on the effect of noise on learning rates. Both baseline learning rates derived from simplified trend learning model and model-betas (β) in parameter from the alpha-beta filter were significantly lower in high noise blocks compared to low noise blocks. Contrary to our hypothesis, we found no correlations between depressive symptoms (measured by CES-D) or trait anxiety and the adaptation of learning rates to either momentum or noise.

Additionally, we observed that momentum significantly increased pupil dilation, consistent with an effect on phasic activity following trial outcomes. In contrast, noise influenced tonic activity, with larger overall pupil size in high-noise conditions compared to low-noise conditions. This finding ran counter to our original hypothesis that noise would reduce tonic pupil size.

Yu and Dayan (2005) proposed that norepinephrine signals unexpected uncertainty, while acetylcholine signals expected uncertainty, offering a framework in which the brain differentiates meaningful change from random fluctuations. Our results invite a complementary interpretation. Specifically, we observed that structured volatility in the form of momentum was associated with phasic pupil responses, whereas unstructured volatility (noise) was associated with tonic pupil changes. This pattern suggests that phasic arousal mechanisms may serve to track directional, meaningful change in the environment, while tonic arousal mechanisms may instead reflect the broader level of environmental randomness or unpredictability.

We must emphasise the limitations of this interpretation. Although there is strong evidence that pupil size correlates with LC-NA activity (Joshi et al., 2016), pupil dynamics are also influenced by other neuromodulatory systems. For instance, serotonin has been shown to modulate pupil size differently depending on levels of

uncertainty (Cazettes et al., 2021). Thus, while our findings are consistent with the idea that distinct arousal processes contribute to differentiating structured volatility from noise, they do not demonstrate a direct causal mapping. Rather, the balance between tonic and phasic arousal mechanisms, as indexed by pupil activity, might contribute to the brain's ability to discriminate between meaningful trends and random fluctuations.

A possible alternative interpretation is that phasic activity may be a response to the most salient environmental changes, regardless of uncertainty type (O'Reilly et al., 2013). In environments with momentum, the pattern of volatility becomes salient, prompting norepinephrine neurons to signal this momentum. In noisy settings, excessive randomness might mask this signal, making phasic responses harder to detect. Instead, changes in noise around a central point become salient, as each trial may involve mean shifts.

We also found that trait anxiety was associated with reduced pupil dilation in response to momentum. This finding aligns with previous work showing that individuals with higher anxiety exhibit less pupil responses to volatility (Browning et al., 2015), suggesting a broader pattern of reduced sensitivity to environmental changes in anxious individuals. Together, these findings indicate that trait anxiety is linked to a deficit in the use of higher-order statistics about the structure of the environment more generally. In particular, the results presented in this chapter extend previous work by demonstrating that reduced pupil dilation in anxious individuals is not limited to volatility, but also applies to more structured, directional forms of change such as momentum. This diminished pupillary response could reflect a blunted NE response, potentially indicating reduced detection of environmental uncertainties.

There were no significant positive correlations between pupillary responses, at either the phasic or tonic levels, and behavioural adaptation to momentum and noise. This dissociation suggests that behavioural and physiological adaptations maybe driven by distinct underlying processes. Another possibility is that behaviour and pupil responses operate on different timescales. Whereas behavioural learning

rates reflect a gradual accumulation of evidence across trials, pupil responses tend to be more transient. As such, the average beta weights of momentum might not align cleanly with longer-term behavioural learning dynamics.

An additional explanation may lie in the model design itself. Both the simplified trend-learning model and the alpha-beta filter demonstrate the asymmetric learning rate effects observed in the model-free analysis, supporting positive model selection. However, the effect of momentum seem to be captured by both $LR_{baseline}$ and LR_{Δ} from the simplified trend learning model and α and β parameters in the alpha-beta filter. This non-specificity may explain the absence of correlations between behavioural and pupillary adaptation to momentum.

4.5 Conclusion

In this chapter, we explored how individuals adapt to momentum, a form of linearly structured volatility and noise within an environment, and the role that learning rate modulation plays in this process. Our findings indicate that participants adapt their learning rates when momentum is present, but only under low-noise conditions. This adaptation highlights the human capacity to integrate patterns and differentiate between significant and random environmental changes.

Phasic and tonic pupil dilation might provide insights into how physiological responses correspond to expected and unexpected uncertainties. Specifically, we found that phasic pupil responses tracked momentum, while tonic pupil responses were more associated with noise. The results align with the dual-mode hypothesis of the locus coeruleus-norepinephrine system, where phasic activity enhances attention to salient changes and tonic activity signals ongoing background states.

5

Conclusions and General Discussion

In Chapter 1, we raised several outstanding questions within the field of learning under uncertainty and its association with symptoms of mood disorders. To address these questions, we designed experiments that manipulated both volatility and noise. In this section, we discuss the answers to these outstanding questions.

5.1 Answers to outstanding questions in Chapter 1

5.1.1 How well could human learners differentiate between noise and volatility?

Our findings suggest that people adapt to volatility and noise independently. This is evidenced in Chapter 3, where we observed significant main effects of both volatility and noise on learning rates, but no interactions between them. The lack of an interaction effect indicates that participants adjusted their learning rates in response to changes in volatility regardless of the noise level, and vice versa. This pattern suggests that learners were not confusing the two sources of uncertainty, as their learning behavior indicates they were responding differently to noise and volatility, even if distinction was not necessarily consciously accessible.

However, results from Chapter 2 complicate this interpretation. The effect of noise on learning disappears when the overall noise level was high, suggesting that participants struggled to adapt when the environment becomes highly random. One possibility is that participants mistook noise for volatility, which implies that the ability to differentiate between noise and volatility breaks down under high randomness. Alternatively, participants could have simply stopped adapting to noise normatively when it reached a certain threshold. Some aspects of this observed pattern may also reflect artefacts of task design: in addition to changes in noise level, we also manipulated other parameters such as bucket size, which could have influenced how uncertainty was perceived and processed in some versions of the task.

Taken together, these interpretations suggest that while learners generally treat noise and volatility as distinct sources of uncertainty, their ability to do so may be constrained under conditions of extreme noise.

5.1.2 What are the neural processes underlying the attribution and signalling of volatility and noise?

In Chapter 3, we assessed the effect of noise and volatility on pupil activities. While previous studies have typically operationalised volatility using change-point paradigms that introduce abrupt shifts in the environment (e.g., Browning et al. (2015) and Nassar et al. (2012)), we instead implemented a gradual, unstructured form of volatility using a Gaussian random walk. Our findings indicated that in individuals with low anxiety, pupil area (tonic activity) increased in high-volatility environments. Noise also appeared to increase tonic pupil area, although this effect did not reach statistical significance.

In Chapter 4, we investigated a different form of volatility, i.e., momentum, which involved a linearly structured change in the environment. In contrast to unstructured volatility, momentum led to a significant positive phasic effect on pupil dilation. Noise again produced a significant increase in tonic pupil area. These contrasting effects of momentum and unstructured volatility suggest that different types of volatility might engage distinct arousal systems , with structured

volatility engaging rapid, transient responses, and unstructured volatility prompting more sustained, background arousal.

Taken together, these results offer new insights into how the brain signals and responds to uncertainty. Specifically, we did not observe any significant phasic pupil response to gradual, drifting volatility. Instead, increased volatility led to a sustained tonic pupil response in individuals with low anxiety. This suggest that when environmental volatility lacks a clear structure or identifiable change-point, the LC-NE system may promote sustained arousal under unexpected uncertainty rather than rapidly adapting to discrete events.

These findings build on prior studies that have linked phasic pupil dilation to volatility-driven learning, where volatility was produced by a sudden change point (Browning et al., 2015; Nassar et al., 2010, 2012). In contrast, the gradual, unstructured volatility in our task did not elicit phasic responses, suggesting that not all forms of volatility are treated as discrete surprise signals. Instead, as shown in Chapter 3, unstructured volatility was likely not perceived as an isolated learning cue (e.g., a reversal or change-point) but rather as an ongoing condition requiring heightened vigilance, akin to noise, which also increased tonic pupil activity without eliciting phasic responses. However, when volatility was structured as momentum, we observed a significant phasic pupil response, suggesting that momentum may function as a salient cue that correlates with arousal-linked mechanisms thought to support flexible learning. This interpretation is consistent with human evidence linking pupil dilations to adaptive learning adjustments (Jepma & Nieuwenhuis, 2011; Nassar et al., 2012) and animal work showing correlations between phasic locus coeruleus activity, pupil size, and behavioural updating (Aston-Jones & Cohen, 2005; Joshi et al., 2016), though in both cases the relationship remains correlational rather than causal.

Altogether, these findings point to an important distinction: volatility is not a unitary construct. Rather, the brain appears to differentiate between types of volatility, with structured and unstructured forms eliciting distinct patterns of arousal. Unpredictable, gradual volatility was associated with tonic pupil responses,

which have been linked to shifts in engagement and disengagement with tasks (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011). In the present context, we interpret this tonic engagement as consistent with sustained vigilance in the face of uncertainty. By contrast, directional or structured volatility such as momentum was associated with a primarily phasic pupil response, which suggests heightened sensitivity to directional trends in the environment.

Finally, while pupillometry provides a useful readout of arousal, the underlying neuromodulatory processes are complex. Many neurotransmitters, including NA, ACh, and serotonin, influence pupil dilation, making it difficult to confidently infer activities of neurotransmitters through pupillometry alone. Nonetheless, the distinct tonic and phasic patterns observed across conditions suggest more nuanced neuromodulatory contributions than those captured by existing frameworks such as Yu and Dayan (2005) or Nassar (2024).

5.1.3 How do human learners adapt under linearly structured volatility?

In Chapter 4, we investigated how human learners adapt to momentum, a linearly structured form of volatility. To capture this adaptation, we examined trial-wise learning rates differentiated by the direction of prediction errors. We found that learning rates were significantly higher for prediction errors in the same direction as the momentum compared to those in the opposite direction – but only under low noise conditions. This suggests that human learners are sensitive to linear trends in environmental changes and can integrate this structure to optimize their predictions when the signal is clear. To further explore this process, we modified the Rescorla-Wagner learning model to include a learning process that estimated shifts in outcome location over recent trials. This allowed us to infer a trial-wise estimate of perceived momentum. We observed that participants' estimated momentum was significant higher in high momentum blocks than in no momentum blocks.

Complementing these behavioral findings, our pupillometry data (discussed earlier) showed a phasic pupil response to estimated momentum, suggesting that

participants not only respond to linearly structured volatility behaviorally but also exhibit a corresponding physiological response. Together, these results indicate that human learners are sensitive to momentum and can integrate higher-order uncertainty into their learning processes.

5.1.4 Are adaptations of learning to noise, volatility, and linearly structured volatility associated with depressive and anxiety symptoms?

Our study investigated whether trait anxiety and depressive symptoms are associated with behavioural and physiological adaptation to volatility and noise. While our primary analyses found no evidence that these symptom dimensions influenced adaptation, an exploratory analysis revealed that individuals with greater capacity for anticipating future pleasure tended to show less adaptation of learning rate to noise under low-volatility conditions. Although this association did not survive correction for multiple comparisons, the pattern raises the possibility that higher anticipatory hedonic ability may be linked to a tendency to interpret random fluctuations as meaningful. This interpretation is consistent with Alloy and Abramson (1979) observation that non-depressed individuals were more likely to overestimate action–outcome contingency, particularly when outcomes were desirable. In the context of our task, such a bias could reflect a cognitive tendency to maintain a sense of control when outcomes are driven by noise rather than true volatility.

In terms of pupillometry responses, we observed in Chapter 3 that individuals with higher trait anxiety showed reduced tonic pupil responses to volatility, and those with lower trait anxiety displayed tonic pupil responses that tracked changes in volatility. Similarly, in Chapter 4, we found reduced phasic pupil dilation to momentum in individuals with higher trait anxiety. These findings are consistent with previous research suggesting that high trait anxiety is associated with reduced pupillary responses to meaningful changes in the environment (Browning et al., 2015). Given the empirical evidence that the LC-NE system plays a central role in signaling unexpected uncertainty, this diminished pupil response in anxious

individuals may reflect disrupted neuromodulatory signalling, potentially impairing their ability to maintain vigilance and flexibly respond to changing conditions.

Across Chapter 3 and 4, however, we did not find strong evidence that trait anxiety or depressive symptoms were related to behavioral adaptation of learning to volatility or noise. This contrasts with previous research (e.g., Browning et al. (2015) and Gagne et al. (2020), which has reported high trait anxiety associated with less adaptation of learning rate to volatility. One plausible explanation for this discrepancy is the difference in task structure: unlike previous work that used change-point paradigm (where sudden shifts or reversal signal volatility), our task used a drift paradigm, in which volatility was manipulated gradually and continuously.

In our task, volatility is manipulated gradually and continuously over time, requiring participants to integrate information across multiple trials in order to infer the underlying structure. The task was deliberately designed so that volatility and noise were difficult to distinguish. This ambiguity was intended to enhance sensitivity to subtle individual differences in adaptive learning. In particular, if noise and volatility could be easily differentiated, most participants would likely adapt successfully, reducing our ability to detect variation between adaptive and maladaptive learning styles. However, one could argue that if maladaptive learning under uncertainty contributes to depression and anxiety, then the most clinically relevant impairments may not emerge in complex or ambiguous scenarios, but rather in situations where the source of uncertainty is clearly signaled. In such cases, a failure to adapt appropriately may reflect a more fundamental impairment in learning.

To explore this possibility more directly, future research could compare responses to volatility driven by change points and by gradual drift within the same participants. Such a design would provide a stronger test of whether the ability to respond to more clearly defined environmental changes is linked to affective symptoms, and whether ambiguity may mask meaningful differences in how individuals learn and regulate arousal under uncertainty.

5.2 Strengths

This thesis presents several key strengths in its experimental design, methodological rigor and analytical approaches, which contributes to a deeper understanding of how humans learn under uncertainty and how these processes relate to mood disorders.

5.2.1 Experimental Manipulation of Volatility and Noise

One of the major strengths of this work is the explicit and systematic manipulation of both volatility and noise. Unlike prior studies that often estimated noise through modelling or set noise levels significantly lower than volatility, we designed experiments that directly manipulated these two types of uncertainties in a controlled manner (Chapter 2 Method). In particular, we tested several versions of volatility and noise levels and ensured that the variation induced by noise in high noise blocks (defined as changes in coin positions from one trial to the next) was of similar magnitude to the variation caused by volatility in high volatility blocks. By matching the magnitude of variability across these conditions, we ensured that participants could not simply rely on the size of the variability to distinguish between noise and volatility (Nassar et al., 2012). This allowed for a clearer interpretation of how participants adjusted their learning strategies, confirming that they were indeed responding to the type of uncertainty rather than to superficial cue like magnitude of variability.

Furthermore, the method used to generate and select experiment schedules is a valuable tool for future studies aiming to create new experimental stimuli. In Chapter 2, we developed a set of selection criteria that allowed us to filter through 37,000 candidate schedules that varied in volatility and noise parameters:

1. Ensuring the trajectory of the mean in high volatility blocks was not unidirectional, to avoid confounding volatility with directional trend.

2. Matching angular differences across high volatility and high noise blocks, so that participants could not rely solely on the magnitude of variability to distinguish between the two sources of uncertainty.

3. Using simulation based on the Rescorla-Wagner model to estimate how different learning rates would affect performance. Schedules were then selected to maximize expected volatility and noise effects while minimizing the dispersion of learning rates, thereby increasing the likelihood that participants would exhibit normative adaptation patterns.

For instance, we applied the method developed in Chapter 2 to create new schedules for Chapter 4 when studying momentum, modifying the first selection criterion to favor schedules with clear directional information. This approach is also useful for experiments that require repetition of the coin-catching task across multiple sessions or for testing boundary condition of which human learners stop adapting normatively to noise. By providing a structured way to manipulate and control uncertainty in behavioral experiments, our methodology offers a foundation for further research into learning and decision-making under uncertain conditions.

5.2.2 Prescreening Participants

Another notable strength is the prescreening of participants based on depressive symptoms. This approach ensured that we obtained a sample that included individuals with a wider range of depressive symptom levels, allowing us to investigate the potential associations between learning adaptations and mood disorder symptoms more effectively. By selecting participants in a way that captures variability in depressive symptoms, we increased the sensitivity of our analyses and reduced the risk of floor or ceiling effects that might obscure meaningful relationships.

5.2.3 Comparison of Different Ways of Computing Learning Rate

This thesis also contributes methodologically by comparing different ways of computing learning rate across both model-free and model-based approaches. Learning rate is a key parameter in the uncertainty literature, reflecting how much learners update their beliefs based on new information. However, its estimation can vary substantially depending on model choice and analytical method.

We first compared two model-free methods for estimating learning rate. The first method calculated learning rate on a trial-by-trial basis, using the ratio of belief updating to prediction error. This approach yields a learning rate estimate for each individual trial. The second method used a regression-based approach, where prediction error was used to predict changes in belief. In this case, learning rate is estimated as the slope of the regression line, computed over a block of trials.

When adopting the trial-wise learning rate approach, prior studies have often excluded extreme learning rates or truncated learning rates to fall between 0 and 1. In contrast, our study included all data points, recognizing learning rates out of the range of 0 and 1 could represent genuine adaptation strategies rather than mere outliers or noise. Notably, approximately one third of the trial-wise learning rates in our data fell outside the 0 to 1 range. This suggests that such values are not rare anomalies but may indicate adaptive strategies, particularly in dynamic or misleading environments. Therefore, our approach provides a more comprehensive pictures of how learners adjust to uncertainty on a trial-by-trial basis.

We also conducted a systematic evaluation of model-based learning approaches. These included the Rescorla-Wagner model, a trend learning model, and a simplified trend model. When linear structure was imposed on volatility, we explored alternative models such as the alpha-beta filter and the proportional-integral controller. Each of these frameworks makes different assumptions about how learners respond to uncertainty, enabling us to test the robustness of our findings across models.

A key strength of our approach was the comparison of the strengths and limitations of each method. We utilised positive model selection alongside model recovery and parameter recovery to determine the most appropriate models for capturing learning adaptation under uncertainties. Because we systematically compared all these approaches, we can say with confidence that the effects we observed were not the result of cherry-picking data or selectively choosing models that fit our hypotheses. Instead, our findings reflect genuine patterns in the data, independent of how we have processed or selected our raw data. Furthermore,

our model-based approach used as few assumptions as possible. This rigorous comparative analysis enhances the interpretability of our results and emphasizes the importance of selecting appropriate analytical approaches (either model-based or model-free) when investigating adaptive learning processes.

In sum, this methodological comparison highlights the importance of choosing appropriate analytical tools, whether model free or model based, when investigating adaptive learning. Our approach reveals how different modeling choices can shape our interpretation of learning behavior and underscores the value of trial wise analysis for capturing nuanced adaptation to uncertainty.

5.2.4 Replicability and Robustness Across Samples

Finally, this thesis demonstrates the replicability of key findings across different participant samples and experimental conditions. The results from Chapter 2, where online participants showed robust adaptation to volatility, were replicated in a controlled in-lab setting in Chapter 3. This strengthens the reliability of our conclusions and suggests that our findings are not limited to a specific sample or context. Moreover, the iterative nature of our experiments, spanning eight different versions, allowed us to refine our paradigm and improve the reliability of our measures over time.

Together, these strengths position this thesis as a rigorous investigation into learning under uncertainty. By combining experimental manipulations, computational modeling, pupillometry, and individual differences in mood symptoms, this work provides a multidimensional perspective on how humans adapt to uncertainty. These contributions lay the groundwork for future studies to further explore the mechanisms of adaptive learning and their implications for mood disorders.

5.3 Limitations

5.3.1 Correlational Studies

It is important to acknowledge that the studies in this thesis are correlational in nature. We did not find strong behavioural evidence of systematic learning rate

adaptation, and our samples did not include patient populations. Therefore, the findings cannot be taken as evidence that maladaptive learning under uncertainty is causally related to depression or anxiety. Instead, our results demonstrate individual differences in learning: for example, lower anhedonia and a greater ability to anticipate future pleasure were associated with reduced adaptation to noise, while higher trait anxiety was associated with reduced adaptation to volatility. These associations do not establish causal links, nor do they in themselves demonstrate maladaptive learning or clinical levels of depression and anxiety. If future research were to establish a causal relationship, however, this could open new avenues for interventions aimed at modifying learning strategies to prevent or alleviate symptoms. Establishing causality would require experimental designs in which either the learning process itself is directly manipulated (for instance, by training participants to adopt different learning rates) or where mood or anxiety states are experimentally altered to test their impact on learning. In Future Work Suggested by the Results in the Thesis, we outline an experiment designed to begin addressing this question.

5.3.2 Pupillometry Analysis and Confounds

Several confounds must be considered in interpreting our pupillometry results. In our regression models, volatility and noise were treated as independent predictors. However, they are often accompanied by other factors such as surprise and trial outcome, which may also influence pupil responses. For example, surprise, typically operationalized as the absolute value of prediction error, has been shown to elicit early phasic increases in pupil dilation (O'Reilly et al., 2013). In both high volatility and high noise environments, outcomes tend to be more unpredictable, leading to larger prediction errors. Thus, surprise is not randomly distributed but systematically elevated under these conditions, making it difficult to disentangle the contribution of surprise from the effects of volatility and noise themselves.

Similarly, trial difficulty can influence tonic pupil diameter and is often reflected in trial outcomes. In our task, trial outcome was binary (correct versus incorrect),

and performance tended to be worse under high volatility and high noise conditions, as expected. This means that trial outcome is not independent of volatility and noise but instead tracks them closely: participants were more likely to make errors when the environment was more uncertain. In this sense, trial outcome serves as a proxy for task difficulty, and it too may account for some of the variance in pupil diameter.

To formally evaluate these relationships, we included both surprise (absolute prediction error) and trial outcome (correct versus incorrect) as additional regressors in our pupil response models. We found that both variables were highly collinear with volatility and noise. That is, trials with higher volatility or noise also tended to have greater surprise and lower accuracy. This high degree of collinearity complicates interpretation: it becomes difficult to determine whether observed pupil responses attributed to volatility and noise are truly driven by those latent factors, or instead reflect more immediate cognitive consequences like prediction error or performance.

This issue highlights a key challenge in studying arousal responses in complex environments. Volatility, noise, surprise, and difficulty are often intertwined. While our regression models attempt to account for multiple predictors simultaneously, their collinearity means that effects should be interpreted with caution. Future work using experimental manipulations that can separate these variables more clearly, such as holding volatility constant while varying trial outcome or surprise independently, may help clarify their distinct contributions to pupil-linked arousal.

5.3.3 Ecological Validity and the Role of Context in Learning under Uncertainty

The experimental design in this thesis focuses exclusively on short-term adaptation to volatility and noise. While we demonstrated how human learners adjust their learning strategies in response to these uncertainties, it remains unclear how stable these adaptations are over longer periods. In everyday settings, individuals may integrate information over extended time scales, often encountering patterns of stability and change that unfold across days, weeks, or even years. For example, in social interactions, Person A may be consistently nice to Person (B) for several

encounters but unexpectedly act mean on a later occasion. When these interactions are spaced out over time, learners may adopt different strategies for attributing changes to noise versus volatility compared to the rapid trial-by-trial feedback seen in our experiments.

Furthermore, real-world environments provide richer contextual cues that may help individuals distinguish whether observed variation is due to volatility or noise. Lastly, we are asking participants to make a prediction trial by trial. In contrast, our experimental task presents a more constrained setting where participants must infer volatility and noise based solely on one feedback, without additional cues that might be available in naturalistic settings (e.g., Person (A)'s train got delayed so the mean interaction Person B had has nothing to do with the actual quality of their relationship going sour).

Another important difference is the way predictions are made in real life versus experimental conditions. In our task, participants make predictions on a trial-by-trial basis, receiving immediate feedback. Although volatility and noise must still be inferred over multiple trials, the timescale for estimating these uncertainties is relatively short. This likely places fewer demands on memory compared to real-world learning. In contrast, real-life learning often involves integrating information over much longer periods, making memory a more significant factor. Given that human memory is imperfect, individuals may give excessive weight to extreme cases or statistically rare events when forming beliefs about uncertainty. This could lead to systematic biases in how people estimate and respond to volatility and noise outside of a controlled experimental context.

5.4 Future Work Suggested by the Results in the Thesis

To better characterize the boundary conditions at which normative noise adaptation disappears, future studies could implement fine-grained manipulations (as shown in Chapter 2) of noise and volatility levels in the same group of participants. This could provide a more comprehensive understanding of how individuals adjust

their learning strategies in response to environmental uncertainty and its potential implications for mental health.

Future research could recruit both healthy individuals and clinically depressed patients. While the current study focused on a non-clinical population, including clinical populations may increase the likelihood of detecting significant effects and help clarify whether maladaptive learning under uncertainty is a contributing factor to mood disorders or a consequence of them.

To investigate the causal role of adaptation in the development of depression and anxiety, future studies could experimentally induce adaptation biases by exposing participants exclusively to high volatility or high noise conditions over an extended period. This approach could help determine whether adaptation styles causally contribute to mood symptoms and, if so, whether modifying these learning strategies could serve as a potential intervention for mood disorders. potentially inform novel interventions for mood disorders.

Further research should also aim to disentangle confounding factors in the pupillometry analysis. Experimental designs that independently manipulate noise, volatility, and related factors such as surprise and task difficulty will be critical for isolating their unique contributions. For instance, in low-volatility, low-noise blocks, the bucket size could be reduced to 15 degrees, whereas in high-volatility, high-noise conditions, it could be extended to 45 degrees to better control task difficulty. Additionally, surprise could be explicitly manipulated by introducing rare but highly informative feedback events to observe their distinct effects on learning and pupil responses (O'Reilly et al., 2013). Finally, combining pupillometry with neuroimaging techniques may offer deeper insights into the neural mechanisms underlying adaptation to uncertainty.

Lastly, future research could investigate whether individuals with anxiety or depression show impairment when volatility and noise are easily distinguished by relaxing some of the selection criteria listed above. If such impairments are found, this would strengthen the case for causal links between maladaptation to uncertainty

and mood disorders. Furthermore, it could also explain why some studies found the correlation with trait anxiety, while some don't.

5.5 Overall Conclusions

This thesis investigated how humans learn under uncertainty, specifically how they adapt to volatility and noise, and how these learning processes relate to symptoms of mood disorders. We observed distinct effects of volatility and noise on learning behavior and physiological responses. In particular, human learners robustly adapt to both unstructured and linearly structured volatility by increasing their learning. In contrast, adaptation to noise was less consistent and appeared to depend on relative levels of volatility and noise. This suggests that while individuals readily adjust their expectations in response to volatility, they struggle to do so under noise, which may lead to maladaptive learning in certain contexts. Furthermore, our results indicated that volatility and noise elicit distinct pupil-linked neuromodulatory responses, with volatility primarily influencing tonic arousal in low-anxiety individuals and momentum—a structured form of volatility—eliciting phasic pupil responses. While we did not find strong evidence linking trait anxiety and depression to learning rate adaptation, we observed near-significant effects suggesting that individuals with higher trait anxiety may exhibit reduced adaptation to volatility, and those with lower anhedonia exhibited less adaptation to noise.

Taken together, the findings from this thesis highlight the complexity of learning in uncertain environments. Future research should go beyond correlation and determine the causal role of learning under uncertainty in the development of mood disorders. By establishing whether maladaptive learning styles contribute to emotional dysregulation, we can assess whether modifying learning strategies could directly influence mental health outcomes. If certain adaptation styles to volatility and noise are found to play a role in the onset or persistence of mood symptoms, this would open the door for targeted interventions aimed at enhancing adaptive learning.

Appendices

A

Supplemental Information for Chapter 2

A.1 Summary plots of learning rate across experiment for each method

In Chapter 2, we compared three approaches to compute learning rate: trial-wise learning rate, slope as learning rate and learning rate derived from Rescorla-Wagner learning model.

We concluded that model-fitted LR provide the most robust analytical approach and observed a consistent volatility effect ((Figure A.3)). There were also consistent effects on median trial-wise learning rate across experiments (see Figure A.1), while the effect of volatility on learning rate estimated with slope from linear regression was less consistent (Figure A.2).

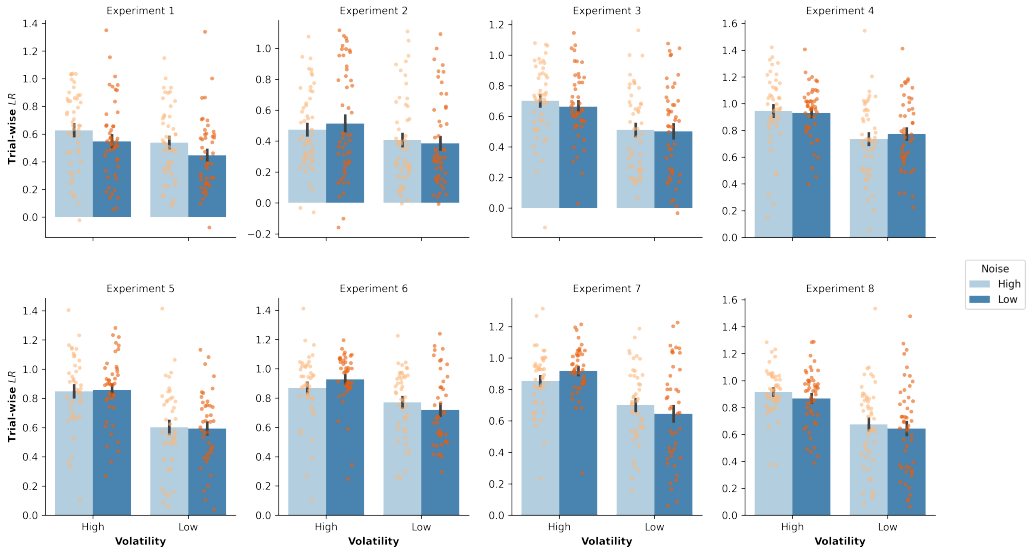


Figure A.1: Adaptation of trial-wise learning rate across eight experiments. Trial-wise learning rates were computed as the change in belief divided by the prediction error. Across eight experiments, we found no consistent adaptation of learning rates to volatility or to noise. Orange points show each participant's median learning rate; error bars represent ±1 SEM (standard error of the mean).

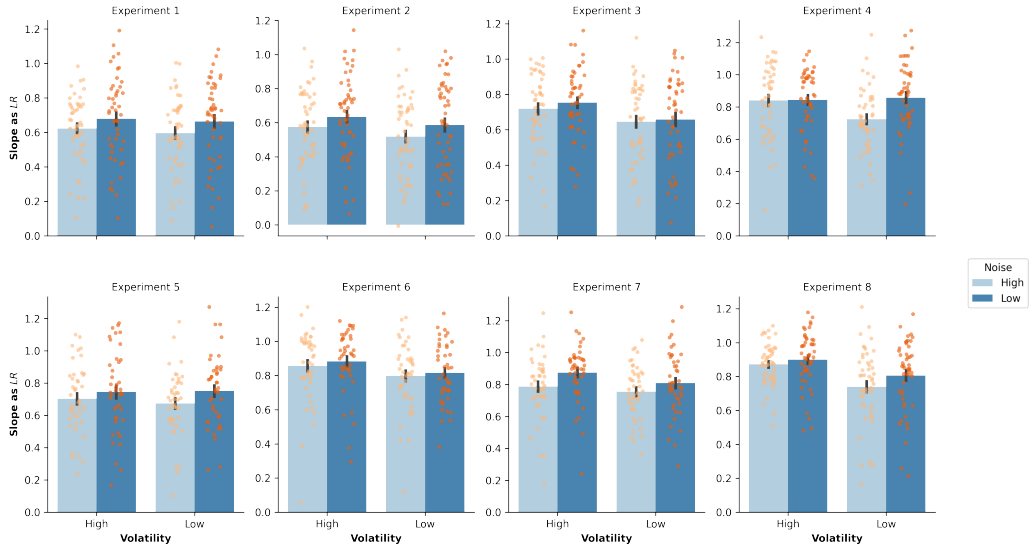


Figure A.2: Adaptation of learning rate as slope across eight experiments. Learning-rate adaptation (slope) across eight experiments. We estimated the learning rate as the slope from a linear regression of change in belief on prediction error. Across experiments, we found no consistent adaptation to volatility or noise. Orange points denote each participant's block-wise learning rate; error bars indicate ±1 SEM (standard error of the mean).

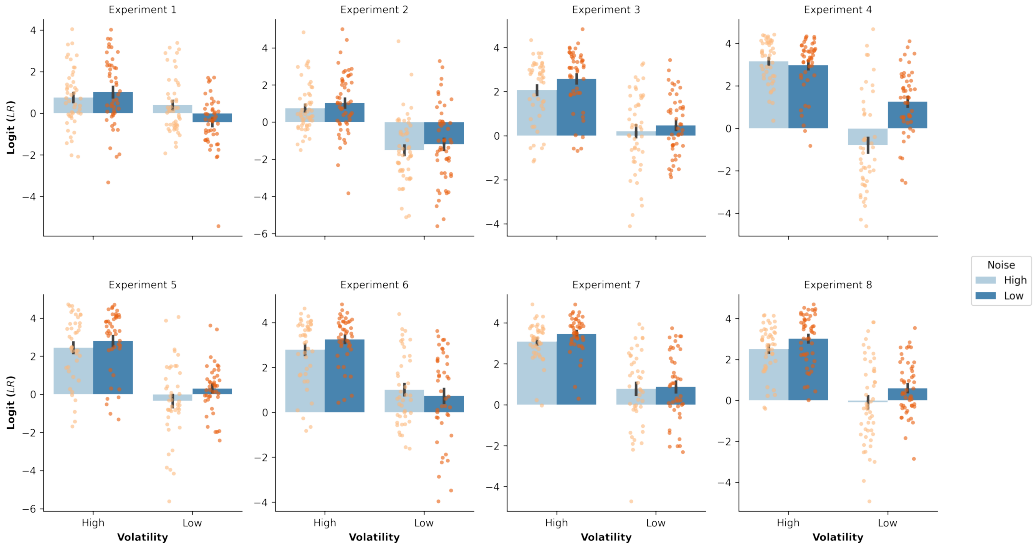


Figure A.3: Adaptation of learning rate across eight experiments derived from Rescorla-Wagner learning model
Learning rates were derived per block by fitting a Rescorla–Wagner learning model to participants' choices. Across eight experiments, learning rates consistently adapted to volatility but not to noise. Orange points show each participant's block-wise learning rate; error bars indicate ±1 SEM (standard error of the mean).

B

Supplemental Information for Chapter 4

B.1 Proportional-integrative-derivative Controller

Ritz et al. (2018) tested a Proportional-Integral-Derivative (PID) controller to describe human trend learning. PID is a widely used control mechanism to manage nuclear reaction. It adjusts the input to a system to maintain the desired setpoint by minimizing the error between the system's current output and the target value. The PID controller operates based on three main components:

Proportional (P) Control

The proportional term produces an output proportional to the current error value. It provides a control action that is directly proportional to the difference between the desired setpoint and the measured output. The proportional gain k_p adjusts the magnitude of the proportional response. Proportional control is similar to the Rescorla-Wagner model. In particular,

$$P_t = k_p \cdot e_t \tag{B.1.0.1}$$

where e_t is the error at time t, the same as the prediction error (PE).

Integral (I) Control

The integral term focuses on the accumulation of past errors. It integrates the error over time and adjusts the system output to eliminate steady-state errors that may persist with proportional control alone. The integral gain k_I determines how aggressively the controller corrects cumulative errors:

$$I_t = k_I \cdot \int_0^t e_\tau \, d\tau \tag{B.1.0.2}$$

Derivative (D) Control

The derivative term predicts future errors based on the rate of change of the error. It provides a control action proportional to the rate of error change, allowing the system to react preemptively to changes in the error. The derivative gain k_D controls the influence of the derivative term:

$$D_t = k_D \cdot \frac{de_t}{dt} \tag{B.1.0.3}$$

Summed together, the PID controller gives a control signal u_t. In nuclear reactors, this control signal is what the PID controller sends to the plant to adjust its behavior and bring the output of the system closer to the desired setpoint:

$$u_t = P_t + I_t + D_t = k_p \cdot e_t + k_I \cdot \int_0^t e_\tau \, d\tau + k_D \cdot \frac{de_t}{dt} \tag{B.1.0.4}$$

Adaptation for Human Cognition

In the Ritz et al. (2018) paper, the authors adapted the PID controller to make it more consistent with human cognition by adding a memory persistence term λ in the integral term. This allowed the model to "forget" and assign less weight to historical prediction errors. The updated control signal is:

$$CB_t = \text{intercept} + k_p \cdot PE_t + k_I \cdot \sum_{n=1}^{t} \lambda^{t-n} \cdot PE_n + k_D \cdot (PE_t - PE_{t-1}) \tag{B.1.0.5}$$

Where:

- CB_t is the change in belief at time t, akin to the control signal.
- $PE_t = \text{Outcome}_t - \text{Belief}_t$ is the prediction error at time t.

Parameter Estimation

In the Ritz et al. (2018) original implementation, the memory persistence term was fixed across participants due to poor parameter recovery. Furthermore, the PI model, i.e., dropping the Derivative term, was concluded as the best fitting model. We thus followed the same practice and ran the PID controller as a linear regression with an intercept and prediction error and sum of all past prediction error weighted by memory persistence as predictors.

We fitted the PI model by block. We did a grid search on 500 possible values of memory persistence ranging from 0.5 and 1 with step size of 0.001 across four blocks. We the picked the value that gives the highest R-squared values by running PI regression with data for all four blocks. We then fitted the regression model with a fixed memory persistence term across all 4 blocks.

The model follows the same implementation and design as the one introduced in Chapter 1. The figure below illustrates the model's ability to replicate participant choices by comparing learning curves derived from participant data with those generated by the model (see Figure B.1).

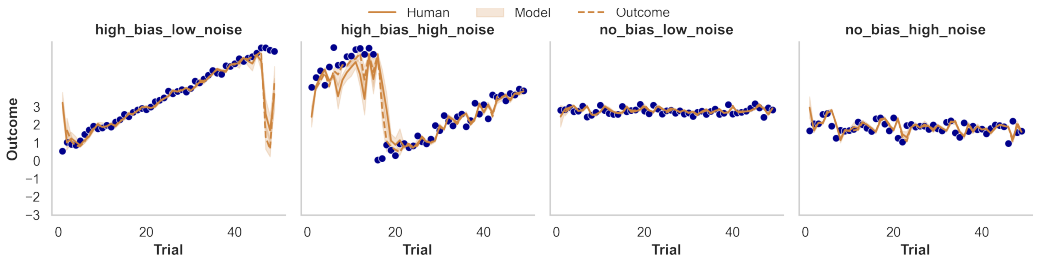


Figure B.1: Comparison of Model-Simulated and Participant Behavioral Data. The plot shown here overlays the choice behaviour of participants (brown solid line) with the behaviour generated using model-fitted parameters (brown dash line) and the output (blue dots). The model output was generated by running the PI controller for each participant using their individually fitted parameter values on the task trial sequences. As shown, the model accurately reflects participant behaviours during the task.

As the model uses historical PEs as part of its prediction, we ran a lagged regression to predict changes of belief with prediction error for the current trial, and prediction errors in the past 10 trials as predictors. We did not observe any evidence that participants used historical PEs when deciding on how much to update their belief since the beta weights for the past 10 trials would be different from 0. We thus excluded PID controller from further analysis (Figure B.2).

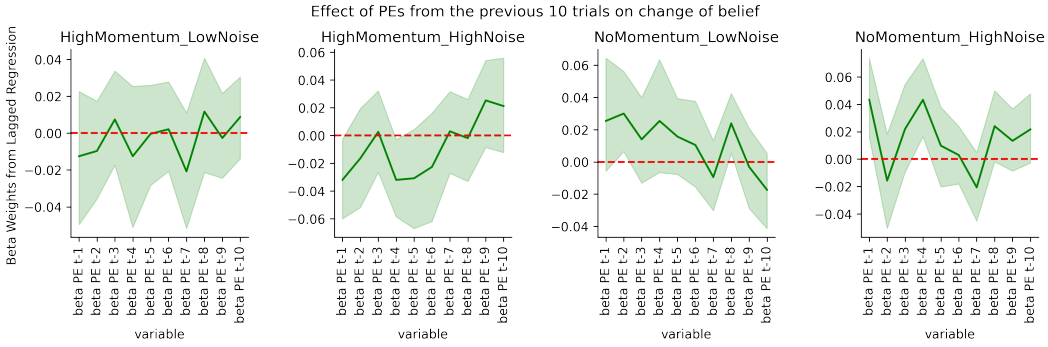


Figure B.2: Effect of PEs from the previous 10 trials on change of belief. Beta weights from a lagged regression model predicting belief updates using PEs from the current trial and the past 10 trials as predictors. Each panel illustrates the distribution of beta weights across participants for a specific lag, with the red dashed line indicating a beta weight of 0. The results show no significant deviation of the beta weights from 0, suggesting that historical PEs did not significantly influence participants' belief updates. Consequently, the PID controller was excluded from further analyses.

Model	Description	Input Data	Free Params	Key Differences
Simplified Trend Learning Model	Adapted Rescorla-Wagner model for circular data that learns from prediction errors and changes in positions.	Prediction and outcome of current trial; outcomes from previous trials	3	Tracks both trial-to-trial changes and prediction errors.
Alpha-Beta Filter	Estimates and predicts system states using position and velocity. Tracks current position and trends over time.	Prediction and outcome of current trial	5	Learns only from prediction errors; combines position (baseline) and velocity (trend learning) updates.
PID Controller	Linear regression model managing error using proportional, integral, and derivative terms.	Belief updates and prediction errors from past 10 trials	5	Accounts for proportional error, past error accumulation, and rate of change. Includes memory term λ for forgetting.

Table B.1: Summary and comparison of three models used in the study. This table outlines the key characteristics of the three models used in the thesis to study belief updating and trend learning. It includes descriptions, input data used, the number of estimated free parameters, and features that differentiate each model.

C
Word and Page Count

C.1 Word count

35449+714+4762 (142/87/860/65) File: Oxford_Thesis.tex

C.2 Page count

Total page count: 205

References

Aberg, K. C., Toren, I., & Paz, R. (2022). A neural and behavioral trade-off between value and uncertainty underlies exploratory decisions in normative anxiety. *Molecular psychiatry*, *27*(3), 1573–1587.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of experimental psychology: General*, *108*(4), 441.

Ashwood, Z. C., Roy, N. A., Stone, I. R., Laboratory, I. B., Urai, A. E., Churchland, A. K., Pouget, A., & Pillow, J. W. (2022). Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, *25*(2), 201–212.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu. Rev. Neurosci.*, *28*(1), 403–450.

Aylward, J., Valton, V., Ahn, W.-Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature human behaviour*, *3*(10), 1116–1123.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4-5), 407–419.

Beck, A. T., Rush, A. J., Shaw, B. F., Emery, G., DeRubeis, R. J., & Hollon, S. D. (1979). *Cognitive therapy of depression*. Guilford Publications.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature neuroscience*, *10*(9), 1214–1221.

Benedict, T., & Bordner, G. (2003). Synthesis of an optimal set of radar track-while-scan smoothing equations. *IRE Transactions on Automatic Control*, *7*(4), 27–32.

Boswell, J. F., Thompson-Hollands, J., Farchione, T. J., & Barlow, D. H. (2013). Intolerance of uncertainty: A common factor in the treatment of emotional disorders. *Journal of Clinical Psychology*, *69*(6), 630–645.

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

Bucci, D. J., Holland, P. C., & Gallagher, M. (1998). Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli. *Journal of Neuroscience*, *18*(19), 8038–8046.

Buhr, K., & Dugas, M. J. (2009). The role of fear of anxiety and intolerance of uncertainty in worry: An experimental manipulation. *Behaviour Research and Therapy*, *47*(3), 215–223.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences*, *107*(32), 14431–14436.

Cazettes, F., Reato, D., Morais, J. P., Renart, A., & Mainen, Z. F. (2021). Phasic activation of dorsal raphe serotonergic neurons increases pupil size. *Current Biology*, *31*(1), 192–197.

Choi, S. W., Schalet, B., Cook, K. F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the bdi-ii, ces-d, and phq-9 to promis depression. *Psychological assessment*, *26*(2), 513.

Cochran, A. L., & Cisler, J. M. (2019). A flexible and generalizable model of online latent-state learning. *PLoS computational biology*, *15*(9), e1007331.

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

Cook, J. L., Swart, J. C., Froböse, M. I., Diaconescu, A. O., Geurts, D. E., Den Ouden, H. E., & Cools, R. (2019). Catecholaminergic modulation of meta-learning. *elife*, *8*, e51439.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). Bold responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204.

Dayan, P., & Yu, A. J. (2002). Expected and unexpected uncertainty: Ach and ne in the neocortex. *Advances in neural information processing systems*, *15*.

Dayan, P., & Yu, A. J. (2005). Norepinephrine and neural interrupts. *Advances in neural information processing systems*, *18*.

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, *17*(4), 335–350.

de Berker, A., Rutledge, R., Mathys, C., Marshall, L., Cross, G., Dolan, R., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, *7*, 10996.

De Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature communications*, *7*(1), 10996.

Demyttenaere, K., & De Fruyt, J. (2003). Getting what you ask for: On the selectivity of depression rating scales. *Psychotherapy and psychosomatics*, *72*(2), 61–70.

Devauges, V., & Sara, S. J. (1990). Activation of the noradrenergic system facilitates an attentional shift in the rat. *Behavioural brain research*, *39*(1), 19–28.

Diaconescu, A., Mathys, C., Weber, L., Daunizeau, J., Kasper, L., Lomakina, E., & Stephan, K. (2014). Inferring on the intentions of others by hierarchical bayesian learning. *PLoS Computational Biology*, *10*(9), e1003810.

Diaconescu, A., Mathys, C., Weber, L., Daunizeau, J., Kasper, L., Lomakina, E., & Stephan, K. (2017). Aberrant perception of environmental volatility during social learning in emerging psychosis. *Nature Communications*, *8*, 14218.

Diederen, K. M., & Schultz, W. (2015). Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of neurophysiology*, *114*(3), 1628–1640.

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as representation of momentum. *Trends in cognitive sciences*, *20*(1), 15–24.

Fan, H., Gershman, S. J., & Phelps, E. A. (2023). Trait somatic anxiety is associated with reduced directed exploration and underestimation of uncertainty. *Nature Human Behaviour*, *7*(1), 102–113. https://doi.org/10.1038/s41562-022-01455-y

Gagne, C., Dayan, P., & Bishop, S. J. (2020). Impaired adaptation of learning to contingency volatility in internalizing psychopathology. *eLife*, *9*, e61387.

Gershman, S. J., & Niv, Y. (2012). Exploring a latent cause theory of classical conditioning. *Learning & behavior*, *40*, 255–268.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science*, *17*(9), 767–773.

Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *Journal of Neuroscience*, *34*(3), 698–704.

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

Hein, T., Herrojo Ruiz, M., & Dolan, R. (2023). Neural correlates of affective forecasting in individuals with high trait anxiety: Evidence from resting-state functional connectivity. *Brain Imaging and Behavior*, *17*(2), 456–468.

Huang, H.-Y., Thompson, W., & Paulus, M. P. (2017). Computational dysfunctions in anxiety: Failure to differentiate signal from noise. *Biological Psychiatry*, *82*(6), 440–446.

Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., den Ouden, H. E., & Stephan, K. E. (2013). Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron*, *80*(2), 519–530.

Jahn, C. I., Varazzani, C., Sallet, J., Walton, M. E., & Bouret, S. (2020). Noradrenergic but not dopaminergic neurons signal task state changes and predict reengagement after a failure. *Cerebral Cortex*, *30*(9), 4979–4994.

Jepma, M., Brown, S. B., Murphy, P. R., Koelewijn, S. C., de Vries, B., van den Maagdenberg, A. M., & Nieuwenhuis, S. (2018). Noradrenergic and cholinergic modulation of belief updating. *Journal of cognitive neuroscience*, *30*(12), 1803–1820.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the adaptive gain theory. *Journal of cognitive neuroscience*, *23*(7), 1587–1596.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221–234.

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nature neuroscience*, *9*(7), 940–947.

Lawson, R., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, *20*(9), 1293–1299.

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature neuroscience*, *14*(10), 1250–1252.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, *164*(1), 177–190.

Marshall, L., Mathys, C., Ruge, D., De Berker, A. O., Dayan, P., Stephan, K. E., & Bestmann, S. (2016). Pharmacological fingerprints of contextual uncertainty. *PLoS Biology*, *14*(11), e1002575.

Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the hierarchical gaussian filter. *Frontiers in human neuroscience*, *8*, 825.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of neuroscience*, *16*(5), 1936–1947.

Moran, R. J., Campo, P., Symmonds, M., Stephan, K. E., Dolan, R. J., & Friston, K. J. (2013). Free energy, precision and learning: The role of cholinergic neuromodulation. *Journal of Neuroscience*, *33*(19), 8227–8236.

Mukherjee, D., Filipowicz, A. L., Vo, K., Satterthwaite, T. D., & Kable, J. W. (2020). Reward and punishment reversal-learning in major depressive disorder. *Journal of Abnormal Psychology*, *129*(8), 810.

Müller, T., Pajkossy, P., Dombai, B., & Racsmány, M. (2019). How uncertain are you? disentangling expected and unexpected uncertainty in pupil-linked brain arousal during reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(6), 1236–1252.

Murphy, F., & Klein, R. (1998). The effects of nicotine on spatial and non-spatial expectancies in a covert orienting task. *Neuropsychologia*, *36*(11), 1103–1114.

Murphy, P. R., O'connell, R. G., O'sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with bold activity in human locus coeruleus. *Human brain mapping*, *35*(8), 4140–4154.

Nassar, M. R. (2024). Toward a computational role for locus coeruleus/norepinephrine arousal systems. *Current Opinion in Behavioral Sciences*, *59*, 101407.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience*, *15*(7), 1040–1046.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, *30*(37), 12366–12378.

Newman, L. A., Darling, J., & McGaughy, J. (2008). Atomoxetine reverses attentional deficits produced by noradrenergic deafferentation of medial prefrontal cortex. *Psychopharmacology*, *200*, 39–50.

Nour, M. M., Dahoun, T., Schwartenbeck, P., Adams, R. A., FitzGerald, T. H., Coello, C., Wall, M. B., Dolan, R. J., & Howes, O. D. (2018). Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. *Proceedings of the National Academy of Sciences*, *115*(43), E10167–E10176.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337.

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *110*(38), E3660–E3669.

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

Overmier, J. B., & Seligman, M. E. P. (1967). Effects of inescapable shock upon subsequent escape and avoidance responding. *Journal of Comparative and Physiological Psychology*, *63*(1), 28–33.

Pajkossy, P., Gesztesi, G., & Racsmány, M. (2023). How uncertain are you? disentangling expected and unexpected uncertainty in pupil-linked brain arousal during reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, *23*(3), 578–599.

Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS computational biology*, *7*(1), e1001048.

Payzan-LeNestour, E., Dunne, S., Bossaerts, P., & O'Doherty, J. P. (2013). The neural representation of unexpected uncertainty during value-based decision making. *Neuron*, *79*(1), 191–201.

Phillips, J. M., McAlonan, K., Robb, W. G., & Brown, V. J. (2000). Cholinergic neurotransmission influences covert orientation of visuospatial attention in the rat. *Psychopharmacology*, *150*, 112–116.

Pinto, L., Goard, M. J., Estandian, D., Xu, M., Kwan, A. C., Lee, S.-H., Harrison, T. C., Feng, G., & Dan, Y. (2013). Fast modulation of visual perception by basal forebrain cholinergic neurons. *Nature neuroscience*, *16*(12), 1857–1863.

Piray, P., & Daw, N. D. (2021). A model for learning based on the joint estimation of stochasticity and volatility. *Nature communications*, *12*(1), 6587.

Piray, P., & Daw, N. D. (2024). Computational processes of simultaneous learning of stochasticity and volatility in humans. *Nature communications*, *15*(1), 9073.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: Knowns and unknowns. *Nature Neuroscience*, *16*(9), 1170–1178.

Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in neuroscience*, *5*, 115.

Pulcu, E., & Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *Elife*, *6*, e27879.

Pulcu, E., & Browning, M. (2019). The misestimation of uncertainty in affective disorders. *Trends in cognitive sciences*, *23*(10), 865–875.

Pulcu, E., & Browning, M. (2025). Humans adapt rationally to approximate estimates of uncertainty. *eLife*.

Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, *1*(3), 385–401.

Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1994). Locus coeruleus activity in monkey: Phasic and tonic changes are associated with altered vigilance. *Brain research bulletin*, *35*(5-6), 607–616.

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature communications*, *7*(1), 13289.

Reimers, S., Donkin, C., & Le Pelley, M. E. (2018). Perceptions of randomness in binary sequences: Normative, heuristic, or both? *Cognition*, *172*, 11–25.

Ritz, H., Nassar, M. R., Frank, M. J., & Shenhav, A. (2018). A control theoretic model of adaptive learning in dynamic environments. *Journal of cognitive neuroscience*, *30*(10), 1405–1421.

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, *111*(33), 12252–12257.

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2015). Dopaminergic modulation of decision making and subjective well-being. *Journal of Neuroscience*, *35*(27), 9811–9822.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature reviews neuroscience*, *10*(3), 211–223.

Sara, S. J., & Segal, M. (1991). Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: Implications for cognition. *Progress in brain research*, *88*, 571–585.

Satti, M. H., Wille, K., Nassar, M. R., Cichy, R. M., Schuck, N. W., Dayan, P., & Bruckner, R. (2024). Absence of systematic effects of trait anxiety on learning under uncertainty. *CCN 2024: the 7th annual conference on Cognitive Computational Neuroscience*.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, *20*(10), 635–644.

Stewart, C., Burke, S., & Marrocco, R. (2001). Cholinergic modulation of covert attention in the rat. *Psychopharmacology*, *155*, 210–218.

Sullivan, G. M., Coplan, J. D., Kent, J. M., & Gorman, J. M. (1999). The noradrenergic system in pathological anxiety: A focus on panic with relevance to generalized anxiety and phobias. *Biological psychiatry*, *46*(9), 1205–1218.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning*, *3*, 9–44.

Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature communications*, *8*(1), 14637.

Vaghi, M. M., Luyckx, F., Sule, A., Fineberg, N. A., Robbins, T. W., & De Martino, B. (2017). Compulsivity reveals a novel dissociation between action and confidence. *Neuron*, *96*(2), 348–354.

Varazzani, C., San-Galli, A., Gilardeau, S., & Bouret, S. (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: A direct electrophysiological comparison in behaving monkeys. *Journal of Neuroscience*, *35*(20), 7866–7877.

Vossel, S., Bauer, M., Mathys, C., Adams, R., Dolan, R., Stephan, K., & Friston, K. (2014). Cholinergic stimulation enhances bayesian belief updating in the deployment of spatial attention. *Journal of Neuroscience*, *34*(47), 15735–15742.

Vossel, S., Geng, J., & Fink, G. (2015). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *The Neuroscientist*, *20*(2), 150–159.

Vossel, S., Mathys, C., Daunizeau, J., Bauer, M., Driver, J., Friston, K., & Stephan, K. (2014). Spatial attention, precision, and bayesian inference: A study of saccadic response speed. *Cerebral Cortex*, *24*(6), 1436–1450.

Vossel, S., Thiel, C. M., & Fink, G. R. (2008). Behavioral and neural effects of nicotine on visuospatial attentional reorienting in non-smoking subjects. *Neuropsychopharmacology*, *33*(4), 731–738.

Williams, J. J., & Griffiths, T. L. (2013). Why are people bad at detecting randomness? a statistical argument. *Journal of experimental psychology: learning, memory, and cognition*, *39*(5), 1473.

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2013). A mixture of delta-rules approximation to bayesian inference in change-point problems. *PLoS computational biology*, *9*(7), e1003150.

Witte, E., Davidson, M., & Marrocco, R. (1997). Effects of altering brain cholinergic activity on covert orienting of attention: Comparison of monkey and human performance. *Psychopharmacology*, *132*, 324–334.

Yoshida, W., & Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. *Neuron*, *50*(5), 781–789.

Yu, A. J., & Dayan, P. (2003). Acetylcholine, norepinephrine, and spatial attention. *Science Direct Working Paper No S1574-034X (04)*, 70278–1.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692.

Zika, O., Wiech, K., Reinecke, A., Browning, M., & Schuck, N. W. (2023). Trait anxiety is associated with hidden state inference during aversive reversal learning. *Nature Communications*, *14*(1), 4203.